



**The Characterisation of
Dynamin-Like Proteins in the
Gastric Pathogen
*Helicobacter pylori***

Emma Mary Dawson

A thesis submitted in fulfilment of the requirements for the degree
of Doctor of Philosophy

January 2018

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Emma Mary Dawson, certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This research is supported by an Australian Government Research Training Program Scholarship.

Signature of candidate:

Emma Dawson, January 2018

ACKNOWLEDGEMENTS

Firstly, I want to thank the ithree Institute in the Faculty of Science at the University of Technology Sydney for providing me with an innovative and ever-changing environment to complete my PhD in. Also, I want to thank the Australian Government for providing me with Australian Postgraduate Award funding throughout my PhD.

I need to thank my primary supervisor Assoc. Prof Iain Duggin for providing me with the opportunity to complete my PhD and for taking me on as an honours student 5 years ago. What a journey we have been on together. To my co-supervisor, Dr Matthew Padula, you provided me with more than just proteomics knowledge and thesis proof reading; you gave me the motivation and the drive to keep going when I was ready to give up on so many occasions. You are a wonderful mentor.

To my right hand, Nina, when you joined our group as an honours student I knew I was the luckiest girl in the world to work alongside you. You drove the project forward and inspired me every day to keep persevering when things weren't working at the time. To the other Duggin Lab-bers past and present, especially James, Greg, Dan and Tamika, thank you so much for your daily support, your constant input and your friendship, I will never forget it. To my undergraduate students, Naomi and Sam, it was a pleasure to have you help me in the lab and in you I have found some wonderful, once in a lifetime, kind of friends.

Over the course of my 10 years at UTS I have been so blessed to have found so many friends I know I will have for life who I want to thank, but none more than the bestest friend a girl could ask for. Samira, you inspire me with all that you do and all that you are. You are a wonderful, talented scientist and on too many occasion you have picked me up, dusted me off and kept me moving forward. Thank you for all that you continue to do for me, but for nothing more than for making me laugh daily. Special mention to Iain Berry, you have instilled in me a little voice that will forever tell me to "say no!" and thank you for reading my thesis!

Ben, there are no words to thank you for what you have done for me over the past 6 months. From reading my thesis, to holding me through the death of my Nan, and loving me unconditionally through it all. I can't explain what it all meant to me. In you I have found all I have ever been searching for. I love you more than anything and can't wait for our life together to truly begin.

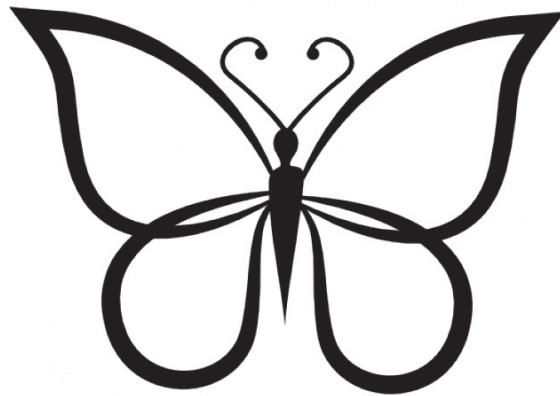
My life-long friends, Alicia, Laura and Elle, you have put up with my constant absence for the past 4 years and, yet your friendship never wavered. I am so grateful to have you in my life and honestly wouldn't have finished without your support. Thank you for the countless conversations, the many nights at the wine bar and the endless love.

To my physie family, through the grief and pain of the loss of a loved one we have become closer and I can't imagine life without you all. Bree, Jono and your beautiful family, you have taught me so much about loving life and taking it all as it comes. You are so wonderful. To Jo Willows and your wonderful family, your weekly check-ins and your support during the writing of my thesis was appreciated more than you realise. Thank you so much for reading it!

Finally, to the people who have put up with so much and their constant, undeniable love and support are the two things that give me a reason to wake up every day. My family, Mum, Dad and Matt, we have been through so much and have come out the other side, closer and stronger. I am the luckiest girl in the world to be a part of such a wonderful family. There are no words to describe how thankful I am for you, but from the bottom of my heart I say thank you.

Throughout this PhD journey, I have learnt that life isn't necessarily about what you do but rather, it is about the people you do it with and you don't know how much time you have with those people. So, I want to dedicate this thesis to my loved ones I have lost during my PhD who supported me, gave me the strength to keep going and loved me beyond anything else.

**Evie and Nan, I did this for you and in the end, I hope I
made you proud**



PUBLICATIONS

1. Duggin, I. G., Aylett, C. H., Walsh, J. C., Michie, K. A., Wang, Q., Turnbull, L., **Dawson, E. M.**, Harry, E. J., Whitchurch, C. B., Amos, L. A. and Lowe, J. (2015). "CetZ tubulin-like proteins control archaeal cell shape." *Nature*, 519 (7543): 362-365.
2. **Dawson, E.M.** (2016) "Dynamin-Like Proteins contribute to membrane integrity and acid tolerance in *Helicobacter pylori*". *ASM Syntrophy*, 17(7): 7.

CONFERENCE PROCEEDINGS

1. **2012 Oral** - 29th RNSH/UTS/USYD Scientific Research Meeting, Sydney
2. **2013 Oral** - 13th East Coast Bacillus, Sydney
3. **2013 Oral** - 30th RNSH/UTS/USYD Scientific Research Meeting, Sydney
4. **2015 Poster** - 32th RNSH/UTS/USYD Scientific Research Meeting, Sydney
5. **2015 Poster** - BacPath 13, Melbourne
6. **2015 Poster** - 40th Lorne Protein Structure and Function, Victoria
7. **2016 Oral** - Gordon Research Seminar, Vermont, USA
8. **2016 Poster** - Gordon Research Conference, Vermont, USA
9. **2016 Oral** - Becton Dickinson Award
 - a. *Finalist*
10. **2016 Oral** – Sydney Protein Group Thompson Prize, Sydney
 - a. *Awarded the Thompson Prize*
11. **2017 Oral** – BacPath 14, Adelaide

ABBREVIATIONS

Ab (α-)	Antibody
AMP	Ampicillin
BDLP1	Bacterial Dynamin Like Protein1
BLAST	Basic local alignment search tool
BSA	Bovine Serum Albumin
BSE	Basal Signalling Element
bp	Base Pair
DDA	Data dependent Acquisition
Δ	Deleted
Dlp1a, 1b, 1, 2	Dynamin-Like Protein 1a, 1b, 1, 2
DIA	Data independent Acquisition
<i>dlp</i>	Dynamin like protein
DNA	Deoxyribonucleic acid
DynA	Dynamin A
ETEC	Enterotoxigenic <i>Escherichia coli</i>
FITC	Fluorescein isothiocyanate
x g	Centrifugal force (unit x gravitational field)
g	Gram
GED	GTPase effector domain
GFP	Green fluorescent protein
GTP	Guanosine triphosphate
HP	<i>Helicobacter pylori</i>
h	Hour (s)
IFM	Immunofluorescence Microscopy
IgG	Immunoglobulin G
IniA	Isoniazid-induced gene A
Kan	Kanamycin
Kb	Kilo basepair (s)
kDa	Kilodalton (s)
L	Litre (s)
LB	Loading Buffer
LeoCBA	Labile enterotoxin output A, B, C
LT	Heat-labile Enterotoxin
M	Molarity
mA	Milliamp (s)
min	Minute (s)
mL	Millilitre (s)

ms	Milliseconds
MQW	Milli-Q water
MS	Mass Spectrometry
Mut	Mutant
MW	Molecular Weight
n	Nano (10^{-9})
OD	Optical Density
OMP	Outer Membrane Protein (s)
OMV	Outer membrane Vesicle (s)
ORF	Opening reading frame
p	Pico (10^{-12})
PAI	Pathogenicity Island
%	Percent
PBS	Phosphate Buffered Saline
PH	Pleckstrin Homology Domain
PCR	Polymerase Chain Reaction
PRD	Proline-Rich Domain
(m)RNA	(messenger) Ribonucleic acid
RT	Room Temperature
rpm	Revolutions per minute
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec	Seconds
SWATH-MS	Sequential Windowed Acquisition of All Theoretical Fragment Ion Mass Spectra Mass Spectrometry
TBS	Tris Buffered Saline
TEMED	N,N,N',N'-tetramethylethylenediamine
Tris	Tris(hydroxymethyl)methylamine
w	Weight
WCL	Whole Cell Lysate
WT	Wild Type
v	Volume
μ-	Micro (10^{-6})

TABLE OF CONTENTS

Acknowledgements	iii
Publications	vi
Conference Proceedings	vi
Abbreviations	vii
Table of Contents	ix
List of Figures	xv
List of Tables	xviii
List of Supplementary Data	xix
Abstract	xx
General Introduction	1
1.1 Membrane Dynamics	2
1.1.1 Endocytosis: An Essential Membrane Remodelling Event	4
1.2 Dynamin: An Integral GTPase Protein involved in Membrane Dynamics	6
1.3 The Dynamin Superfamily	11
1.3.1 Classical Dynamin	11
1.3.2 Functional Diversity of the Dynamin-Related Proteins	12
1.3.2.1 Plant Dynamins	14
1.3.2.2 Mx Proteins	15
1.3.2.3 Mitochondrial associated Dynamins	15
1.3.2.4 Distantly related Dynamins	16
1.4 The Diverse Bacterial Membrane	18
1.5 Bacterial Dynamin-Like Proteins	21
1.5.1 The First Characterised Bacterial Dynamin – BDLP1	22
1.5.2 Diversity of Bacterial Dynamins	25
1.5.3 A Unique Bacterial Dynamin in <i>Bacillus subtilis</i> -DynA	28
1.5.4 Enterotoxigenic <i>E. coli</i> Bacterial Dynamins - LeoCBA	33

1.6 <i>Helicobacter pylori</i> As A Pathogen	40
1.6.1 <i>H. pylori</i> Can Cause Significant Disease in Humans	42
1.6.2 How Does <i>H. pylori</i> Deal With the Acidic Environment of the Stomach?	44
1.6.3 Key Features in the Pathogenesis of <i>H. pylori</i> Infection	47
1.7 Hypothesis And Aims Of This Thesis	51
General Methodology	54
2.1 Bacterial Strains and Plasmids	55
2.2 Bacterial Growth Media	56
2.3 Chemicals, Reagents and Solutions	57
2.4 Bacterial Growth Conditions	59
2.4.1 Normal growth conditions of <i>H. pylori</i>	59
2.4.2 Normal Growth conditions of <i>E. coli</i>	59
2.4.3 Preparation and Transformation of <i>E. coli</i>	59
2.5 General Methods Used in this Work	60
2.5.1 SDS-PAGE Analysis	60
2.5.2 DNA Electrophoresis	61
2.5.3 Preparation of gDNA and plasmid DNA	62
2.5.4 Polymerase Chain Reaction (PCR)	62
2.5.5 Sanger Sequencing	63
2.5.6 General Cloning Protocols	63
Characterisation of Dynamin-like Proteins (DLPs) in <i>Helicobacter pylori</i>	64
Preface	65
3.1 Introduction	66
3.2 Material and Methods	69
3.2.1 Collection of Sequences of the Leo Type Dynamins and the Dynamin Superfamily	69
3.2.2 Phylogenetic Analysis	69
3.2.3 <i>H. pylori</i> screen for Leo-Type DLPs and Phylogenetic analysis	69
3.2.4 DLP Structural Features	70
3.2.5 RNA Purification and Reverse Transcription (RT)-PCR	70

3.3 Results	72
3.3.1 Conservation of the Dynamin Superfamily	72
3.3.1.1 Dynamin conservation in Bacteria	72
3.3.1.2 Bacterial dynamins in the context of the greater superfamily	75
3.3.1.3 <i>H. pylori</i> genomic screen for DLPs	78
3.3.2 Dynamin-Like proteins in <i>Helicobacter pylori</i>	81
3.3.2.1 <i>H. pylori</i> phylogenetic analysis of the DLPs	81
3.3.2.2 <i>dlp</i> operon in <i>H. pylori</i>	83
3.3.2.3 Predicted DLP structural features	87
3.3.3 Expression of <i>dlps</i> in <i>Helicobacter pylori</i>	92
3.3.3.1 RT-PCR analysis of <i>dlp</i> expression	92
3.3.3.2 A DNA Sequence Anomaly in the Overlap Between <i>dlp1a</i> and <i>dlp1b</i>	95
3.4 Discussion	99
Investigation into the Biochemical properties of Dlp1 and Dlp2	106
Preface	107
4.1 Introduction	108
4.2 Material and Methods	111
4.2.1 pHis17- <i>dlp</i> construction	111
4.2.2 Over Expression and Purification of Dlp2-6H	112
4.2.2.1 His Trap Purification	112
4.2.2.2 Ion Exchange Purification	113
4.2.2.3 Gel Filtration Purification	114
4.2.3 Over Expression and Purification of Dlp1a-6H	114
4.2.3.1 Inclusion Body Purification	114
4.2.4 Production of Dlp1a-6H and Dlp2-6H antibodies	115
4.2.5 Testing of Antibodies on <i>H. pylori</i> whole cell lysate (Western Blot)	116
4.2.6 Subcellular fractionation	117
4.2.7 Immuno-Fluorescence Microscopy (IFM)	118
4.2.8 Liposome Binding Assay	119
4.2.9 Co-Pelleting and Co-Sedimentation Assay	119

4.2.10 Bacterial Two Hybrid System	120
4.2.11 Gel Filtration complex Studies	122
4.3 Results	123
4.3.1 Purification of Dlp2	123
4.3.2 Purification of Dlp1a	132
4.3.3 <i>In Vivo</i> Expression Of Dlp1a and Dlp2	137
4.3.4 Subcellular Localisation of Dynamin-Like Proteins in <i>H. pylori</i>	141
4.3.5 Lipid Interaction Studies of <i>H. pylori</i> DLPs	148
4.3.6 Investigating potential interactions of the DLPs	151
4.4 Discussion	158
What is the function of Dynamin-like Proteins (DLPs) in <i>Helicobacter pylori</i>?	165
Preface	166
5.1 Introduction	167
5.2 Materials and Methods	170
5.2.1 Plasmid Construction	170
5.2.2 <i>H. pylori</i> Transformation	173
5.2.3 Whole Genome Sequencing and Data Analysis of B128.7.13 strains	174
5.2.4 Haemolysis of Blood Agar	174
5.2.5 Membrane integrity staining	175
5.2.6 Growth Curve in neutral conditions	176
5.2.7 Exposure to Acid	176
5.2.8 One point measurement of Growth in Acid	177
5.2.9 Growth Curve in Acid	177
5.3 Results	178
5.3.1 Creation of B128.7.13 Δ <i>dlp</i> strains	178
5.3.2 Creation of a complemented B128.7.13 Δ <i>dlp</i> Strain	182
5.3.3 Confirmation of B128.7.13 Δ <i>dlp</i> and B128.7.13 Δ :: <i>dlp</i> Strain Sequences	184
5.3.4 Δ <i>dlp</i> causes an Increase in Blood Agar Haemolysis	189
5.3.5 Haemolysis is due to a Membrane Defect in Δ <i>dlp H. pylori</i>	191
5.3.6 <i>H. pylori</i> Δ <i>dlp</i> is Sensitive to Acid Exposure	195
5.3.7 <i>H. pylori</i> Δ <i>dlp</i> has a Minimal Growth Defect in Neutral Conditions	200

5.3.8 <i>H. pylori</i> Δdlp is unable to grow in an Acidic Environment	201
5.4 Discussion	208
Exploring the Global proteomic changes to B128.7.13 in response to acid	217
Preface	218
6.1 Introduction	219
6.2 Materials and Methods	222
6.2.1 Preparation of samples for SWATH	222
6.2.2 Data Dependent Acquisition (LC/MS/MS) for Ion Library Construction	223
6.2.3 Data Independent Acquisition (SWATH) and Data Analysis	224
6.3 Results	226
6.3.1 Ion Library Construction	226
6.3.2 Principle Components Analysis of SWATH-MS Data	229
6.3.3 Summary of Significant Protein Changes during SWATH acquisition	231
6.3.4 Proteome Changes to B128.7.13 WT After an acidic Shock (pH5)	233
6.3.5 Proteome Changes to B128.7.13 Δdlp After an acidic Shock (pH5)	237
6.3.6 Proteome Changes under normal conditions (pH7) to B128.7.13 Δdlp	240
6.3.7 Proteome Changes After an acidic Shock (pH5) To B128.7.13 Δdlp	245
6.4 Discussion	250
6.4.1 Proteomic differences between B128.7.13 WT and B128.7.13 Δdlp	254
6.4.2 Proteomic changes in B128.7.13 WT upon acid shock	260
6.4.3 The Shared Acid Response	262
6.4.4 The B128.7.13 Δdlp membrane is significantly altered in response to acid shock	262
General Discussion	270
7.1 The Ins and Outs of Dynamin-Like Proteins in <i>H. pylori</i>	272
7.1.1 Aim I: To investigate the conservation of dynamin-like proteins in <i>H. pylori</i>	272
7.1.2 Aim II: To characterise the expression and biochemical features, such as localisation, lipid binding and interactions of dynamin-like proteins from <i>H. pylori</i>	273
7.1.3 Aim III: To explore the function of dynamin-like proteins in <i>H. pylori</i>	276

7.2 What does this mean for furthering our understanding of <i>H. pylori</i> ?	279
7.3 Concluding Remarks	280
Additional Data (Appendix)	282
Chapter III	283
Chapter V	283
Chapter VI	283
References	284

LIST OF FIGURES

Figure 1.1 Important Membrane Dynamics	3
Figure 1.2 Different Pathways of Endocytosis in Eukaryotic Cells	5
Figure 1.3 The GTPase Dynamin Dimer Structure	8
Figure 1.4 Two Proposed Models of Coupled Dynamin GTPase Activity and Lipid Binding	10
Figure 1.5 Dynamin Superfamily and their Conserved Domain Architecture	13
Figure 1.6 Diverse Functions of the Dynamin Superfamily	17
Figure 1.7 Gram-positive and Gram-negative Bacterial Membranes	20
Figure 1.8 BDLP1 has the Conserved Dynamin Structural Features	24
Figure 1.9 DynA is a Fused Bacterial Dynamin	29
Figure 1.10 DynA plays a role in the Maintenance and Integrity of the Cell Wall	32
Figure 1.11 LeoA is a Bacterial Dynamin involved in Vesicle Formation in ETEC	36
Figure 1.12 <i>leoCBA</i> operon in ETEC Encodes Two Bacterial Dynamin-like Proteins	39
Figure 1.13 <i>Helicobacter pylori</i> causes Infection in the Stomach	41
Figure 1.14 <i>H. pylori</i> Infection can Lead to Many Diseases in Humans	43
Figure 1.15 Acid Adaption Mechanisms of <i>H. pylori</i>	46
Figure 1.16 <i>H. pylori</i> Pathogenesis is Multifactorial	50
Figure 3.1 Phylogenetic Analysis of Leo-Type Dynamins in Bacteria	74
Figure 3.2 Phylogenetic Analysis of the Dynamin Superfamily	77
Figure 3.3 Genomic Screen of <i>H. pylori</i> for Leo-Type Dynamins	80
Figure 3.4 <i>H. pylori</i> DLPs form a Distinct Two Branch Phylogenetic Tree	82
Figure 3.5 <i>H. pylori</i> <i>dlp</i> Operon Structure in Selected Strains.	84
Figure 3.6 <i>H. pylori</i> 26695 Dlp Operon Features	86
Figure 3.7 Structural Analysis of Dlp2 using Phyre2 Prediction	89
Figure 3.8 Structural Analysis of Dlp1 using Phyre2 Prediction	91
Figure 3.9 RT-PCR analysis of <i>dlp</i> Operon Expression in 26695	94
Figure 3.10 Overlapping Region between <i>dlp1a</i> and <i>dlp1b</i> Sequence Anomaly	96
Figure 3.11 Sequencing Evidence of both <i>dlp1a/1b</i> and <i>dlp1</i>	97

Figure 4.1 pHis17- Plasmid Map	111
Figure 4.2 Dlp2-6H Over Expression	124
Figure 4.3 HisTrap Purification of Dlp2-6H	126
Figure 4.4 Ion Exchange Purification of Dlp2-6H	127
Figure 4.5 Further HisTrap Purification of Dlp2-6H	128
Figure 4.6 Final Purified Dlp2-6H	129
Figure 4.7 Possible Dlp2 Dimer Interaction	131
Figure 4.8 Dlp1a-6H Over Expression	132
Figure 4.9 Final Inclusion Body Containing Dlp1a-6H	133
Figure 4.10 Refolding and Elution of Dlp1a-6H	134
Figure 4.11 BSA Standard Curve for Estimation of Dlp1a-6H Concentration	136
Figure 4.12 <i>In vivo</i> Expression of Dlp2 in both 26695 and B128.7.13	138
Figure 4.13 <i>In vivo</i> Expression of Dlp1a and Dlp1 in both 26695 and B128.7.13	140
Figure 4.14 <i>H. pylori</i> Whole Cell Fractionation and Localisation of Dlp1 and Dlp2	142
Figure 4.15 IFM of <i>H. pylori</i> 26695 for Localisation of Dlp1 and Dlp2	144
Figure 4.16 IFM of <i>H. pylori</i> B128.7.13 for <i>in vivo</i> Localisation of Dlp2	146
Figure 4.17 IFM of <i>H. pylori</i> B128.7.13 for <i>in vivo</i> Localisation of Dlp1a	147
Figure 4.18 Liposome Deformations in the Presence of Dlp2-6H	149
Figure 4.19 Co-Sedimentation of Liposomes with Dlp2-6H and Co-pelleting of Dlp2-6H only	150
Figure 4.20 X-Gal Plate Results for BacTH Assay	153
Figure 4.21 Quantification of Potential Dlp Interactions	155
Figure 4.22 Preliminary Interaction between Dlp1a and Dlp2	157
Figure 5.1 General Plasmid Map for all Deletion Constructs	172
Figure 5.2 Overview of the Planned <i>H. pylori</i> Dynamin Deletion Strains	179
Figure 5.3 Confirmation of Cassette Insertion into B128.7.13	181
Figure 5.4 Confirmation of the Complementation of B128.713 Δ dlp	183
Figure 5.5 Whole Genome Alignment of B128.7.13, Δ dlp and Δ ::dlp Strains	185
Figure 5.6 dlp Operon in B128.7.13 WT, Δ dlp and Δ ::dlp Strains	188
Figure 5.7 Haemolysis of Blood Agar in the Presence of Δ dlp Strain	190
Figure 5.8 Δ dlp shows Evidence of a Fragile Membrane	192

Figure 5.9 Quantification of Δdlp Membrane Permeability	194
Figure 5.10 Δdlp is Unable to Overcome an Acid Exposure	197
Figure 5.11 Additional Urea can Rescue Δdlp in Acid	199
Figure 5.12 Δdlp Shows No Growth Defect in Neutral pH Conditions	200
Figure 5.13 Δdlp is Unable to Grow in Acidic pH Conditions	202
Figure 5.14 Additional Urea can Rescue Δdlp Growth in Acid	204
Figure 5.15 Acidic Growth Curve and Corresponding pH Neutralisation	206
Figure 6.1 Coverage of the Predicted <i>H. pylori</i> B128.7.13 Proteome	227
Figure 6.2 Variance in the SWATH-MS Sample Groups	230
Figure 6.3 Summary of Significant Hits of SWATH-MS.	232
Figure 6.4 SWATH-MS Analysis of Proteome changes to B128.7.13 WT After an Acidic Shock	234
Figure 6.5 SWATH-MS Analysis of Proteome changes to B128.7.13 Δdlp After an Acidic Shock	238
Figure 6.6 SWATH-MS Analysis of Proteome changes to B128.7.13 Δdlp Under Neutral Conditions	244
Figure 6.7 SWATH-MS Analysis of Proteome changes to B128.7.13 Δdlp Under Acidic Conditions	249
Figure 6.8 The Ion Intensity of PROKKA_01140 – Ferritin	257
Figure 6.9 Acid Tolerance Transcriptional Regulation by the ArsS/R system	259
Figure 6.10 The Ion Intensity for OMPs PROKKA_00299 and PROKKA_01348	265
Figure 6.11 <i>H. pylori</i> Flagella Machinery	268

LIST OF TABLES

Table 1.1 Bacterial Dynamins that have been Identified to Date	27
Table 2.1 General Bacterial Strains used in this Work	55
Table 2.2 General Plasmids used in this Work	55
Table 2.3 Bacterial Growth Media	56
Table 2.4 Commonly Used Chemicals in this Work	57
Table 2.5 Aqueous Buffers and Solutions used in this Work	58
Table 2.6 SDS-PAGE Components for ‘in house’ Gel	61
Table 3.1 Primers used in RT-PCR	71
Table 4.1 Primers Used for Amplification of Dlps for pHis17 Plasmid construction	112
Table 4.2 Antibodies Used throughout this Work	116
Table 4.3 Primers Used for Amplification of Dlps for BacTH Plasmid construction	120
Table 4.4 Testing Panel for BacTH Assay	121
Table 5.1 Primers Used for Amplification of Flanking Region of <i>dlp</i> for pBSII Plasmid construction	171
Table 5.2 Summary of FlagStat Analysis of the Alignment of B128.7.13 MiSeq Reads	186
Table 6.1 Significant Proteome Changes to B128.7.13 WT After an Acidic Shock	233
Table 6.2 Significant Proteome Changes to B128.7.13 Δ <i>dlp</i> After an Acidic Shock	237
Table 6.3 Significant Proteome Changes to B128.7.13 Δ <i>dlp</i> Under Neutral Conditions	241
Table 6.4 Significant Proteome Changes to B128.7.13 Δ <i>dlp</i> Under Acidic Conditions	246

LIST OF SUPPLEMENTARY DATA

(On the USB provided)

- S1** *H. pylori* (1997) *dlp1a*, *dlp1b* and *dlp2* as well as *H. pylori* (2013) *dlp1* nucleotide sequences
- S2** Conservation screen for Dynamin-Like proteins in *H. pylori*
- S3** Transmembrane predictions of selected *H. pylori* DlpS
- S4** Completely Labelled *H. pylori* dynamin (back-to-back) phylogenetic tree
- S5** *H. pylori* B128.7.13 predicted proteome annotated by PROKKA
- S6** Membrane Integrity Individual Filter Images from the DeltaVision Elite
- S7a** B128.7.13 SWATH Ion Library
- S7b** WT pH7 verses pH5 all quantified proteins
- S7c** WT pH7 verses pH5 significantly (< 0.05) quantified proteins
- S7d** Mut (Δdlp) pH7 verses pH5 all quantified proteins
- S7e** Mut (Δdlp) pH7 verses pH5 significantly (< 0.05) quantified proteins
- S7f** WT verses Mut (Δdlp) at pH7 all quantified proteins
- S7g** WT verses Mut (Δdlp) at pH7 significantly (< 0.05) quantified proteins
- S7h** WT verses Mut (Δdlp) at pH5 all quantified proteins
- S7i** WT verses Mut (Δdlp) at pH5 significantly (< 0.05) quantified proteins
- S7j** WT verses Mut (Δdlp) pH5 Quantified Flagella Proteins

ABSTRACT

In eukaryotic cells, it has been well established that large GTPases of the dynamin superfamily are important drivers of membrane curvature and constriction during membrane fusion and fission processes such as clathrin-coated endocytosis (1, 2). More recently, dynamins have been identified in a number of bacteria including *Bacillus subtilis*, *Nostoc punctiforme* and *Escherichia coli* (3-5). However, despite their putative identification in over 1500 bacterial species to date, minimal understanding of their biological function has been established in bacteria (6). One such bacterial dynamin, a previously characterised enterotoxigenic *E. coli* virulence factor, LeoA has been shown to be involved in outer membrane vesicle formation for the release of heat-labile LT toxin (5, 7). Upstream of LeoA there were two other dynamins identified, LeoC and LeoB, encoding a full-length dynamin only when expressed together. This is a seemingly split ORF, a first for dynamin (5). The closest homologs to these genes are found in the gastric carcinogenic pathogen *Helicobacter pylori* and we have termed them *dlp1a*, *dlp1b* and *dlp2* respective to their LeoCBA homologs. Therefore, this thesis endeavoured to investigate for the first time the presence of dynamin-like proteins (DLPs) in *H. pylori*, an important and surprisingly understudied human pathogen that is known to be dependent on membrane dynamics for virulence (8-10). It is also a pathogen that survives and thrives in one of the harshest environments, the acidic environment of the human stomach, where membrane integrity is essential to its protection.

The investigation into the presence of DLPs in *H. pylori* undertaken in this thesis was achieved through a diverse series of aims. Firstly, a bioinformatic survey undertaken to determine the conservation of the DLP genes in the *H. pylori* genome highlighted conservation greater than 90% in *H. pylori*, unprecedented for any known bacterial dynamins. For the first time, *in vivo* expression of *dlps* was shown in *H. pylori* using detection of the mRNA transcripts as well as the protein products via Western blotting. Interestingly, a single-nucleotide indel was identified in a region of overlap between *dlp1a* and *dlp1b* that differs between sub-isolates of *H. pylori*, leading to either the split or the joined forms of *dlp1* (the full-length gene). Furthermore, the analyses of *dlp1* and *dlp2* mRNA and their protein

products were consistent with this genotypic variation and might represent a novel regulation mechanism.

A preliminary investigation into the biochemical features of the DLPs such as localisation, self-interaction and lipid binding capabilities was commenced. In strains containing the fused *dlp1*, Western blotting and immuno-fluorescence microscopy suggested that Dlp1 is membrane-associated, whereas Dlp2 is approximately equally distributed between the cytosol and membrane. Also, initial studies suggested that Dlp2 may dimerise, interact with liposomes, as well as showing potential for interaction with its operon partner Dlp1a; all fundamental dynamin features. The gene arrangement and presence of a predicted transmembrane helix in Dlp1, similar to the membrane-binding domain of eukaryotic dynamin, suggested that these proteins might act as a hetero-complex involved in bacterial membrane dynamics.

Finally, the previously uncharted function of DLPs in *H. pylori* was explored using both direct and global approaches. Knockout of the *dlp* operon in *H. pylori* caused a high proportion of cells to develop a compromised cell membrane in dye-penetration assays, and minimal capacity to overcome acidic exposure as well as minimal growth in acidic conditions. Utilising a global proteome approach, SWATH-MS, one of the most comprehensive quantitative proteomics study to date of the acid response in wild-type and knock-out strains was achieved. This analysis highlighted the potential role that dynamins may have in acid adaptation, with proteins involved in processes such as motility and cytoplasmic neutralisation detected in higher abundance in the absence of dynamin. Taken together, this suggests that *H. pylori*, with a compromised membrane in the absence of dynamin, is attempting to overcome the higher influx of acid into the cell by “escaping” to conditions that are more neutral.

Overall, this thesis provides the first evidence that this important human pathogen, *H. pylori* utilises membrane remodelling events in some capacity driven, by dynamins, to maintain membrane integrity. A feature that is necessary to ensure continual pathogenesis in its niche environment the stomach of over half the world’s population.

CHAPTER I

GENERAL INTRODUCTION

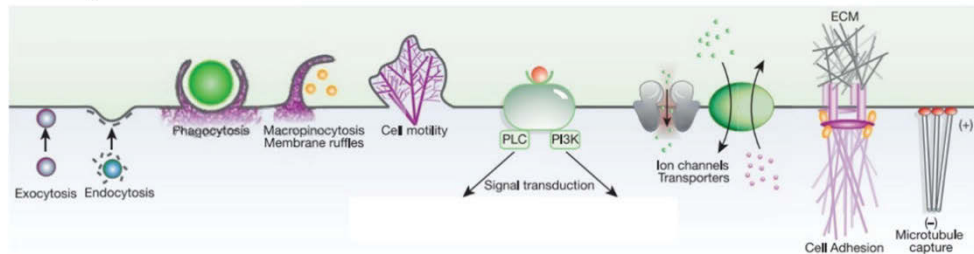
1.1 MEMBRANE DYNAMICS

Biological membranes are highly dynamic, semi-permeable barriers that encompass cells from all domains of life, and are essential for providing protection to the intracellular environment from the extracellular milieu. The membrane comprises the major boundary outlining the cytoplasm and controls the transport or diffusion of a myriad of small molecules and macromolecules between the intra- and extracellular spaces. Within the eukaryotic cell, organelles, such as the mitochondria, Golgi apparatus and the nucleus, are also membrane-bound to organise and safeguard intracellular trafficking and other fundamental biological processes.

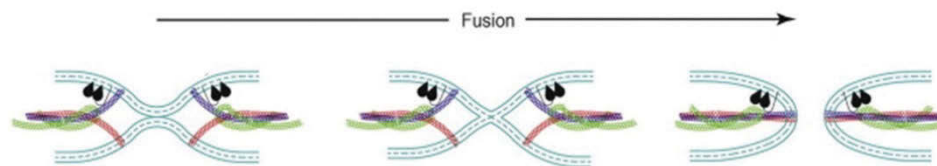
The constant import and export of molecules is fundamental to maintaining intracellular homeostasis, and for this, the ability to dynamically remodel the lipid bilayer membrane is crucial. Transport proteins embedded within the membrane facilitate the passive and active movement of ions, small molecules, or macromolecules across the membrane. For the majority of ions, metals and organic nutrients this is done through the use of specific channels and ATP-driven pumps that regulate their uptake and influence cellular osmotic balance. The transport of large macromolecules, and even other organisms including bacteria, and viruses across the membrane occurs via two of the most well-studied membrane remodelling events, exocytosis and endocytosis (*Figure 1.1A*).

Endocytosis is the process whereby contents of the plasma membrane and extracellular fluids are delivered to the intracellular milieu via membrane-bound vesicles, with exocytosis being essentially the reverse of this process (11, 12). These two processes require the same two basic underlying mechanisms; membrane fusion (*Figure 1.1B*) and fission (*Figure 1.1C*). Membrane fusion describes the merging of the hydrophobic cores of two initially distinct lipid bilayers to form one interconnected structure (13), whereas membrane fission is the process in which a cellular membrane that is initially contiguous is split into two distinct membranes (14).

(A) Membrane Dynamic Events



(B) Protein Mediated Membrane Fusion



(C) Protein Mediated Membrane Fission

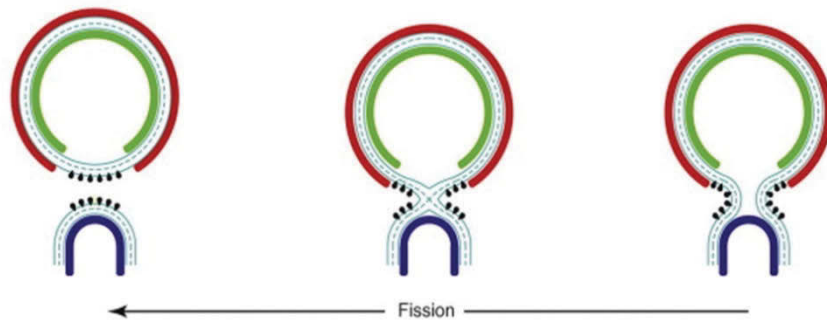


Figure 1.1 Important Membrane Dynamics

(A) There are a number of important events that occur at the membrane including endo- and exocytosis, phagocytosis, membrane ruffling, cell motility, signal transduction, transport of ions and cell adhesion. Each of these events requires extensive remodelling of the membrane. Adapted from (15). (B) Membrane fusion occurs when two initially separate and apposed membranes merge into one by undergoing a sequence of intermediate transformations. (C) Membrane fission is the division of an initially continuous membrane into two separate ones through the narrowing or 'pinching' down on the neck of the membrane. Adapted from (16).

1.1.1 ENDOCYTOSIS: AN ESSENTIAL MEMBRANE REMODELLING EVENT

Endocytosis is the process by which the membrane invaginates to translocate and internalise a molecule/s from the external milieu or cell surface into the cell cytoplasm, via an intracellular vesicle. Endocytosis can be carried out by a number of independent pathways, seen in *Figure 1.2A*. These include Clathrin-dependent, Caveolin-dependent, or Clathrin- and Caveolin-independent endocytosis; of these the most well studied and most relevant to this work is Clathrin-dependent endocytosis (*Figure 1.2A, red box*) (11, 17).

Clathrin-dependent endocytosis, also known as receptor-mediated endocytosis, is the uptake of material into the cell from the surface using clathrin-coated vesicles, a process utilised by all eukaryotic cells (11, 17, 18). Receptor-mediated endocytosis controls constitutive and stimulated internalization of many receptors with roles in cellular homeostasis, growth control, cell differentiation and synaptic transmission (17). Additionally, this pathway is also exploited by viruses and bacteria to enter cells (17). This pathway is highly versatile, and a variety of cargo can be packaged into these endocytosed vesicles. This pathway begins with the binding of a specific molecule to its respective transmembrane signalling receptor (18, 19). Following the activation of the signalling cascade, the clathrin-coated inner leaflet of the membrane curves inwards. This curvature eventually develops to a point where it can be pinched off to create an endocytic vesicle, incorporating not only the signalling receptor but also a sample of extracellular milieu (11). This clathrin-coated vesicle formation and internalisation moves through five distinct stages; initiation, cargo selection, coat assembly, scission and uncoating (17). Of these five stages, the scission or “pinching” off from the membrane to form the endocytic vesicle, a process requiring both membrane fusion and fission to occur, is of most relevance for this study (*Figure 1.2B*). This process requires a large assembly of proteins, but the first structural protein involved in the essential membrane-remodelling step during membrane fission is the GTPase protein, dynamin (20).

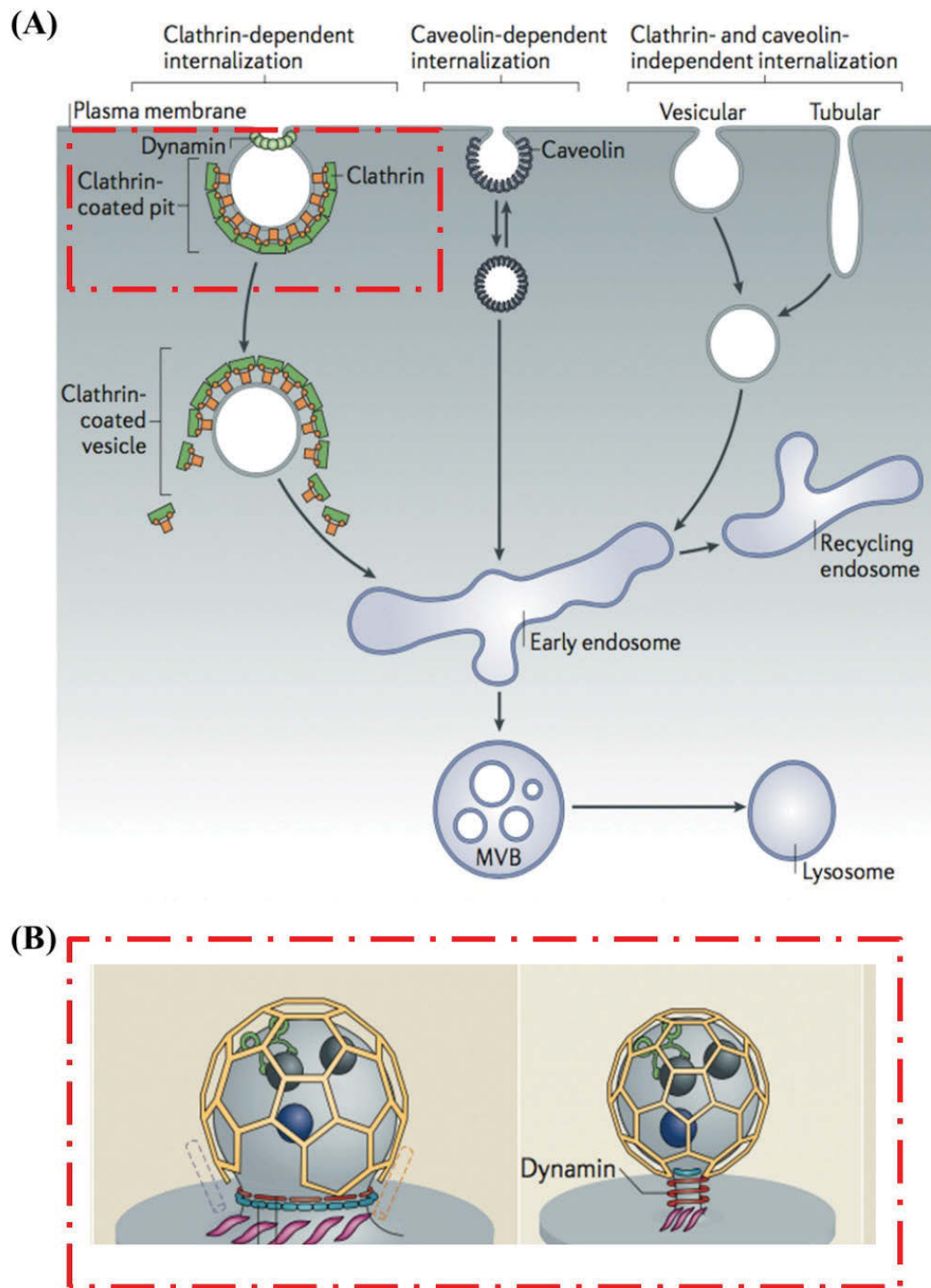


Figure 1.2 Different Pathways of Endocytosis in Eukaryotic Cells

(A) There are three distinct pathways of endocytosis, clathrin-dependent, caveolin-dependent and clathrin- and caveolin-independent internalisation. Once the cargo is internalised it is commonly trafficked into endosomes, where it can be sent back to the surface of the cell or to other cellular compartments for degradation (i.e. to the lysosome). (B) Specifically, for the pinching-off of the clathrin-coated vesicle (red box) and subsequent endocytosis of the vesicle encapsulated in the clathrin cage, the GTPase protein dynamin is recruited to the neck of the budding vesicle. It is here that it self-polymerises on the membrane and, upon GTP hydrolysis, induces membrane scission of the vesicle from the membrane. Adapted from (17).

1.2 DYNAMIN: AN INTEGRAL GTPASE PROTEIN INVOLVED IN MEMBRANE DYNAMICS

The discovery of dynamin came with identification of temperature sensitive mutants in the locus termed *shibire* in *Drosophila melanogaster* that were first shown to induce paralysis in the fly (21, 22). Specifically, it was shown that this mutation blocked endocytosis in a temperature dependent manner, preventing the reformation of vesicles after the release of neurotransmitters and the accumulation of arrested ‘collared’ endocytic pits at the plasma membrane (23). Later it was shown that this particular locus encoded a dynamin-like gene forming the first evidence for the involvement of dynamin in vesicle trafficking within a cell (24). Since these fundamental studies, the mechanism behind this essential function has been explored in depth, highlighting the importance of the structural features of dynamin in carrying out this function.

Dynamin is a ~100 kDa protein comprised of five distinct structural domains; a GTPase domain, a stalk, a Pleckstrin Homology (PH) domain, a GTPase Effector Domain (GED) and a Proline Rich Domain (PRD) (25-30). Each of the members of the greater dynamin superfamily show a distinct function within the eukaryotic cell and these differences are related to the presence and absence of particular structural domains (1, 2). Starting at the N-terminus, the GTPase domain contains the residues necessary to form the active site for GTP

binding and hydrolysis. This domain sits adjacent to a helical bundle, termed the bundle-signalling element (BSE), which links the GTPase domain to the GED (*Figure 1.3*). The BSE is comprised of residues from both the GTPase domain as well as the C-terminus of the GED domain (29-31). This region is followed by the stalk, or middle domain, involved in dimerisation. The dynamin homo-dimer assembles in a cross-like pattern, as seen in the lower panel of *Figure 1.3* (32, 33). The “foot” of the dynamin protein, otherwise known as the PH domain, binds acidic phospholipids, in particular phosphoinositide-4, -5 bisphosphate (PIP₂), on the inner side of the plasma membrane (27, 28). The interaction between lipids and the PH domain is crucial for membrane remodelling (34). On the C-terminus is the PRD, containing an array of Pro-X-X-Pro motifs that facilitate interactions with other proteins and cellular structures including the cytoskeleton (35-39). The interactions of dynamin proteins with both lipid and themselves (dimerisation) are of comparatively low affinity but have all been shown to increase in affinity upon polymerisation at the site of function (40, 41).

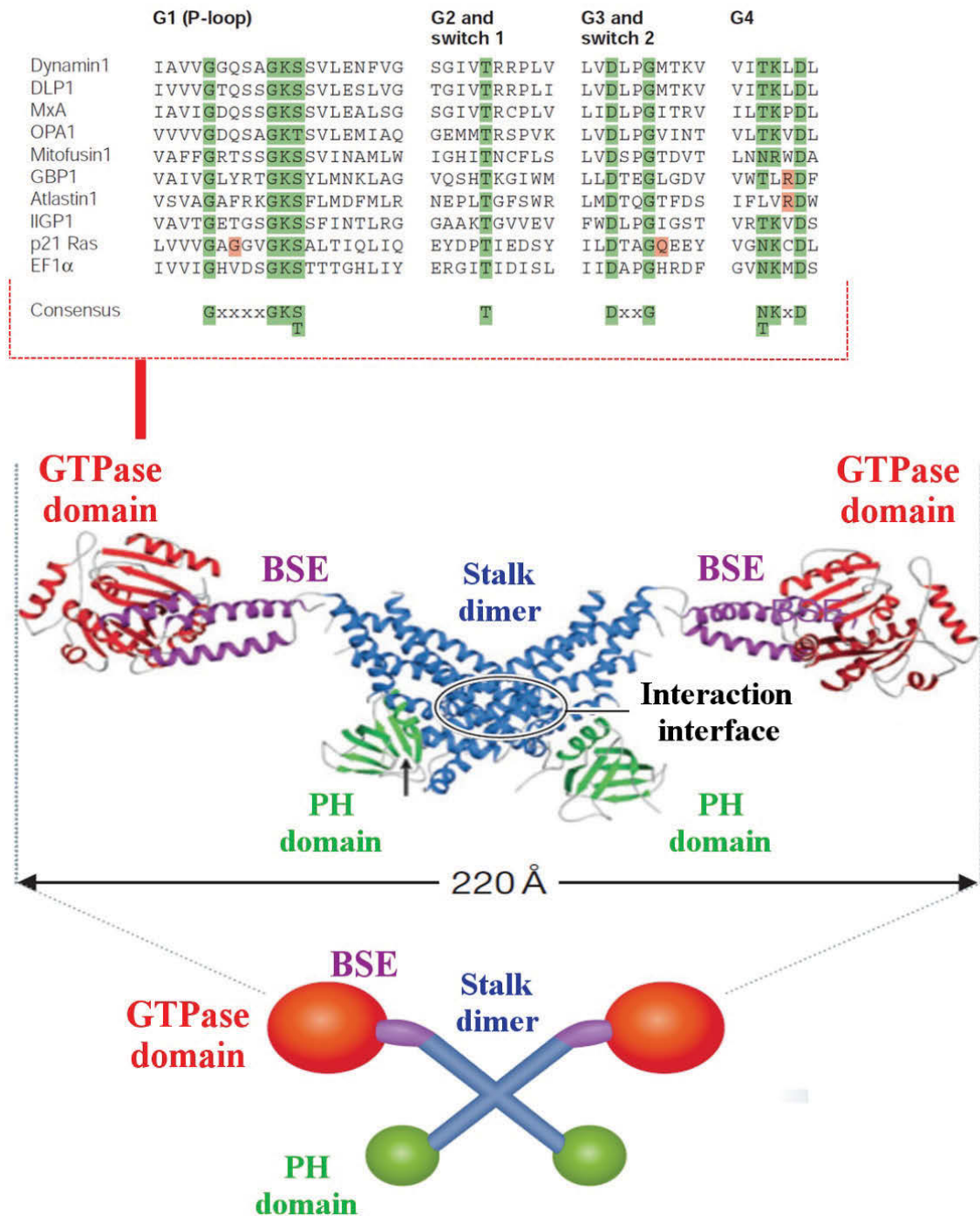


Figure 1.3 The GTPase Dynamin Dimer Structure

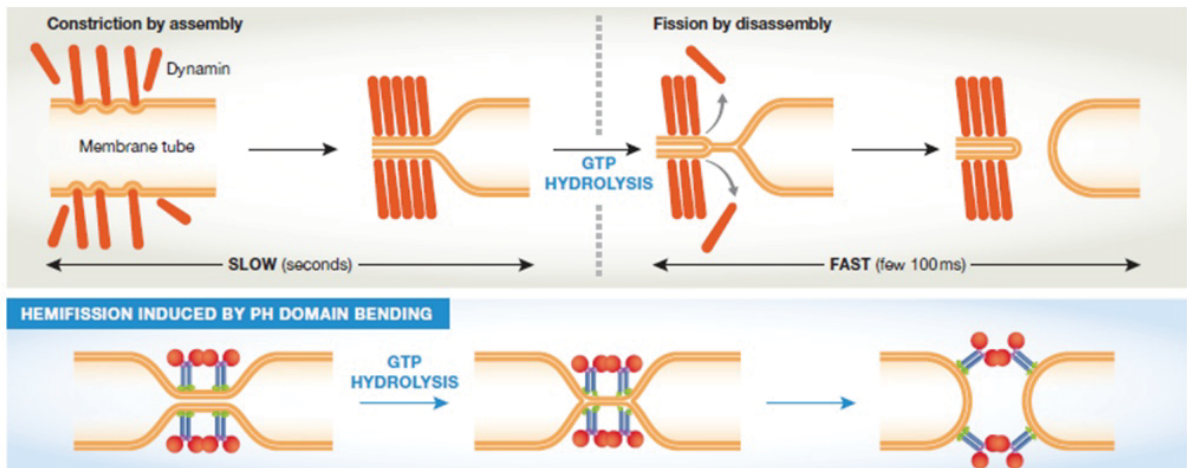
Dynamin is a GTPase protein that has a number of structural features including but not limited to a GTPase domain (red), basal signalling element (BSE) (purple), stalk domain (blue) and a pleckstrin homology (PH) domain (green). Across the entire dynamin

superfamily, the GTPase motif (G1, G2, G3 and G4) is conserved with the consensus for representative members of the family shown above (green amino acids). As can be seen in the crystal structure of the dynamin dimer, interaction between dynamin monomers occurs at the interaction interface via the stalk domain. The lower part of the figure shows a schematic representation of this dimerisation. Adapted from (2, 33).

The energy generated by GTP hydrolysis is essential for the dynamic membrane functions of dynamins. The GTPase domain contains a conserved core that consists of approximately 300 amino acids that form a mixed six-strand beta sheet that is enclosed by five alpha helices (29, 31). Only a single GTP molecule will bind per GTPase domain but the binding residues are spread across the entire domain and these are; G1 –Gly-X-X-Gly- Lys-Thr-Ser, G2 - Thr, G3 – Asp-X-X-Gly and G4 – Asn-Lys-X-Asp. These conserved residues are present in the GTPase domains of all members of the dynamin superfamily, and it is thought that most members of this superfamily can bind GTP (1, 2). Interestingly, despite the presence of these residues across diverse members of the superfamily, many of the dynamin-like proteins have a relatively low affinity for GTP and have an even lower affinity for GDP, the hydrolysed product, compared to classical dynamins (42, 43). While the catalytic mechanism that underlines GTP hydrolysis is generally conserved across the greater dynamin family (*Figure 1.3*, upper panel), the details of the catalysis differs between different family members.

There has been intense interest and ongoing debate as to how dynamin uses the interplay between lipid binding properties and GTP hydrolysis to catalyse membrane structural changes (33). The two recent models of dynamin-mediated membrane fission are outlined in *Figure 1.4*. Both models try to resolve the different properties of dynamin that have been established over the many years of research. One model invokes a two-stage mechanism where membrane constriction is mediated by dynamin assembly, and then membrane fission is mediated by dynamin disassembly (*Figure 1.4A*) (44, 45). The second, more well-known model, is the constriction/ratchet model, whereby dynamin is considered a molecular motor using GTP hydrolysis to “twist” and “squeeze” the membrane (*Figure 1.4B*) (25, 26, 29, 46). Currently this remains an unresolved question, highlighting a clear gap in the understanding of how dynamins carry out their function.

(A) Model One – Two Step



(B) Model Two – Constriction/Ratchet

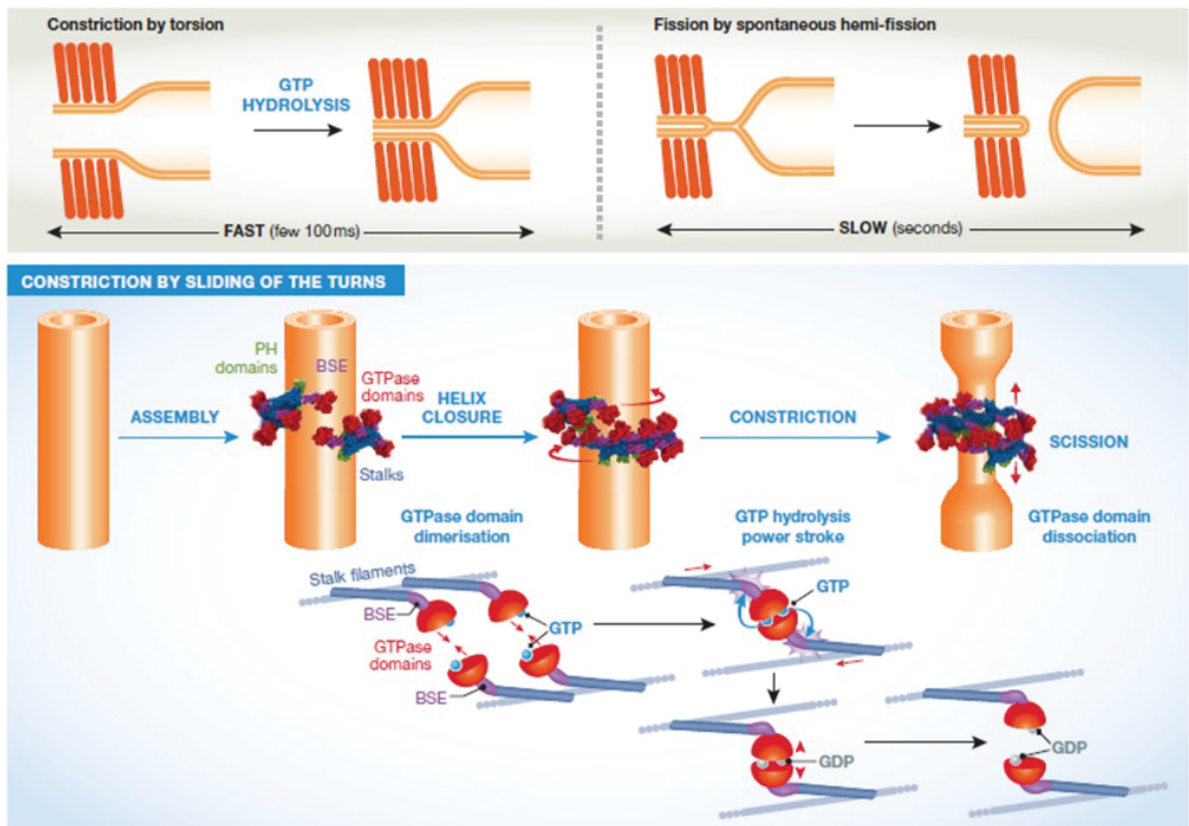


Figure 1.4 Two Proposed Models of Coupled Dynamin GTPase Activity and Lipid Binding

(A) In the first stage, dynamin assembles in a nucleotide-loaded conformation, causing the protein to be in a constricted state forming hemi fission intermediates. Then upon GTP hydrolysis when opposing GTPase domains of neighbouring proteins interact to release the phosphate to form GDP, the binding of dynamin to the lipids weakens and the hemi fission vesicle intermediates progress through to complete fission. This model takes into consideration that dynamin will disassemble quite quickly upon GTP hydrolysis. (B) The second model, termed the constriction/ratchet model, proposes that dynamin acts as a motor. This means that GTP hydrolysis produces mechanical energy allowing a “squeeze” and “twist” motion of the dynamin helix. The direct interactions between opposing GTPase domains act as molecular motors in continuous cycles of association and dissociation. The sliding of the helical turn is triggered and subsequent constriction and twisting of the dynamin helix occurs resulting in the scission of the membrane. Adapted from (33)

1.3 THE DYNAMIN SUPERFAMILY

Despite being studied for over 20 years, many aspects of the dynamin superfamily remain elusive, including both structural and functional aspects. However, it quickly becomes apparent though that viewing dynamin solely as an endocytic protein is overly simplistic. Instead, dynamin should be and is considered the prototypical member of a superfamily of proteins called the dynamin-related proteins, or dynamin-like proteins, that carry out a plethora of membrane remodelling functions throughout the cell (*Figure 1.5*).

1.3.1 CLASSICAL DYNAMIN

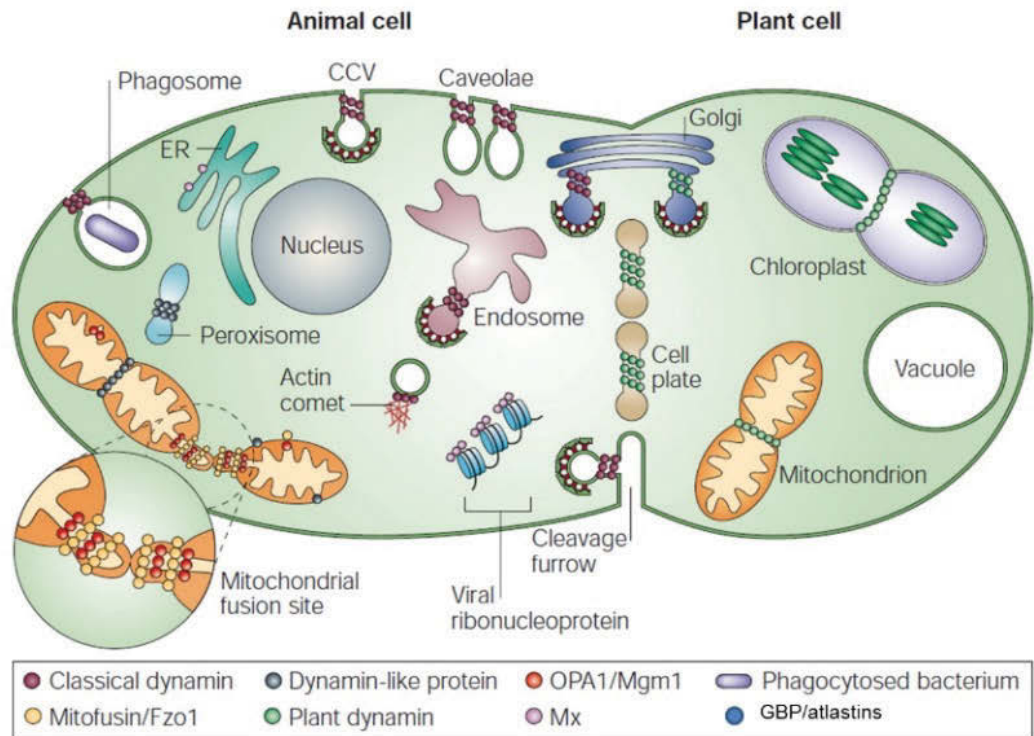
The most well known dynamin is the classical dynamin, of which there are three distinct isoforms encoded in mammalian genomes (1, 47). In most eukaryotes these three isoforms of dynamin share 80% sequence similarity to each other and have similar structures; however, their production and activity are independently regulated (2). Each of these proteins have the typical five-domain architecture; described below/above (29). The first isoform, Dynamin 1, is highly expressed in neurons and is involved in endocytosis during

neurotransmitter uptake at the plasma membrane (48). It is generally considered that Dynamin 1 is not present in other tissues in the body (2). The second main isoform, Dynamin 3 is also considered to be localised to neuronal cells, albeit at much lower concentrations compared to Dynamin-1, as well as being found elsewhere in the body at low concentrations such as the testes and the lung (49, 50). The final classical dynamin isoform, Dynamin 2 is expressed throughout many if not all tissues in the body (51). Dynamin 2 could be carrying out general endocytosis for all cells, and Dynamin 1 and Dynamin 3 would appear to be required for specialized functions in neurons (1). Adding further complexity, each of these three classical dynamins also may exist as multiple splice variants, resulting in a large number of variant isoforms that presumably have differing sub-functions (52). Abnormal dynamin function has been implicated in the development of genetic diseases, such Charcot-Marie-Tooth disease with mutations in Dynamin 2 (53, 54), and changes in Dynamin 1 seen in Epilepsy models (55, 56).

1.3.2 FUNCTIONAL DIVERSITY OF THE DYNAMIN-RELATED PROTEINS

Members of this superfamily that are not classed as “classical” dynamins are termed dynamin-related proteins and by definition, each of these are missing one or more of the domains seen in the classical dynamins. Dynamin-related proteins include the sub-families, dynamin-like proteins (*dlp*), Mx proteins, ARC5-like proteins (plants), Vps1-like proteins (yeast), Optic atrophy 1 (OPA1), mitofusins/Fzo1 (fuzzy onions 1) and Guanylate-binding proteins (GBPs) and their close homologs, Atlastins. As outlined in the upper panel of *Figure 1.5* upper panel, each of these dynamin-related proteins carries out a different role in the cell, including vesicle budding, organelle division, cytokinesis and protection against invading species like bacteria and viruses (2). The overarching theme of membrane remodelling underpins their biological roles. This functional diversity can be related to the modular domain architecture of dynamin superfamily members (*Figure 1.5 lower panel*)

(A) Diverse Dynamin Superfamily



(B) Dynamin Superfamily Domain Architecture

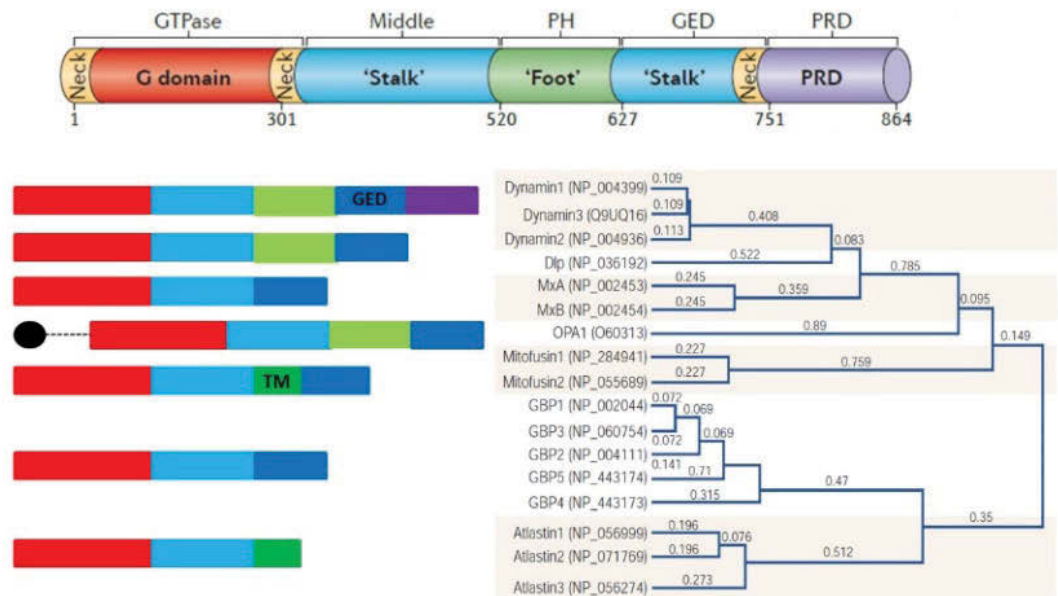


Figure 1.5 Dynamin Superfamily and their Conserved Domain Architecture

(A) The dynamin superfamily carry out membrane remodelling functions throughout both the animal and plant cell. These functions include the budding of clathrin-coated vesicles at the plasma membrane (classical dynamins, maroon) and division of organelles such as the mitochondria (Dynamin-like proteins (Dlps), grey and OPA1 (optic atrophy 1), red and mitofusins, yellow). The Mx family (pink) is induced by interferons and confers resistance against RNA viruses. Plants contain many different dynamin proteins (light green) that have functions in the formation of the cell plate or/and chloroplast division. (B) Schematic of the domain architecture of the dynamin superfamily coloured according to classical domains; GTPase domain (red), a middle domain (light blue), a GTPase effector domain (GED) (dark blue) and lipid binding domains, pleckstrin-homology (PH) domain (green), a transmembrane domain or a sequence for lipid attachment (green). Classical dynamins contain an additional proline-rich domain (PRD) (purple). Dynamin superfamily members have been grouped according to their domain structure and their accession numbers are shown. Adapted from (2)

1.3.2.1 Plant Dynamins

There are a number of plant-specific dynamins that are involved in organelle division, such as chloroplast division (*Figure 1.5*). They lack the classical PRD and PH domains (2). In plant cytokinesis, the ARC5-like and DymA dynamin-related proteins are involved in the formation of the tubular-vesicular networks, the precursors to the cell plate, as well as having a role in the retrieval of excess membrane through vesicle budding (*Figure 1.6*) (2, 57, 58). Interestingly, in eukaryotic cells, one of the classical dynamin isoforms of Dynamin 2 has been implicated in this type of function in the completion of cytokinesis through a cycle of phosphorylation at division sites (59). This is postulated due to the close relation of Dynamin 2 to the cytoskeleton, with the rearrangement of microtubules, a fundamental protein structure, essential in the successful mitosis and meiosis of a eukaryotic cell (59).

1.3.2.2 Mx Proteins

The Mx family of dynamin-related proteins are missing the same domains as the plant dynamin-related proteins, i.e. the PRD and PH domains (2). They have been shown to be expressed in response to type I interferons and play a vital role in viral protection of the host cell (*Figure 1.6*) (60). In humans, the dynamin-related protein, MxA has been shown to possess a diverse activity against a range of viruses (61, 62). The exact mechanism behind how Mx proteins carry out their antiviral activity is still not known. It is thought that MxA may interact with the nucleocapsid of the virus (63, 64) by associating with the smooth ER, a site where viruses can pass through. MxA may interact with viruses at this location to stop their replication (65).

1.3.2.3 Mitochondrial associated Dynamins

There are three protein sub-families in the dynamin superfamily, Dynamin-like protein (*dlp1*), OPA1 and mitofusins, that are implicated in the division of organelles particularly the mitochondria (2). The mitochondria, which has both an inner and outer membrane, is constantly undergoing fusion and fission for a number of biological processes, such as the turnover of integral membrane proteins and the correct distribution of mitochondria throughout the cell (66). Dlp1, of which there is only one homolog per organism, lacks the PRD domain of the classical dynamins and is the only one of these three sub-families shown to be involved in membrane fission (*Figure 1.6*) (67, 68). Dlp localises to the mitochondrial outer membrane and places pressure on both membranes of the mitochondria to induce division (67, 69).

The other two dynamin sub-families, OPA1 and mitofusins, have been implicated in membrane fusion in the mitochondria (*Figure 1.6*) (1, 2). OPA1 contains an additional targeting sequence to the mitochondria, and is associated with the outer surface of the mitochondrial inner membrane, where it is thought to play a role in linking the two membranes together (70-72). Mitofusins are present in all eukaryotic cells and have a predicted transmembrane domain in the place of the PH lipid-binding domain in classical

dynamins (73, 74). They are located on the outer surface of the outer membrane and have been found to interact with OPA1 (75). The two proteins are most likely functioning together to carry out membrane fusion by regulating and/or creating areas of high curvature in the organelle through their lipid binding capabilities and their protein-protein interactions.

1.3.2.4 Distantly related Dynamins

GBPs and their close relatives the Atlastins are arguably the most distantly related members of the dynamin superfamily, with only the GTPase domain showing similarity (76). Atlastins have been shown to catalyse membrane fusion of the endoplasmic reticulum (77, 78). GBPs are another example of dynamin proteins being regulated by an interferon response (Type II), similar to the Mx family, and have been shown to play a role in resistance against intracellular pathogens (79).

As seen above, the eukaryotic dynamin superfamily is made up of several structurally similar proteins that carry out an extremely diverse GTPase-dependent membrane remodelling function within the cell. There are still many unanswered questions when it comes to the specific mechanoenzymatic properties of dynamin in terms of the coupling of membrane remodelling and GTPase activity. Further study of different members of the dynamin superfamily will continuously add to the understanding of the important biological process that is membrane remodelling.

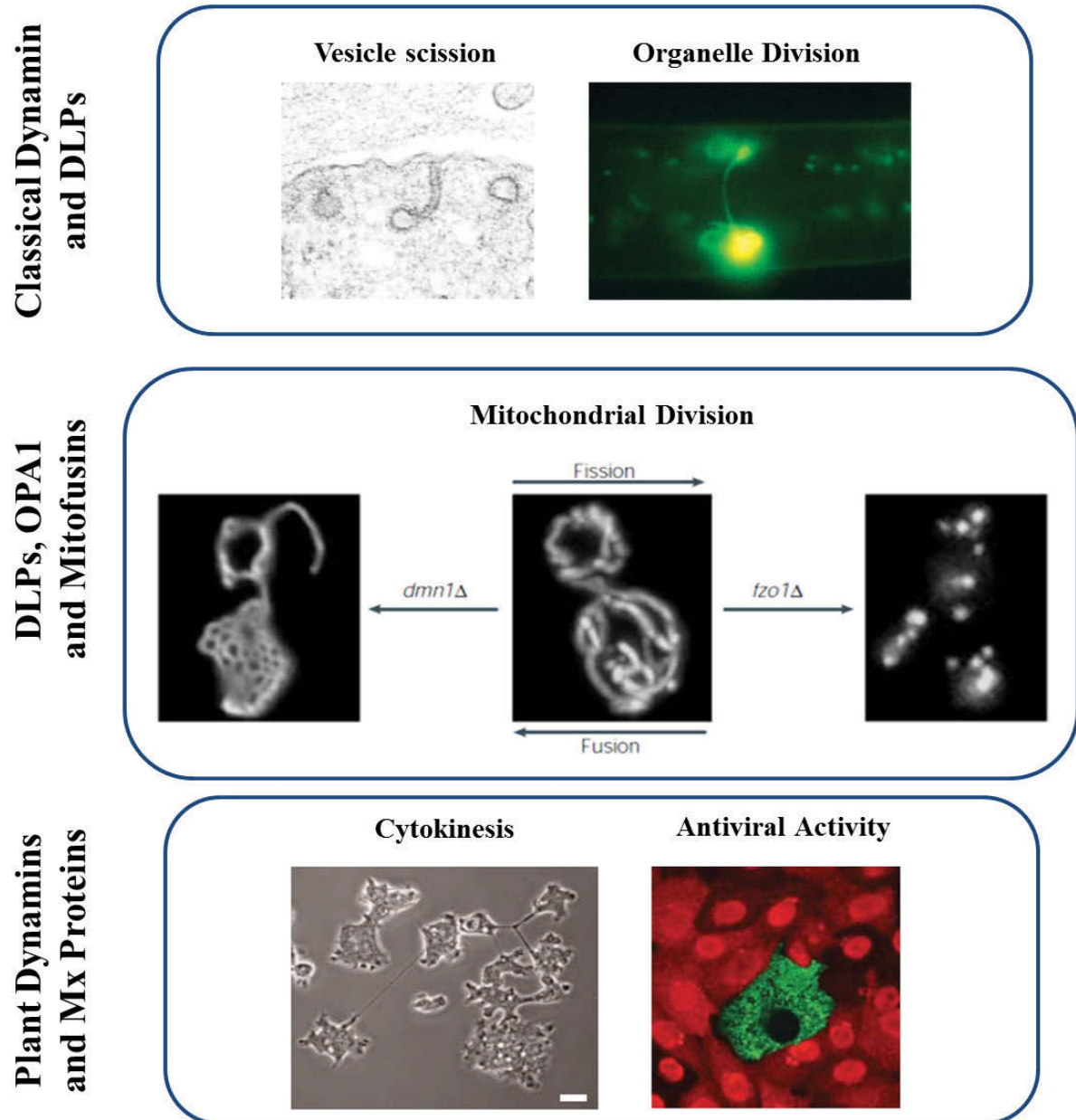


Figure 1.6 Diverse Functions of the Dynamin Superfamily

Top Left - Classical dynamin can be seen wrapped around the neck of an invaginating membrane pit. Top Right - With a mutation in *Dlp1*, divided nematode cells remain connected by the outer membrane of the mitochondria. A defect in the division of the organelle. Middle - The important interplay between the sub families involved in mitochondrial division is seen with a deletion of *Dlp1* (*dmn1*) inhibiting the fission of the membrane, forming 'networks' of mitochondria. Whereas a deletion of a mitofusin (*Fzo1*) has the opposite effect with increased

fragmentation or fission evident, with no ability to oppose this process with fusion of the membrane. Bottom Left - Failed division of soil-living amoebae can be seen in the absence of dynamin with 'bridges' formed between the cytoplasm of divided cells. Bottom Right - Successful anti-viral activity is seen (green) in a cell transfected with MxA compared to neighbouring cells that do not have protection, labelled for a viral protein (red). Adapted from (2)

1.4 THE DIVERSE BACTERIAL MEMBRANE

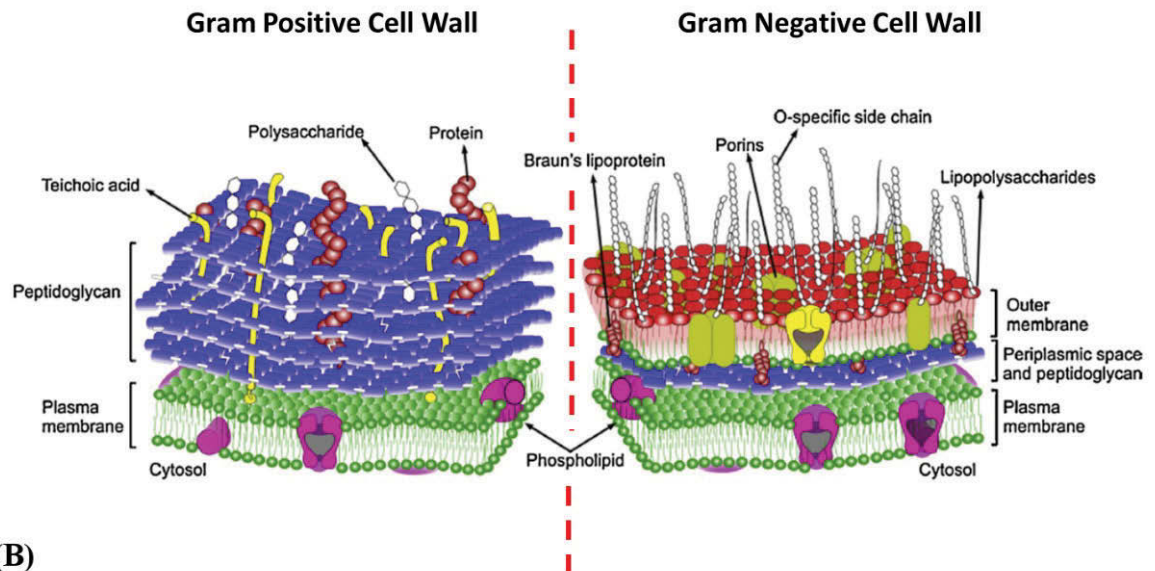
Despite their minimal genomes when compared with eukaryotes, bacteria are strikingly complex organisms that possess subcellular structures crucial for their survival and pathogenicity. Many of these features are distinctively bacterial and are not found among the other domains of life, Archaea or Eukarya. For both Gram-positive and Gram-negative bacteria, their cytoplasmic membrane is the essential semi-permeable barrier to the external environment that regulates transfer of nutrients and virulence factors, as well as provides protection from stresses and predators (80). The cell envelope is composed of the plasma membrane (inner membrane) and cell wall (primarily composed of peptidoglycan). As in other organisms, this wall provides both structural integrity and protection to the cell. Gram-negative bacteria also have an outer membrane that houses and forms key modulators important in host pathogen interactions such as LPS and outer membrane vesicles.

The Gram-positive envelope, seen in *Figure 1.7A*, consists of a thick peptidoglycan layer, which can make up 95% of the cell envelope (81). In the peptidoglycan layer, a number of different components can be found including proteins, polysaccharides and teichoic acids. Teichoic acids are polymers of ribitol and glycerol phosphate and are only located on the surface of Gram-positive bacteria (82, 83). The exact function of teichoic acids is not fully understood. However, in host organisms they are believed to act as surface antigens for eukaryotic toll-like receptors (83, 84). There are a number of membrane remodelling events, seen in *Figure 1.7B*, that are important in Gram-positive cells, including cytokinesis, spore engulfment and maintenance of membrane integrity. Much of the membrane dynamics taking

place in these types of remodelling events is provided by interactions between proteins and the plasma membrane (85).

In Gram-negative bacteria, the envelope consists of inner and outer membranes that are separated by a space termed the periplasm (86) (*Figure 1.7A*). It is within this periplasm that the relatively thin peptidoglycan layer can be found, making up only 5-10% of the cell envelope compared to Gram-positives (81). Interestingly, the two membranes have different structures and compositions; the outer membrane consists of an asymmetrical bilayer made up of inner-leaflet phospholipids and outer-leaflet lipopolysaccharides (LPS), and integral outer membrane proteins (OMP). The proteins in the outer membrane can form protein pore structures, called porins, which are large enough to allow diffusion of some molecules across the membrane into the periplasm (87). The inner membrane of a Gram-negative cell is primarily a phospholipid bilayer, likewise consisting of integral proteins (88). As pictured in *Figure 1.7B*, there are numerous membrane remodelling events that occur in the Gram-negative cell including cytokinesis, outer membrane vesicle biogenesis and release as well the remodelling of the OM to form long tubes that have been implicated in long range electron transport (nanowire) and more distant cell-to-cell contacts (85). Membrane structural changes require interaction between numerous proteins and the inner and outer membranes, to induce curvature and change (85).

(A)



(B)

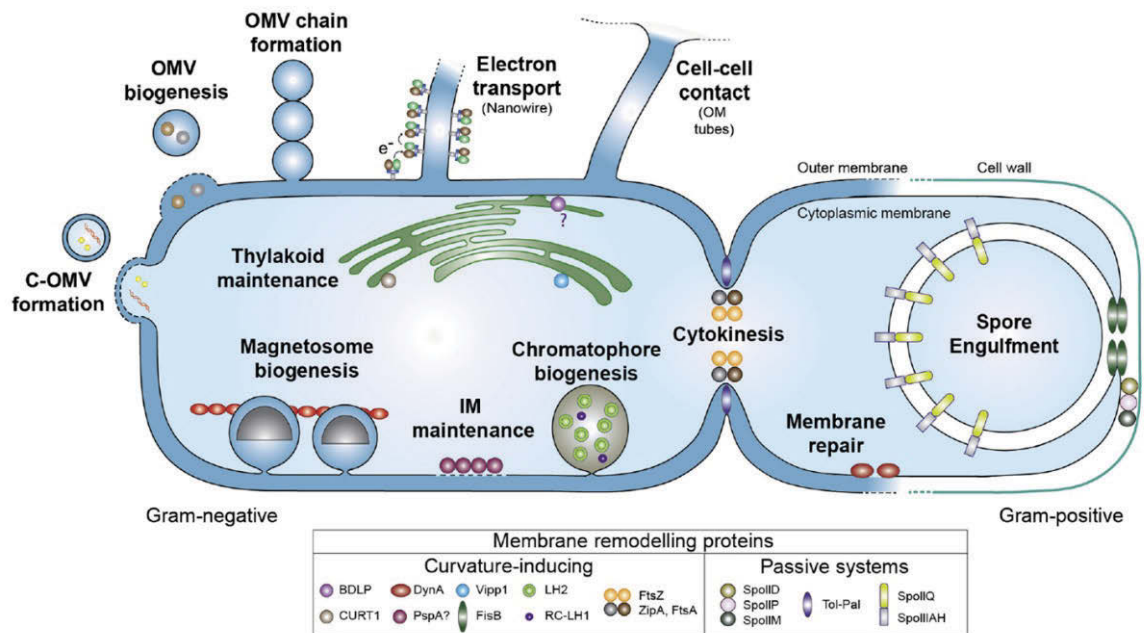


Figure 1.7 Gram positive and Gram-negative Bacterial Membranes

(A) A labelled schematic diagram of the cell envelope of Gram positive (left) and Gram-negative (right) bacteria. Gram-positive bacteria have a single membrane with a thick layer of peptidoglycan whereas Gram-negative bacteria have a double membrane separated by a thin layer of peptidoglycan in the periplasmic space. Adapted from (89). (B) Both Gram positive (right) and Gram-negative (left) bacteria undergo a range of different membrane remodelling events (pictured) including cytokinesis, outer membrane vesicle formation and

membrane repair and maintenance. These events are carried out by a number of different membrane binding proteins such as bacterial dynamins as well as passive systems such as Tol-Pal. Adapted from (85).

The bacterial cell wall differs from archaea and eukaryotes by the presence of the sugar-amino acid polymer, peptidoglycan, located immediately outside of the cytoplasmic (inner) membrane. The peptidoglycan layer is one of the key components that vary between Gram positive and Gram-negative bacteria as mentioned above (80, 81). This layer maintains cell rigidity and structural integrity of the cell, two important things the cell must possess to survive in a range of environments. Along with this, the peptidoglycan layer controls turgor pressure in the cell by resisting osmotic forces in the face of differing environmental conditions (90). The membrane remodelling events crucial to the survival of the cell mainly occur through proteins interacting with the membrane in response to both intra and extracellular signals. One such protein with an established function in membrane dynamics in eukaryotic cells that may also be carrying out a similar membrane remodelling function in bacteria is dynamin.

1.5 BACTERIAL DYNAMIN-LIKE PROTEINS

Until recently, the dynamin superfamily was thought to be restricted to eukaryotic cells, but is now known to be widespread among many bacterial species, with nearly 1500 dynamin-like protein (DLP) sequences, mostly uncharacterised, across over 900 bacterial species identified up to 2012 (6). Bacterial dynamins have the strongest relationship to the eukaryotic dynamin related proteins in the mitofusins and atlastins families (6). Whilst the bacterial dynamins share quite a low sequence identity to the eukaryotic members, they possess striking structural similarities. An interesting feature of bacterial dynamins is that they are almost always encoded in pair of tandem homologs in the bacterial genome (6). In one bacterial dynamin, DynA of *Bacillus subtilis*, a gene fusion of two standard DLPs has been observed, thus encoding a fused pair of full-length dynamin-homologs (3, 5). The function of dynamin-like proteins in bacteria, as well as their biochemical characteristics, remain relatively poorly understood and speculative.

1.5.1 THE FIRST CHARACTERISED BACTERIAL DYNAMIN – BDLP1

The first Bacterial Dynamin-Like Protein 1 (BDLP1) was identified in a cyanobacterium, *Nostoc punctiforme* (4). The crystal structure (*Figure 1.8A*) showed that this protein possessed the main structural features of a dynamin, including a N-terminal GTPase domain and a predicted C-terminal hydrophobic transmembrane “paddle”. To begin to understand bacterial dynamins, functionality of the GTPase domain was investigated. Using recombinant BDLP1, GTPase activity was shown to be similar to that of the previously characterised eukaryotic dynamin family members with a K_m of 68.6 μM . This was further confirmed by analysis of a GTPase P-loop mutation (K82A), located within the active site, which greatly reduced the GTPase activity (*Figure 1.8B*) (4). Interestingly it was found that BDLP1 had no detectable GTPase activity when bound to liposomes, suggesting a much lower GTP turnover in the process of membrane remodelling (4). Additionally, BDLP1 cannot efficiently bind liposomes when GDP is bound, suggesting the role of hydrolysis as a mechanism of displacement of BDLP1 from the membrane, a previously noted concept in eukaryotic dynamins (45, 91, 92).

The first studies of bacterial dynamins showed that BDLP1 from *N. punctiforme* can interact with lipid membranes *in vitro* and generate narrow protein-lipid tubules (91). BDLP1 decorated the lipid structures in a non-hydrolysable-GTP dependent manner, suggesting that the tubules might represent transient structures in a dynamic lipid remodelling process. An electron cryomicroscopy structure of the BDLP1-lipid bilayer tubule (91), showed that the paddle domain binds to the lipid membrane and an extended form of BDLP1 interacts with the neighbouring monomer to form a “cart-wheel” like architecture (*Figure 1.8C*). The crystal structures of BDLP1 the first of any dynamin superfamily protein both with and without nucleotide revealed a compact folded molecule (*Figure 1.8A*). By comparison with the lipid tubule structure, where an open “T-shape” form is adopted, a large repositioning of domains was observed in the crystal structure, through a conformational change in the central hinge regions of the protein (*Figure 1.8C*). It was proposed the trunk opens up first, followed by the rotation or opening up of the G domain at the top of the neck (*Figure 1.8D*). Both changes would suggest that nucleotide and lipid binding drive these substantial

intramolecular rearrangements though there are a few differences between the crystal structure and the density map that were not accounted for, highlighting that this model is not yet perfect (91). Despite these structural studies of BDLP1, minimal research has been conducted into its biological function.

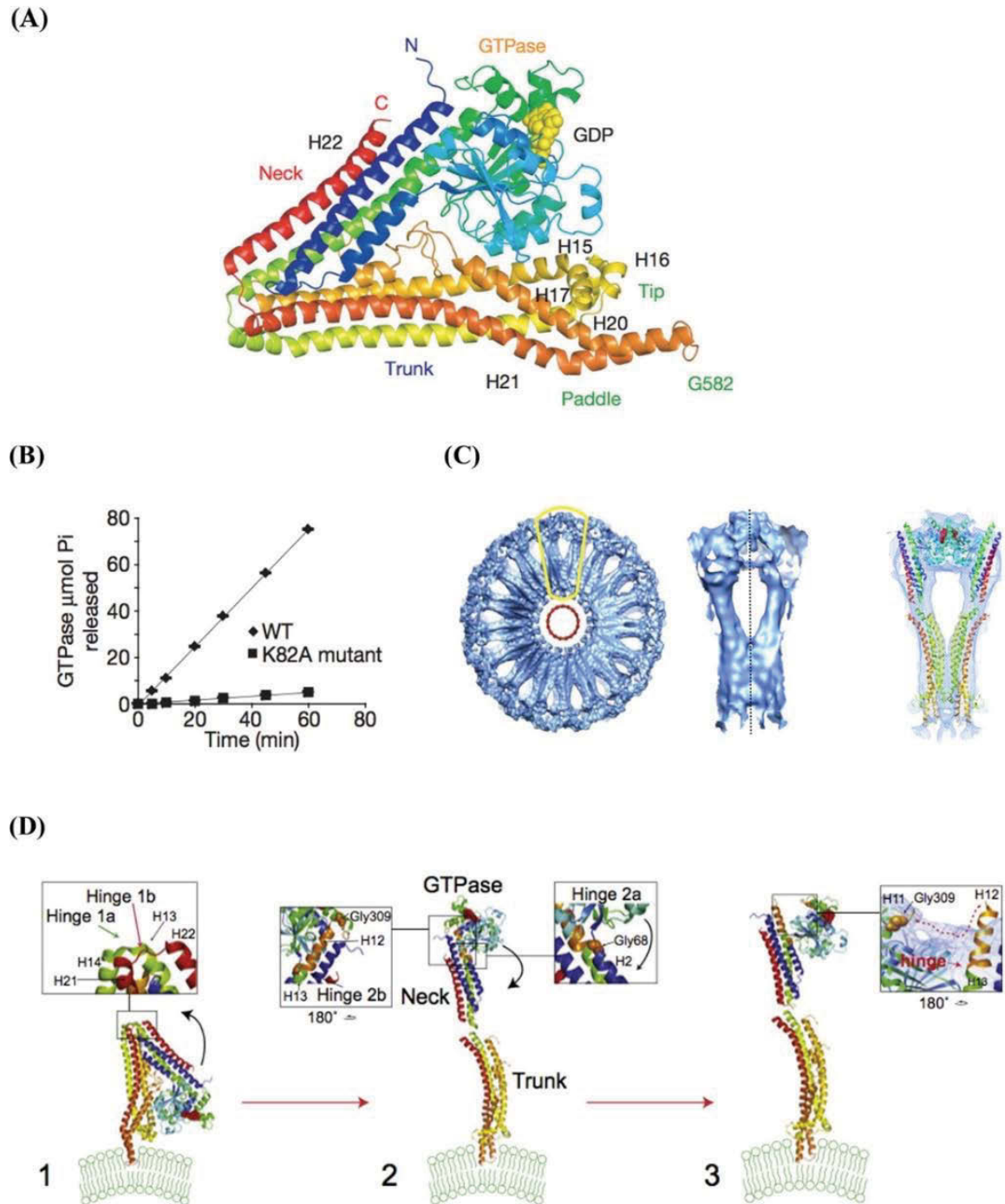


Figure 1.8 BDLP1 has the Conserved Dynamin Structural Features

(A) The BDLP1-GDP monomer and domain assignment (GTPase, Neck, Trunk and Paddle) is rainbow-coloured from the N (blue) to the C (red) terminus. (B) GTP hydrolysis time

course comparison of wild-type (WT) and K82A (Walker A) mutant BDLP1. (C) Model of the helical BDLP1 filament fitted into the native tube reconstruction shown as a cross-section, with a close-up image of the region outlined in yellow showing the fit of two BDLP1 molecules that form the dimeric asymmetric unit of the reconstruction. (D) A three-step model of the coupling of GTP hydrolysis and lipid binding using both the GTP and GDP bound BDLP1 crystal structures. The sequence of domain rearrangements is unknown however; most changes can be accommodated with two-hinged movements. Adapted from (4, 91).

1.5.2 DIVERSITY OF BACTERIAL DYNAMINS

To date only a few bacterial dynamins have had their structure, function and/or their biochemical properties investigated. Along with BDLP1, as seen in *Table 1.1*, these include CrfC from *Escherichia coli*, IniA and IniC from *Mycobacterium tuberculosis* and DynA and DynB from *Streptomyces venezuelae*, and due to this limited research into bacterial dynamins there is minimal information available on the physiological role dynamin has within a bacterial cell.

The *yjdA* gene of *E. coli* encodes for a putative DLP known as CrfC, possessing approximately 30% identity to the structure of BDLP1. CrfC contains three of the previously mentioned structural domains of the dynamin superfamily; a GTPase domain, GED and middle domain. Interestingly, its transmembrane domain appears to have been substituted with a functional replicase clamp binding motif (93). The replicase clamp, or sliding clamp, is a ring-like structure which surrounds DNA and acts as a scaffold for the recruitment of proteins during DNA replication (94). Whilst the mechanism remains speculative, CrfC appears to contribute to the equal segregation of sister chromosomes following replication and cell division, suggesting the bacterial DLP functions are not limited to lipid remodelling events (93). Interestingly, homologs of CrfC have been identified recently in *H. pylori* genomes, signifying the potential conservation of this type of function for dynamin across different bacterial species (95).

The *iniBAC* operon of *M. tuberculosis* encodes for a putative DLP, known as isoniazid-induced gene A or IniA (96). *IniA* expression is induced upon exposure to the antibiotic isoniazid (INH) and ethambutol (EMB) and is associated with drug-resistance (97). There is a very basic connection between the dynamins a predicted GTPase and transmembrane domain (97). Interestingly, a low-resolution electron-microscopy structure of IniA on a lipid surface showed that it resembles a pore-like hexamer not yet described in the superfamily. Nevertheless, the positioning of the 6 subunits would be consistent with possible dynamin-like subunits assembled into a donut-shaped hexamer. This has led to the speculation that IniA could be part of a porous multi-drug resistance pump (97).

Recently two genes in *S. venezuelae*, Sven2472 and Sven2471, now termed DynA and DynB respectively, have been shown to be involved in yet another function for dynamins within the bacterial cell, namely the stabilisation of the cell division (FtsZ) ring during sporulation (98). Both DynA and DynB were localised in an FtsZ-dependent manner to the site of septation and were needed for stabilisation of the Z-ring by interacting with a number of members of the cell division machinery (98).

From the examples given above, it is clear that dynamins already have diverse functions throughout the bacterial cell but there is limited research on the mechanisms or even the conservation of the genes across different species. Nonetheless, there have been somewhat more in-depth studies carried out on two other identified bacterial dynamins, DynA from *Bacillus subtilis* and LeoCBA from *Enterotoxigenic Escherichia coli* (ETEC) and these will be discussed in the subsequent sections.

Table 1.1 Bacterial dynamins that have been Identified to Date

Dynammin	Organism	Structure Determined? ¹	Biochemically Defined? ²	Function	Reference
BDLP1	<i>Nostoc punctiforme</i>	Yes	Yes	Unknown	(4)
CrfC	<i>Escherichia coli</i>	No	Yes	DNA replication/segregation	(93)
HP0965 and HP0966	<i>Helicobacter pylori</i>	No	No	Unknown	(95)
IniAC	<i>Mycobacterium tuberculosis</i>	Yes	Yes	Drug Efflux	(97)
LeoA	<i>Enterotoxigenic Escherichia coli</i>	Yes	Yes	Vesicle Formation	(5, 7)
LeoC and LeoB	<i>Enterotoxigenic Escherichia coli</i>	No	No	Unknown	(5)
HP0731, 0732 and 0733	<i>Helicobacter pylori</i>	No	No	Unknown	(5, 99)
DynA	<i>Bacillus subtilis</i>	No	Yes	Membrane integrity	(3, 100)
DynA and DynB	<i>Streptomyces venezuelae</i>	No	Yes	FtsZ stabilisation during cell division	(98)

1 Crystal structure or EM structure has been published

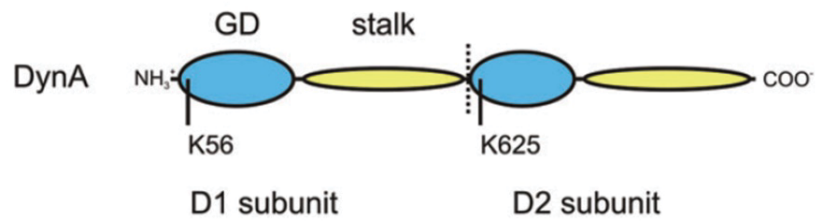
2 GTPase or lipid assays have been published

1.5.3 A UNIQUE BACTERIAL DYNAMIN IN BACILLUS SUBTILIS -DYN A

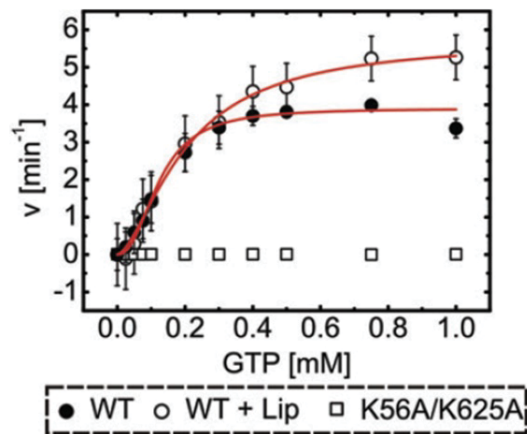
In *B. subtilis*, a protein encoded by the *ypbR* locus was discovered to be a dynamin-like protein and termed DynA (3). DynA is a 138 kDa protein, which is the result of the fusion of two ancestral dynamin genes, termed D1 and D2 (*Figure 1.9A*). Interestingly, the full-length protein is not processed into the two-individual dynamin-like subunits post-translationally. This suggests that the protein functions as a dual-dynamin unit, and components of the other known bacterial two-dynamin systems (*Section 1.5.4*) (which appear to be separately translated) might act in concert, as putative multimeric complexes, to carry out their function also (3).

An interesting finding surrounding the GTPase function of DynA came with GTP binding studies of DynA demonstrating a non-independent activity of the two GTPase domains. This was demonstrated when mutation of the individual GTPase domain P-loops (i.e. K65A and K625A) prevented their GTPase activity, and yet normal binding of GTP was observed (3). Interestingly, when both P-loops are mutated in a DynA double-mutant, GTP binding was eliminated, indicating some sort of interplay between the two domains (3) (*Figure 1.9B*). DynA was also shown to form dimers and polymerise on lipid membranes, in a nucleotide independent way (3), and membrane-association increased GTPase activity (*Figure 1.9B*). Strikingly, DynA, and more specifically D1, showed a capacity for inducing liposome membrane fusion. These impressive fusion events were dependent on the presence of magnesium ions (Mg^{2+}) and it was postulated that the ions reduced the repulsion produced by the negatively charged phospholipids, bringing the membranes closer together and permitting DynA to catalyse the lipid fusion (3) (*Figure 1.9C*).

(A) DynA is a Fused Dynamin



(B) DynA is a GTPase



(C) DynA binds to liposomes

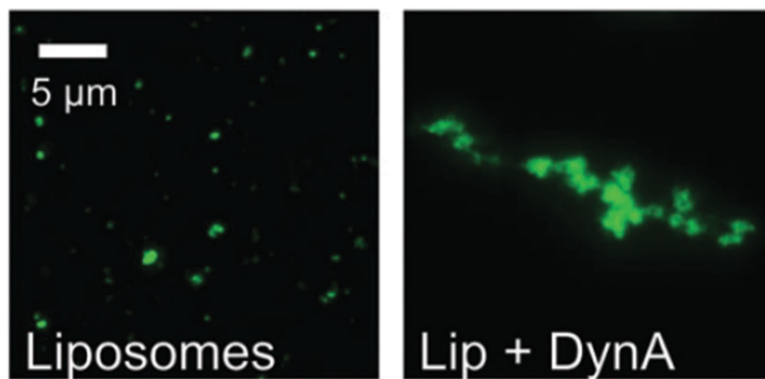


Figure 1.9 DynA is a Fused Bacterial Dynamin

(A) Graphic representation of the fused DynA protein, which contains two GTPase domains (crucial P-loop residues indicated) and a stalk region that possibly encompasses the neck, trunk and tip, similar to BDLP1. (B) GTPase activity of DynA in the presence of liposomes

(white) is increased compared to in the absence (black). When both P loop residues are mutated there is no activity detected. (C) DynA is able to tether fluorescently labelled *E. coli* liposomes (right) compared to liposome only control (left). Adapted from (3)

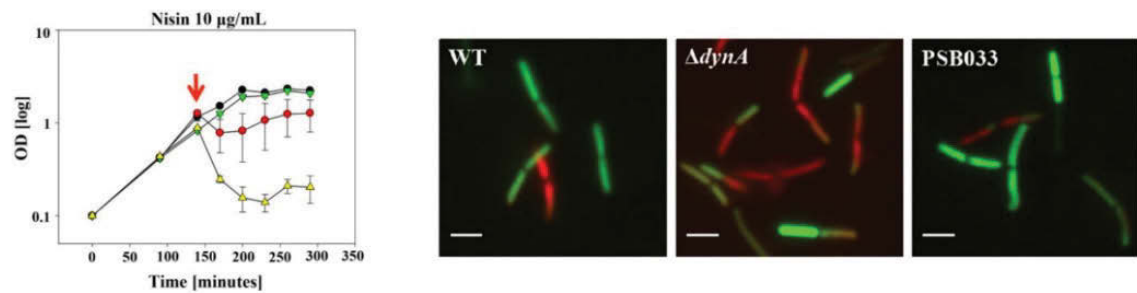
To begin to understand the potential role that DynA might have in the cell, the sub-cellular localisation of the protein was investigated. This was done through the use of the expression of DynA-GFP under the control of the native promoter. Under this system, it was found that DynA-GFP preferred to localise at the site of cell division (septation) and was completely disrupted in a deletion strain of the known division protein, MinJ (3). MinJ has been observed previously as a topological determinant of cell division in *B. subtilis*, aiding in the medial positioning of the FtsZ ring (101). However, there were no observed growth defects and no morphological changes to the cells or their nucleoids in a $\Delta dynA$ strain under normal conditions (3). This led to speculation of a potential role for DynA in cytokinesis under alternative specific environmental conditions, but a more recent study found that 5.5% of cells showed double septa or mini-cells during normal growth, which were never observed in the wild-type (WT) strain (102), indicative of a moderate cell division defect in the DynA mutant. In support of this, an interaction screen to look for binding partners of DynA identified a number of known cell division proteins (103). This supports the potential function of DynA in the stabilisation of the division machinery under normal conditions, a function similar to what has been established for DynA/B in *S. venezuelae* (98).

The functionality of DynA has since been further determined, through investigation of the phenotype of $\Delta dynA$ under environmental stress conditions, bearing in mind that *B. subtilis* is primarily a soil bacterium. Many bacteria have evolved mechanisms that regulate membrane re-modelling events to maintain cell integrity under a range of different stress conditions, with one survival strategy being to regulate membrane fluidity and permeability by altering the pre-existing fatty acids in the membrane (104, 105). In the soil environment, *B. subtilis* quite frequently comes into contact with antibiotics and bacteriophages. Therefore, $\Delta dynA$ was exposed to antibiotics that induced membrane stress, such as nisin, and it was clear that $\Delta dynA$ was not able to overcome this stress, compared to WT (Figure 1.10A). When $\Delta dynA$ was exposed to cell wall synthesis targeting antibiotics, including vancomycin, there

were no detected growth differences (100). To further investigate the effect of nisin-induced membrane stress, a live/dead staining technique was utilised. WT, $\Delta dynA$ and $\Delta dynA::dynA$ (complemented strain) were with a membrane permeable nucleic acid stain (green) and a membrane impermeable nucleic acid stain (red), seen in *Figure 1.10A*. Thus, in these experiments, cells with intact membranes appeared green and cells with damaged membranes appeared red. After exposure to nisin, it was clear that $\Delta dynA$ cells possessed damaged membranes, with a 70% increase in the number of cells staining red when compared to the normalised WT cells (100).

It is proposed that wild-type *B. subtilis* undergoes rapid membrane or lipid re-arrangement driven by the action of DynA to fill the nisin-induced membrane gaps (*Figure 1.10B*). This was consistent with the previous observation that DynA can fuse lipid membrane *in vitro* (3). Therefore, it is exciting to postulate that DynA functions by fusing the membrane back together when recruited to the membrane, specifically for repairing a pore or disruption in the bacterial membrane.

(A) DynA has a Function in Membrane Integrity



(B) Proposed Model of DynA Function

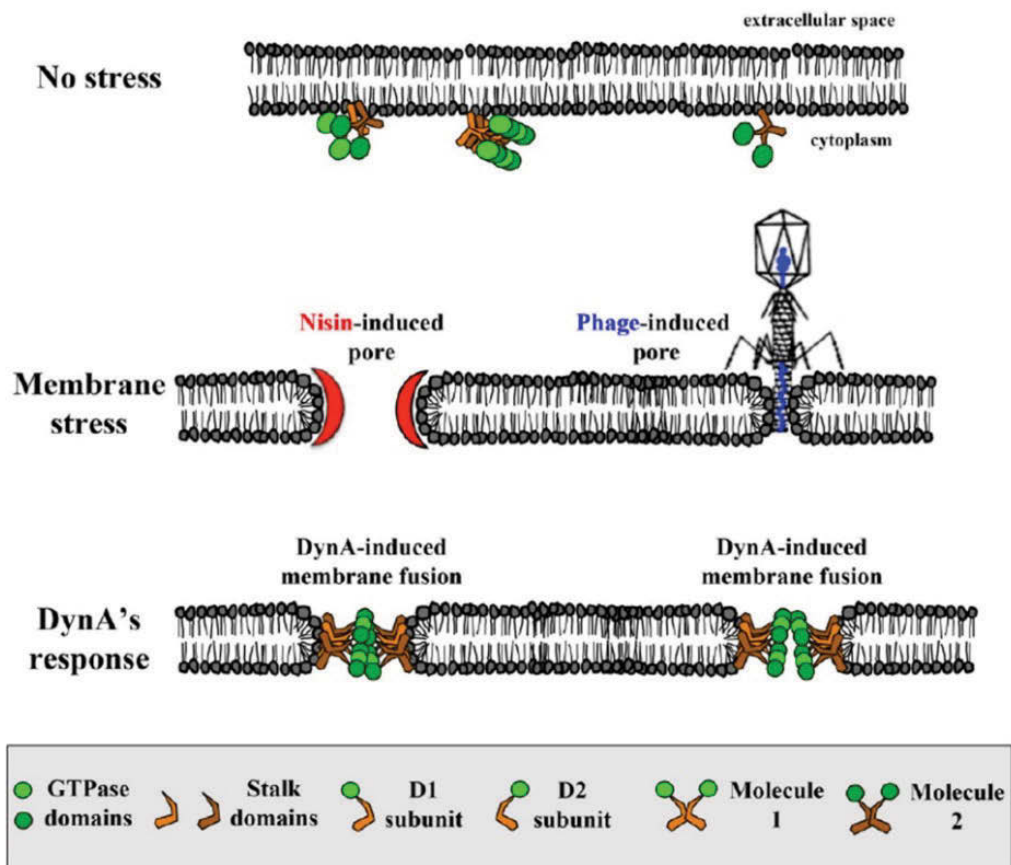


Figure 1.10 DynA plays a role in the Maintenance and Integrity of the Cell Wall

(A) When a membrane targeting antibiotic, nisin, is added to exponentially growing cells (arrow), a rapid decline in culture optical density is observed compared to untreated cells (Green and black). The *B. subtilis* WT (red) is much less affected by nisin than the Δ dynA

strain (yellow). When the wild-type, Δ dynA, and complemented (PSB033) strains were examined with a membrane permeable nucleic acid stain (green) and a membrane impermeable nucleic acid stain (red), to indicate bacterial permeability, it is very clear that in the absence of dynA, there is a higher proportion of dead or permeabilized cells. (B) Proposed model for DynA function. Upon exposure to a membrane stress like antibiotics or a phage, DynA assembles at the site of injury in the membrane, where it will exert a membrane fusion-type activity, and close the gap. Adapted from (100)

1.5.4 ENTEROTOXIGENIC *E. COLI* BACTERIAL DYNAMINS - LEOCBA

Enterotoxigenic *E. coli* (ETEC) causes a toxin-induced disease that produces significant morbidity in both humans and livestock. ETEC expresses several enterotoxins whose combined activity causes the symptoms of disease, and despite their importance in infection very little is known about the enterotoxin release from the bacteria and its subsequent delivery to the host. Outer membrane vesicles (OMVs) have been hypothesised to be one way for release and delivery.

A gene named “heat liable enterotoxin output A” (*leoA*) was found to be encoded as part of a pathogenicity island (PAI) in ETEC, which also encodes a putative adhesin known as Tia (99, 106). The clinical outcome of bacterial infections can rest upon the presence or absence of pathogenicity islands (PAIs). For example, the presence of the *cag* PAI of the gut pathogen *Helicobacter pylori* can be the determining factor between peptic ulcer disease versus asymptomatic infection (107). Remarkably, the deletion of *leoA* was shown to result in reduced secretion of labile enterotoxin (LT), which accumulated in the periplasm instead. The *leoA* deletion strain also resulted in the reduction of fluid accumulation in a rabbit ileal loop model of gastrointestinal infection indicating minimal secretion of the enterotoxin (LT) into the supernatant of the culture (Figure 1.11B, right panel) (99). Interestingly, LeoA has also been shown to be important in the formation of OMVs in ETEC (7). Therefore, given its presence on a PAI that also encodes an adhesin and a function that links to LT secretion through a role in the vesicle formation/release, LeoA represents a potential virulence factor

in ETEC. However, its importance for general ETEC-mediated infection is most likely low as it is only present in less than 3% of clinical ETEC isolates identified to date (7).

Until recently, no function reminiscent of vesicle scission during endocytosis in eukaryotic cells by dynamins had been assigned to the bacterial dynamins. However, LeoA was recently identified as a dynamin homolog (5). Whilst LeoA possesses low sequence homology to classical dynamins and some other bacterial dynamins, its crystal structure shows all the hallmarks of a DLP, including a GTPase domain, BSE, trunk and hydrophobic paddle/tip (*Figure 1.11A*) (5).

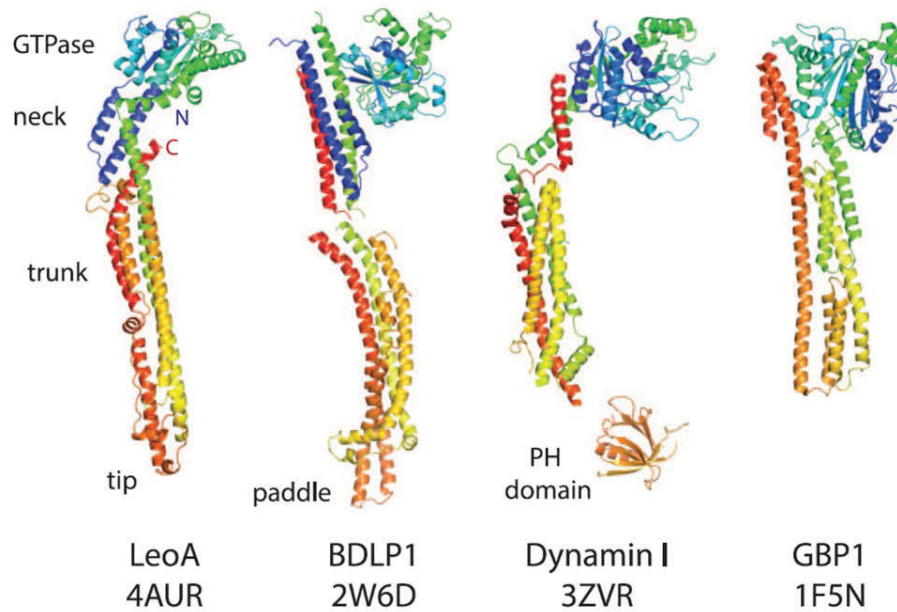
A common feature of dynamins is the ability of the protein to bind and hydrolyse GTP, which is required for membrane fusion/fission. Yet, there is a conflict in the literature surrounding the capacity of LeoA to bind and hydrolyse GTP. LeoA was first reported to carry out GTP hydrolysis, which was decreased as expected with a P-loop active site mutation (7), but these findings could not be reproduced in a subsequent independent study (5). More recently GTP binding as well as GTP hydrolysis by LeoA was not able to be detected by Michie et al, though it was noted that the LeoA active site was blocked in the accompanying crystal structure, suggesting an additional factor may be needed to allow nucleotide binding (5).

Another feature of dynamins is their ability to bind and induce membrane curvature however no interaction of LeoA with lipids was observed *in vitro*, using a range of different lipid sources, and under a range of conditions including the presence of metal ions and/or nucleotides and crowding agents (5). This suggests that something required to reflect a likely *in vivo* function is missing in the *in vitro* system currently, such as polymerisation or interaction with other proteins needed for the function of LeoA.

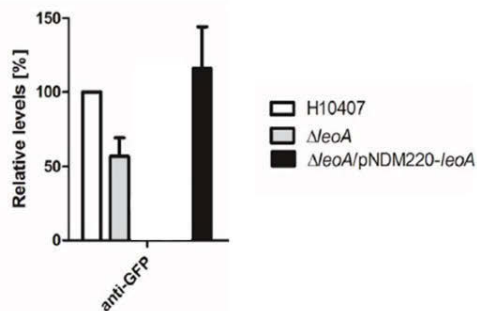
Since the initial discovery of LeoA, its function has mostly been studied in the context of LT secretion. In ETEC, LT, which is associated with the membrane, is thought to be delivered to the host through the blebbing and pinching off of the membrane to form OMVs, a process highly reminiscent of exocytosis in eukaryotes (108). OMVs are ~20-250 nm in diameter and are assumed to be derived from the Gram-negative cell envelope components (109). They

have been found to be produced by numerous pathogenic bacteria including *Pseudomonas aeruginosa* (110), *Helicobacter pylori* (111), *Niesseria* spp. (112), *Vibrio* (113), *Salmonella* spp. (114) and *Campylobacter jejuni* (115). Whilst it remains controversial as to whether OMVs are a by-product of cell breakdown or are deliberately produced by viable cells, the observation of this process across many bacterial genera suggests the latter (116). These bacteria may utilise OMVs in pathways essential for survival during infection and continual pathogenesis in their host (109). There is no single and clearly-defined mechanism to explain OMV biogenesis in prokaryotes, raising the question as to whether LeoA is involved in this process in ETEC strains, potentially through the fission of the membrane during vesicle formation. In support of this hypothesis, decreased numbers of OMVs have been observed in the supernatant of $\Delta leoA$ mutants compared to the wild-type, using electron microscopy (7). This was further explored with the construction of a tat-GFP reporter system, where GFP was tagged with the tat-export sequence, allowing GFP to be transported across the inner membrane to the periplasm and subsequently packaged into OMVs. In this system, it was noted that the tat-GFP levels in vesicle fractions decreased by 50% in the supernatants of $\Delta leoA$ mutants, and increased in whole cell lysates (*Figure 1.11B left panel*) (5). This reduction strongly suggests that the GFP protein, in the absence of LeoA, is not secreted efficiently, but rather is being retained in the cell, supporting the notion that LeoA is involved in OMV formation and/or release in ETEC.

(A) LeoA has Structural Homology to Dynamin



(B) LeoA has a role in vesicle formation



Strain	LT (ng ml ⁻¹)	
	Periplasmic	Supernatant
wt ETEC	15.2 ± 4.5	199.7 ± 17.1
ETEC $\Delta leoA$	27.0 ± 7.0*	34.5 ± 4.3†
wt ETEC/p $leoA$ -FLAG	32.3 ± 9.2*	124.7 ± 8.2†
ETEC $\Delta leoA$ /p $leoA$ -FLAG	24.8 ± 12.0*	126.3 ± 10.1†
wt ETEC/p $leoA\Delta G$ -FLAG	20.5 ± 8.4	32.7 ± 5.7†
ETEC $\Delta leoA$ /p $leoA\Delta G$ -FLAG	7.2 ± 4.4*	31.2 ± 6.4†

*Significantly different from wt; Student's *t*-test, $P < 0.05$.

†Significantly different from wt; Student's *t*-test, $P < 0.005$.

Figure 1.11 LeoA is Bacterial Dynamin involved in Vesicle Formation in ETEC

(A) The crystal structure of LeoA shows strong structural homology to the dynamin superfamily (BDLP1, Dynamin 1 and GBP1, with PDB identifiers below corresponding structures), and includes a GTPase domain, a neck/trunk and a lipid binding paddle/tip. Adapted from (5). (B) Quantified Western blot (left) of the culture supernatant (vesicle fraction) shows that in the absence of LeoA the level of 'secreted' twin-arginine reporter GFP plasmid (*tat*-GFP) is reduced by approximately 50% which can be reversed with a supply of LeoA (pNDM220-*leoA*). Previously a reduction in the release (supernatant

fraction) of toxin (LT) was evident in the absence of LeoA, which was able to be rescued (table, right). Also, this release was dependent on the activity of the GTPase domain of LeoA. Adapted from (5, 7)

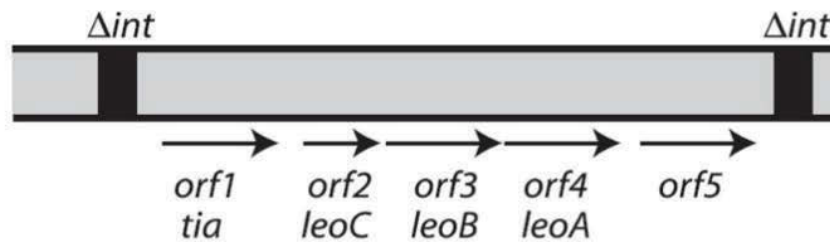
A role in OMV biogenesis and release is not the only postulated function of LeoA; a possible alternative role is in surface adhesion and infection of host cells. This stems from interactions noted using pull down assays with outer membrane proteins (OMPs), OmpX and OmpA, known *E. coli* virulence factors associated with these key infection stages (5, 7, 117-119). The importance of these interactions between LeoA and the OMPs remains undetermined, although in the absence of LeoA, there is a reduction of the levels of OmpX and OmpA in outer membrane fractions (7). Intriguingly there is yet another putative function for LeoA in motility. Deletion of LeoA promoted accumulation of a key component of the bacterial flagellum FliC in the periplasm and blocked its release to the membrane (120). This was supported by the observation that the motility of ETEC was dampened in the absence of LeoA (7). However, this phenotype may be a pleiotropic effect of the deletion, given that most motile pathogenic *E. coli* isolates do not encode LeoA (7). The common features amongst OMV biogenesis and release, OMP interactions and effects on components of the flagella suggest that the membrane is the main site of action by this bacterial dynamin.

The subcellular localisation of the LeoA was expected to help elucidate its function. LeoA was found to localise to the periplasmic compartment of ETEC by subcellular fractionation methods (5). Periplasmic localisation is quite perplexing as LeoA does not possess a signal peptide for export to the membrane, and this compartment of the cell reportedly lacks GTP, that would seemingly be required for functional activity of many dynamins (5, 109). Furthermore, dynamins typically bring about negative curvature to the membrane with which they are associated. To do this LeoA would be expected to localise to the outer membrane of the cell and to date no studies have looked for any bacterial dynamins here (6). Immunofluorescence microscopy was also used to investigate the localisation of LeoA with a punctate pattern seen, with a potential preference for cell poles and the mid cell (5). From all this, only further localisation studies will help bring to light the function of bacterial DLPs.

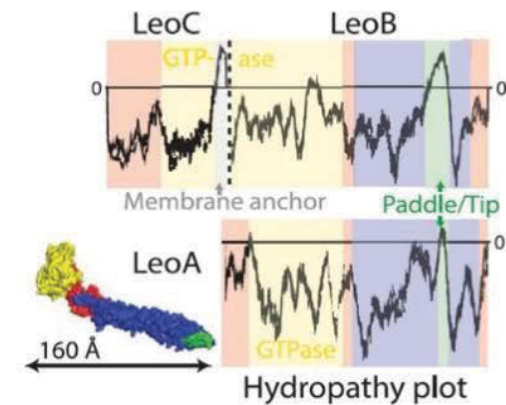
An extremely interesting discovery surrounding LeoA was the identification of two genes upstream of LeoA in the *tia* locus, now termed LeoC and LeoB (*Figure 1.12A*). These genes seemingly encode two parts of a full-length DLP (LeoCB) (*Figure 1.12C*) (5). The LeoC gene furthest upstream, encodes part of the GTPase domain and includes a putative hydrophobic insert that might be a membrane-anchor. The remainder of the GTPase domain is encoded by the downstream gene, LeoB, along with a strongly predicted hydrophobic membrane-binding region (the “paddle”) (*Figure 1.12B*). This is the first instance of a bacterial dynamin gene being split and the biological significance behind this is still not known.

With the knowledge that bacterial dynamins mostly come in pairs in the genome, a LeoCB fusion might function together with LeoA. Also, with the incomplete dynamin function of LeoA, and with no detectable GTP binding or hydrolysis and no lipid binding observed, it was speculated that LeoCBA all work together as a hetero-multimeric complex in ETEC (5). Curiously, the closest structural homologs to the ETEC LeoCBA proteins are found in the important human pathogen, *Helicobacter pylori*.

(A) Leo CBA operon



(B) Leo CBA hydrophobicity



(C) LeoC and LeoB form a full length dynamin

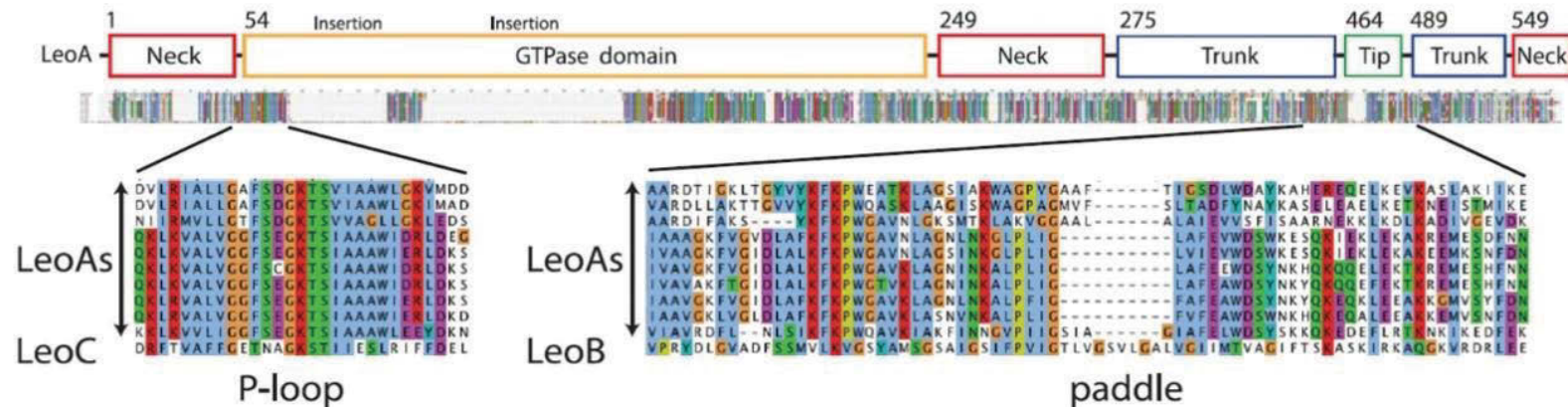


Figure 1.12 *leoCBA* Operon in *ETEC* Encodes Two Bacterial Dynamin-like Proteins

(A) Schematic diagram of the Tia locus in *E. coli* H10407 (ETEC), containing *leoC*, *leoB* and *leoA*. (B) Surface plot and hydropathy plots of *LeoA* coloured by domains: yellow, GTPase domain; red, neck domain; blue, trunk domain; green, putative paddle region. TMpred highlights predicted transmembrane domains for both *LeoA* and *LeoB* (green). (C) *LeoC* and *LeoB* align well against the entire length of *LeoA*. *LeoC* contains the GTPase domain and *LeoB* contains the paddle for lipid binding. It seems that ETEC encodes two dynamins in tandem, with one split into two chains, termed *LeoC* and *LeoB*. Sequences aligned: *LeoA* (*E. coli* ETEC H10407), WP_011717023.1 (*Shewanella* sp. ANA-3), WP_007214706.1 (*Bacteroides cellulosilyticus*), WP_001006159.1 (*Helicobacter pylori*), WP_001006151.1 (*Helicobacter pylori*), WP_014535968.1 (*Helicobacter pylori*), WP_000787447.1 (*Helicobacter pylori*), WP_001006093.1 (*Helicobacter pylori*), WP_000787451.1 (*Helicobacter pylori*), WP_005966123.1 (*Fusobacterium periodonticum*). Adapted from (5).

1.6 *HELICOBACTER PYLORI* AS A PATHOGEN

Helicobacter pylori is a human-specific gram-negative pathogenic bacterium that survives within the gastric mucosal layer in the stomach of over half of the world's population. It is one of the only bacteria that thrives in the extremely harsh environment of the stomach in the low pH found there.

The stomach is a muscular organ that is a part of the gastro intestinal tract (GIT) that extends from the oesophagus and empties into the small intestine. The epithelium of the stomach is covered in a thick mucus layer that is thought to protect the cells from the harsh acid environment. The organ itself does not have a consistent cellular composition and is made up of subregions that differ in their cellular structure with three main glandular regions; the fundus (upper), corpus (middle) and antrum (lower). As seen in Figure 1.13, *H. pylori* colonises very specific areas once it enters via the oral route through the oesophagus that are thought to be a little more neutral in pH; firstly, the mucosal layer specifically in the corpus and atrium glands (121, 122). Interestingly, it is well established that these areas of the stomach are quite different in composition, however the important factors that promote and harbour growth of *H. pylori* in the stomach are unknown. The corpus region contains a range of different cell types including

but not limited to pit mucous cells, which produce the mucous layer both inside and outside the glands, and HCl-secreting parietal cells. The antrum does not have the same variety of cells as seen in the corpus, with the specific lack of the ability to secrete acid. However, the antrum still has mucous cells in the pit as well as the neck to protect the epithelium cells. In some animal models, it seems that *H. pylori* shows more initial growth in the antrum and therefore there are higher levels of infection detected in this sub region, which may just be attributed to the anatomy of the stomach (123).

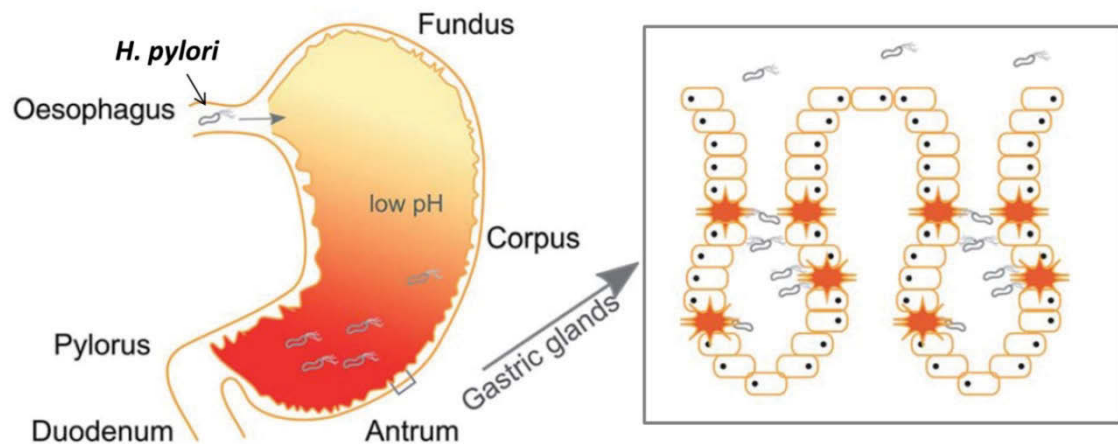


Figure 1.13 *Helicobacter pylori* causes Infection in the Stomach

H. pylori enters the environment it will colonise, the stomach, through the oesophagus, colonising the corpus and the atrium with preference for the atrium. Each of the regions of the stomach have different cellular compositions, but a close view of the gastric glands of the atrium (right) shows that *H. pylori* colonises both deep within the glands as well outside in the mucus layer that overlays the cells. Once *H. pylori* has colonised the gastric epithelium, damage such as apoptosis can occur (stars). Adapted from (122).

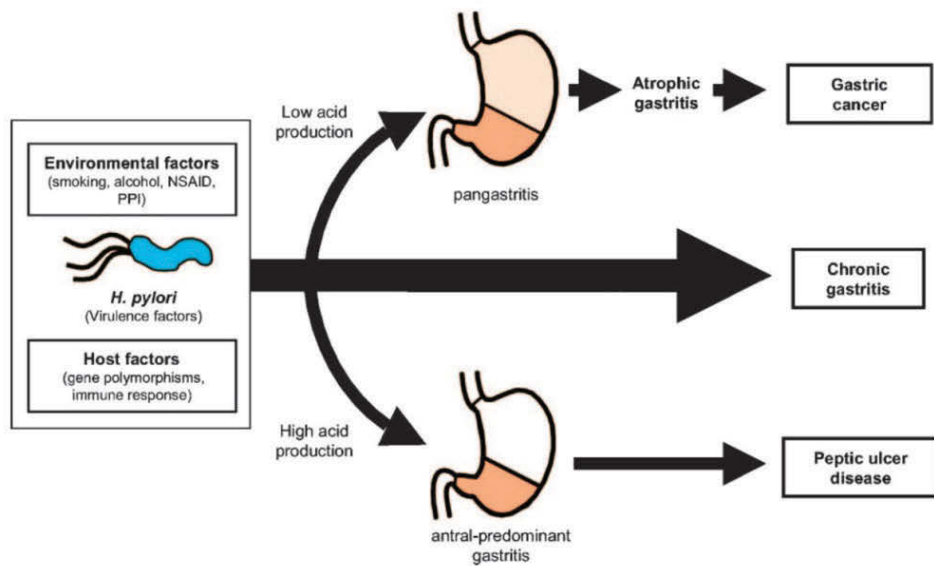
1.6.1 H. PYLORI CAN CAUSE SIGNIFICANT DISEASE IN HUMANS

In 1983, Robin Warren and Barry Marshall identified a curved bacillus organism in almost all active chronic gastritis patients screened (124). These unidentified organisms were found in over half of their routine gastric biopsy patients in large enough numbers to be seen using traditional histology and light microscopy. Using morphology, they were seen to resemble *Campylobacter* species, were isolated using *Campylobacter* techniques and were concluded to potentially be a part of the genus *Spirillum* (124). Warren and Marshall proceeded to term these organisms *Campylobacter pyloridis*, based on their location in the stomach and their morphology. However, since then many other examples of bacteria with similar characteristics have been noted in a number of animals, with 16S RNA sequencing indicating they were not a part of the *Campylobacter* species and therefore the genus *Helicobacter* was established (125). *H. pylori*, as it is now termed, has been identified as the causative agent of duodenal and gastric ulcers (126). Furthermore, chronic infections of *H. pylori* have been shown to increase the risk of developing gastric cancer and mucosa-associated tissue lymphoma (MALT) and for this reason it has been classified as a Type 1 Carcinogen by the World Health Organisation (127).

There are a number of gastroduodenal diseases that are associated with *H. pylori* infection with the colonisation of the gastric mucosal layer resulting in the development of chronic gastritis in all infected individuals (128, 129). The clinical outcome of infection is highly dependent on many variables such as the genotype of *H. pylori*, genetic predisposition and environmental factors (*Figure 1.14A*). Due to these different types of variables, only a subset of these patients will progress to further disease, including ulcer disease and gastric cancers such as gastric adenocarcinoma and MALT lymphoma (129) (*Figure 1.14B*).

The different diseases caused by *H. pylori* are highly dependent on a number of different pathogenic features that remain quite controversial and are consistently being explored in the literature. These features can include the ability for host colonisation and persistence as well as a number of virulence factors that *H. pylori* releases to sustain disease.

(A)



(B)

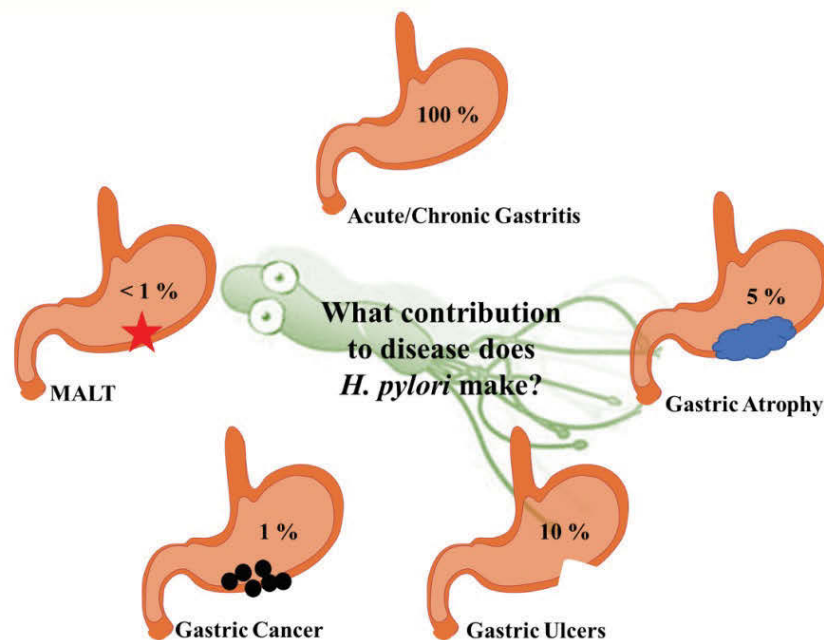


Figure 1.14 *H. pylori* Infection can lead to Many Diseases in Humans

(A) Graphical representation of the factors that contribute to the different pathologies and the outcomes associated with *H. pylori* infection including gastric cancer, chronic gastritis and gastric ulcers. Adapted from (130). (B) There are a range of diseases that *H. pylori* contributes to in varying degrees from 100% of cases of gastritis to 10% of gastric ulcers cases to 1% of gastric cancer cases (129).

1.6.2 HOW DOES *H. PYLORI* DEAL WITH THE ACIDIC ENVIRONMENT OF THE STOMACH?

H. pylori not only survives but also thrives in arguably one of the harshest environments found in the body, the stomach. To do this it has evolved a number of defence mechanisms, seen in *Figure 1.15*, that allow it to combat the stresses that come with living in such a niche including; motility, chemotaxis and acid tolerance mechanisms (122). The obvious environmental factor that *H. pylori* overcomes consistently is the acidic nature of the stomach however there are other environmental factors, like diet and antibiotics that the bacterium must also overcome. An important dietary factor that *H. pylori* is exposed to is high salt which in turns leads to exposure to oxidative stress (131). Interestingly, it has been shown that *H. pylori* exposed to high salt has an increased capacity to cause cancer in a mouse model (132). Furthermore, a high salt diet in humans has been stated as a contributing risk factor for gastric cancer (133).

However, the principal environmental stress that *H. pylori* combats daily, more successfully than any other bacterium to cause and sustain infection, is acid. The most studied acid adaptation pathway in *H. pylori* is the urease pathway, despite there being several other ways the bacterium can overcome acid stress such as chemotaxis, motility and intracellular pH homeostasis systems like AmiE/AmiF (122, 134, 135).

The urease pathway is utilised by *H. pylori* to neutralise the HCl (hydrochloric acid) in close proximity to itself. The urease enzyme, built from two individual protein subunits UreA and UreB, makes up 10% of the entire protein mass of the bacterium, highlighting the importance of this enzyme and this pathway to its survival in this environment (136). In this pathway, host-derived urea diffuses into the cytoplasm through the proton dependent pump, UreI (137). There it is converted into ammonia and carbon dioxide (CO₂) by the urease enzyme. Both molecules then diffuse into the periplasm where CO₂ mostly diffuses out of the cell, but some is retained by the cell and is converted by the periplasmic enzyme α -carbonic anhydrase (α CA) into bicarbonate (138). This bicarbonate forms a buffer that helps maintain the periplasmic pH at around 6.1. The other product of the urease pathway, ammonia, diffuses through to the periplasm where it reacts with the incoming H⁺ ions (HCl acid) and neutralises them, creating a zone of neutral pH

directly around *H. pylori* (Figure 1.15A) (139, 140). This pathway is the first line of defence that *H. pylori* uses to combat short exposures to acid stress.

In addition to the urease pathway, *H. pylori* has the ability to actively and rapidly move away from regions of low pH using chemotaxis and motility. The bacterium uses these two mechanisms in combination to look for an optimal place to set up stable long-term colonisation (122). The chemotaxis signalling system in *H. pylori* is rather simple and consists of only 4 receptors that can detect multiple signals including acid, urea, amino acids and metals (Figure 1.15B) (123, 141-145). This simple system underlies many of the bacterium's environmental responses but most importantly is the driving force behind the directionality in the bacterial swimming response. A group of uni polar flagella that spans both membranes establishes chemotactic motility in *H. pylori* (143). The flagella are built from approximately 30 proteins that make up the different flagella sections including the filament, the body and the hook that joins them together; these are described in more detail in Chapter VI of this thesis (146, 147). The swimming motion is due to the rotation of the flagella, which is powered by two motor proteins, MotA and MotB (148). Research has focussed on understanding the flagella specifically with the thought that it could form a potential drug target to treat *H. pylori* infection. This is because with no flagella, *H. pylori* loses its ability to effectively colonise the stomach.

H. pylori also utilises a two-component regulatory system, ArsR/S, to sense acid and respond accordingly. This system is activated by low pH and controls the transcription of many genes (109 genes), including the urease genes (135). Other genes that are activated by this system in the presence of low pH are the ammonia producing pathways AmiE and AmiF, as well as the ETEC LeoCBA dynammin homologs (135).

Acid adaptation is an essential feature that *H. pylori* possesses to promote successful colonisation, growth and survival in the harshest environment found in the body for bacteria. At the core of these mechanisms is the membrane of *H. pylori*, which is possibly the most important defensive wall available to combat this stress with all the above processes carrying out their acid adaptive function at both the inner and the outer membrane.

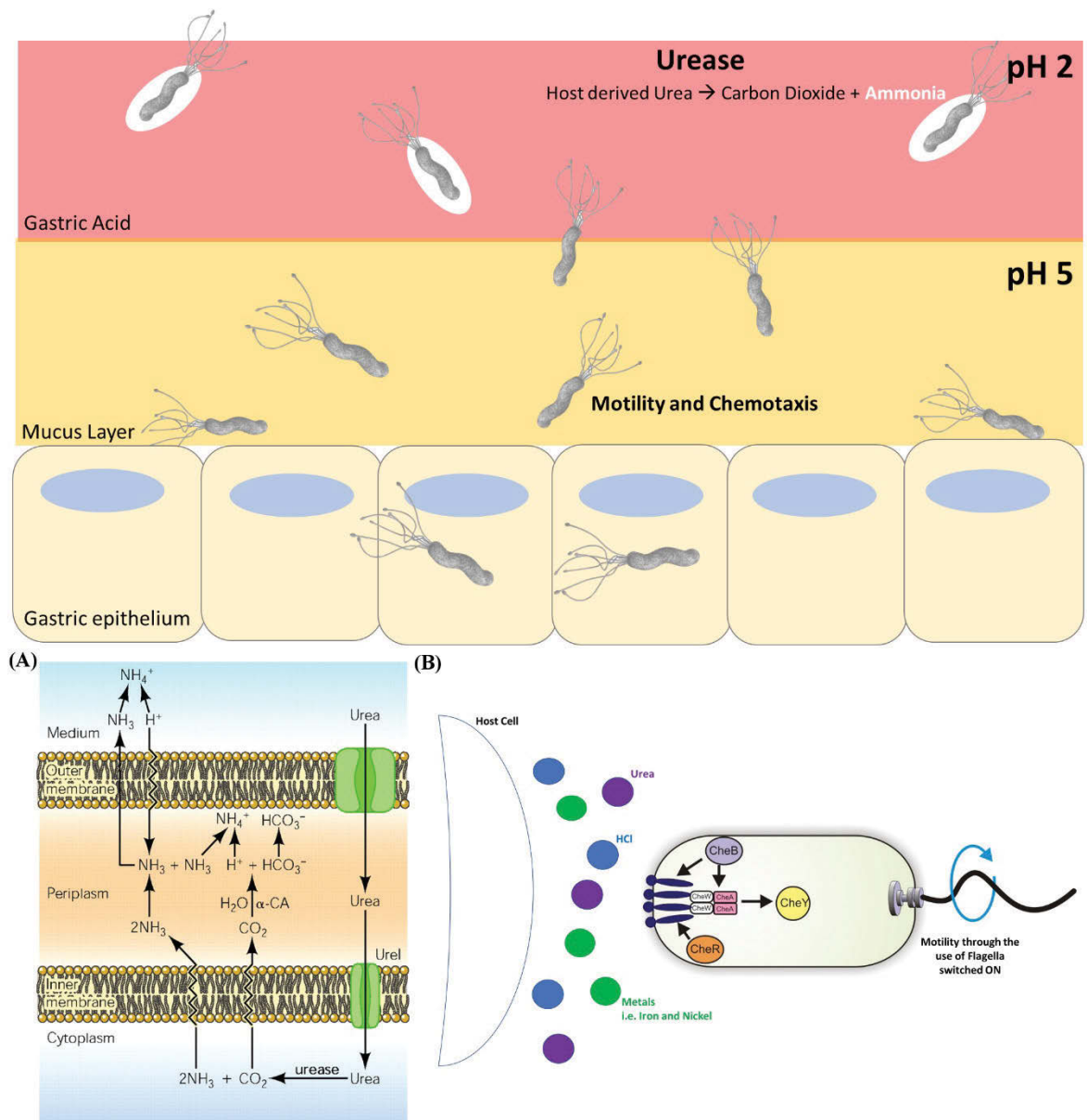


Figure 1.15 Acid Adaptation Mechanisms of *H. pylori*

*Schematic representation of H. pylori's key acid adaptation mechanisms needed for infection. In response to gastric acid, the urease pathway neutralizes the pH in the proximity of the bacterium. (A) Host derived urea passes through the membrane, is converted by urease to carbon dioxide and ammonia which diffuses back to the periplasm to play a role in periplasmic homeostasis and then to the extracellular milieu Adapted from (149). The chemotactic flagella mediated motility towards the epithelium through the mucus layer brings the bacterium in close proximity of the host cell. (B) *H. pylori* senses urea, acid and metals such as Iron and Nickel, to then switch on internal signalling pathways through CheB, CheR and CheY, to activate MotA and MotB to begin flagella motion (122). Adapted from (150)*

1.6.3 KEY FEATURES IN THE PATHOGENESIS OF *H. PYLORI* INFECTION

There are three essential aspects to *H. pylori* pathogenesis; colonisation, adherence and virulence factors. Without these features, *H. pylori* is not able to cause and sustain infection in the population.

Successful colonisation requires two features of *H. pylori* and they are shape and motility (*motility was discussed previously in Section 1.6.2*). The characteristic helical shape of *H. pylori* is both crucial for motility as well as necessary for colonisation. The helical shape is determined by modifications to the peptidoglycan layer in the periplasm of the cell by a number of different proteins, but mainly the Csd 1-6 proteins. It has been shown that in *H. pylori* *csd1* curved rod mutants, *csd3* variably curved rod mutants and *csd4* mutants which exhibit extremely straight shapes (*Figure 1.16A*), were consistently outcompeted in a mouse model by WT strains in colonisation stages. This suggests that the cell shape and/or cell wall changes absent in these mutants are uniquely required during stomach colonisation (151-153).

It is known that adhesion is an essential stage in infection of pathogenic bacteria and therefore *H. pylori* encodes over 30 outer membrane proteins with putative involvement in adhesion that differ from strain to strain, potentially to adapt to the changing environment in the host (154). For *H. pylori* to cause infection it must adhere to the gastric epithelium cell and does so through expression of several well studied outer membrane adhesins, including Blood group antigen binding adhesion protein (BabA/B), sialic acid binding protein (SabA), adherence associated lipoproteins (AlpA/B), outer inflammatory protein (OipA) and Helicobacter outer membrane protein (HopZ) (154). These adhesins establish the initial tight contact between the bacteria and the gastric epithelium as seen in *Figure 1.16B, middle*. BabA is the initial point of contact for *H. pylori* to the host cell and binds to the host cell Lewis B blood group carbohydrate (155, 156). This host cell interaction has been shown to be essential for the delivery of one of the most important virulence factors, cag pathogenicity island (PAI) type IV secretion systems (T4SS), which is commonly encoded within the *H. pylori* genome (154, 157, 158). SabA binds to the host sialyl-dimeric-Lewis X glycosphingolipid which acts as a receptor, interestingly induced by *H. pylori* infection (159). Increased inflammation caused by the binding of *H. pylori* to the host cell by BabA, increases the presence of the SabA receptor and promotes

further binding of the bacterium to the host cell. The presence of this adhesin is commonly associated with more devastating disease such as gastric cancer due to this mechanism of binding (160).

Despite *H. pylori* producing multiple virulence factors, like the adhesins mentioned above, the two most well studied virulence factors are CagA and vacuolating toxin A (VacA). CagA is encoded on the *cag* pathogenicity island, in approximately 50-70% *H. pylori* strains (161, 162). Disease outcomes of *H. pylori* infection can vary in severity depending on the different polymorphisms of CagA present in the infecting strain (107, 161). The PAI is horizontally acquired by the bacterium and encodes for a putative type IV secretion system (T4SS). The T4SS is a multi-protein complex that spans both the inner and outer membrane and in the unique case of *H. pylori* is used to inject CagA into the host gastric epithelial cell (*Figure 1.16B, right side*) (163, 164). Once inside it can act directly on a range of different cellular processes in both an unphosphorylated and phosphorylated state by disrupting signal transduction and gene expression (165). These processes include but are not limited to the formation of tight junctions, polarity, proliferation and the induction of the inflammatory response (166-169).

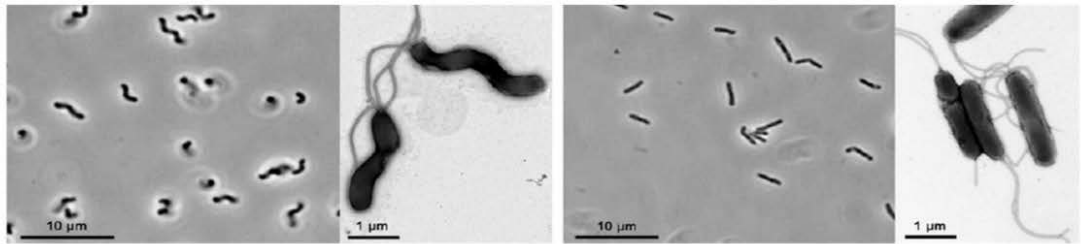
VacA is a pore forming toxin which interacts directly with host cell receptors to trigger cellular vacuolation and is the main mechanism by which *H. pylori* induces host cell apoptosis (*Figure 1.16B left side*) (169, 170). VacA is a type V autotransporter, consisting of a signal peptide for translocation to the cell surface and it also undergoes a series of post translational modifications to form the active and the binding subunits of the protein to carry out function in the host cell (169, 171, 172). The release of VacA from *H. pylori* has been attributed to both active transport via a channel that the C-terminus of the protein forms in the outer membrane, as well as the formation of outer membrane vesicles that package the toxin for delivery to the host (111) (*Figure 1.16B left side*).

Despite no known mechanism for the formation of OMVs in *H. pylori* (*Figure 1.16C*), it is well known that OMVs are enriched with proteins and other cellular components like DNA that enhance the invasive abilities of the bacterium, like *H. pylori*. This in turn promotes efficient internalisation of virulence factors such as VacA at the host interface (*Figure 1.16B left side*) (116, 173). There are number of proposed models of OMV biogenesis, mainly centred around modifications to the peptidoglycan that promote

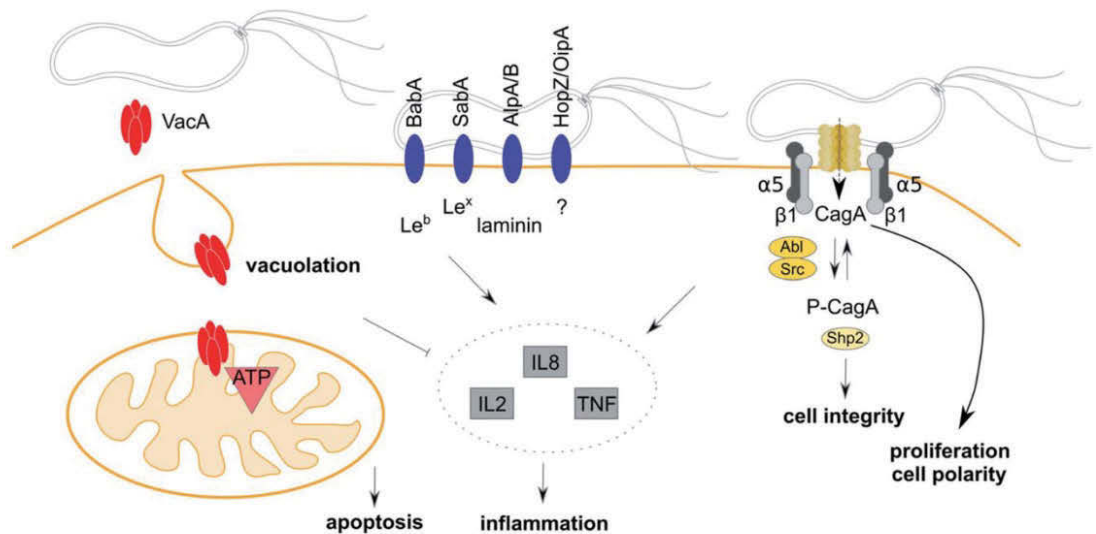
changes to the membrane (174). Also, it would be expected that enrichment of molecules, like LPS and phospholipids that can increase membrane curvature would also induce OMVs formation (175). The integrity of the membrane structure is vital for the formation of OMVs with a range of proteins shown to be involved in the maintenance of this, the most well-known being the highly conserved proteins in the *tol-pal* cluster (176). The mutation or absence of the *tol-pal* cluster has been shown to cause numerous phenotypes including but not limited to, increase vesiculation, altered cell morphologies and changes to motility not only in *H. pylori* but in a range of Gram negative bacteria (174, 177-180). Interestingly, it was found that *tolB* and *pal* mutations in *H. pylori* brought about morphological changes in cell shape as well as affecting membrane integrity (179). Specifically, when *tolB* was mutated there was an increase in OMV production compared to WT with minimal change to the protein content inside the vesicles (179). Therefore, the structure and integrity of the cell membrane is a distinguished and imperative factor in the successful release of OMVs from the cell.

From these examples of cell shape, adhesion, delivery of virulence factors and OMVs, it is clear that the integrity of the *H. pylori* membrane is essential for pathogenesis. It is well established that the *H. pylori* membrane is highly important in all stages of infection, from surviving the acidic environment to adhering and colonising the cells of the stomach. Membrane dynamics and remodelling events are therefore important in the protection of the cell as well as contributing to the pathogenicity of the bacteria, i.e. through the formation of OMVs. It is exciting to postulate that the conservation of the dynamin superfamily in this organism could mean that dynamin like proteins are playing a key role in the remodelling of the membrane in the human pathogen *H. pylori*, a function reminiscent of the classical dynamins.

(A) Cell Shape Changes



(B) Adhesins and Virulence Factors



(C) Outer Membrane Vesicles

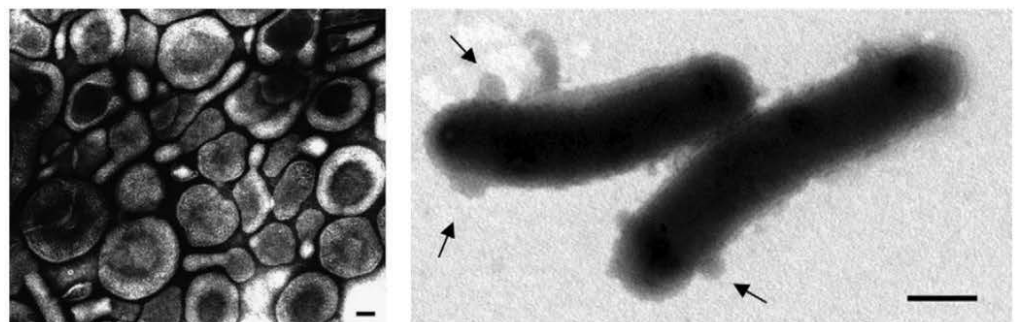


Figure 1.16 *H. pylori* Pathogenesis is Multifactorial

(A) Phase and transmission electron microscopy (TEM) highlight cell shape changes in the absence of *csd4* (right) compared to WT (left) which in turns affects colonisation. Adapted from (151). (B) *H. pylori* possesses multiple factors that assist in adherence to host cells including *BabA*, *SabA*, *AlpA/B* and *HopZ/OipA*. This adherence is key to the delivery of *CagA* via the T4SS, which induces changes to signalling pathways that decrease cell integrity. Independent of attachment, *VacA* is delivered to the host and

induces cell apoptosis. Adapted from (122). (C) Purified outer membrane vesicles (left) show evidence of a bilayer membrane and come in all different sizes. H. pylori uses 'blebbing' or production of OMVs (right arrows) to deliver toxins to the host and form biofilms on the host cells. Adapted from (173).

1.7 HYPOTHESIS AND AIMS OF THIS THESIS

The investigation of the homologs of the ETEC LeoCBA dynamins in *H. pylori* will bring a wealth of knowledge to further our understanding of *H. pylori* infection with many aspects highly understudied. Particularly, the importance of the membrane and how it is maintained and/or remodelled by *H. pylori* to ensure continual infection remains unknown. This research will also help further elucidate the role that these bacterial dynamins (LeoCBA) may have within ETEC. With the Leo dynamins seemingly involved in membrane vesicle formation, studying their function in *H. pylori*, a vesicle forming pathogen, may help provide an understanding of both dynamin function in bacteria and potentially OMV biogenesis, two understudied and poorly understood areas of research. However, with the eukaryotic dynamins possessing similar structural features but characterised to carry out a diverse range of functions, it would be naïve to begin to assign a vesicle formation function to the *H. pylori* dynamins based solely on homology. The knowledge gained by this study could be applied beyond the scope of just *H. pylori* and ETEC dynamins, mainly due to the fact that putative dynamins exist across many bacterial species with a number of different functions already being suggested for this family of proteins. Consequently, furthering our understanding of bacterial dynamins may translate to the identification of novel therapeutic targets in not only the important human pathogen *H. pylori* but potentially other bacteria also.

Therefore, to address the large paucity of knowledge on bacterial dynamins, this thesis had the overarching theme to investigate the presence of dynamins in the human pathogen *H. pylori*. Under this theme was the following hypothesis; that dynamin-like proteins are highly conserved in *H. pylori*, that they are carrying out a key function at the membrane of the cell, and that this function is important in pathogenesis, such as a role in vesicle formation or membrane integrity.

This hypothesis was investigated through the following aims:

- Aim One: To investigate the conservation of dynamin-like proteins in *H. pylori* (Chapter III)
- Aim Two: To characterise the expression and biochemical features, such as localisation, lipid binding and interactions of dynamin-like proteins in *H. pylori* (Chapter III and IV)
- Aim Three: To explore the function of dynamin-like proteins in *H. pylori* (Chapter V and VI)

Minimal studies have been conducted into the conservation of dynamins in bacteria and no analysis to date has been performed into the conservation of dynamins specifically in *H. pylori*. Hence, to begin the investigation into this conservation, a comprehensive search was conducted using psi-BLASTp of the bacterial phyla. Along with this, a number of bioinformatics tools, such as an iterative search tool JackHMMER, membrane prediction software TMHMM and structural predication software Phyre2, were used to establish the presence and nature of *H. pylori* dynamin-like protein specific conservation as well as gene and protein characteristics (Aim One).

Bacterial dynamins have several defining biochemical characteristics such as self-interaction, membrane localisation and lipid binding capabilities. With no prior work on *H. pylori* dynamins, these characteristics were studied using a range of biochemical assays. Firstly, *in vivo* expression of the *dlp* operon was confirmed using both RT-PCR as well as Western blot. Localisation of these proteins *in vivo* was investigated using both subcellular fractionation as well as immuno- florescence microscopy (IFM). Lipid binding capabilities were established using *in vitro* binding assays analysed by microscopy as well as by co-sedimentation assays. Finally, interactions were studied using the Bacterial Two Hybrid System (BacTH) as well as co-pelleting and gel filtration methods (Aim Two).

The conserved function of bacterial dynamins is still speculative, mainly because they are highly understudied. Interestingly, functions are already quite diverse for the bacterial dynamins studied to date; from a role in cell division, to vesicle production to membrane

integrity, reminiscent of their eukaryotic counterparts. Following on from the biochemical characterisation, the focus of this research was to investigate the function of dynamins in *H. pylori*. This was accomplished through the creation of a complete knockout of the *dlp* operon in *H. pylori* and phenotypes tested. These included growth, membrane integrity, acid tolerance as well as acid growth. Each of these phenotypes was analysed using growth assays as well as live microscopy. Furthermore, a global proteome exploration was conducted using quantitative unlabelled mass spectrometry (SWATH-MS) to investigate the effect that the absence of dynamin has on the proteome to try and elucidate a pathway that dynamin may interact with (Aim Three).

From these aims, the data presented in this thesis provides an investigation into the presence of dynamin-like proteins in the human pathogen *H. pylori*.

CHAPTER II

GENERAL METHODOLOGY

2.1 BACTERIAL STRAINS AND PLASMIDS

Tables 2.1 and Table 2.2 list the general bacterial strains and plasmids used in this work, which were already available. Bacterial strains and plasmids produced in this work will be described where appropriate.

All *H. pylori* strains were stored as late log phase cultures with 20% (v/v) glycerol at -80°C in useable aliquots of 0.3 mL. All *E. coli* strains were stored as stationary phase cultures with 16% (v/v) glycerol at -80°C.

Table 2.1 General Bacterial Strains used in this Work

Species	Strain	Relevant Genotype	Reference/Source
<i>H. pylori</i>	26695	Wild-Type	Mitchell, H (UNSW)
<i>H. pylori</i>	B128.7.13	Wild-Type B128 passaged for gerbil adaption	Ferrero, R (Hudson Institute)
<i>E. coli</i>	DH5α	<i>fhuA2</i> Δ(<i>argF-lacZ</i>)U169 <i>phoA</i> <i>glnV44</i> Φ80 Δ(<i>lacZ</i>)M15 <i>gyrA96</i> <i>recA1</i> <i>relA1</i> <i>endA1</i> <i>thi-1</i> <i>hsdR17</i>	Duggin, I (Lab)
<i>E. coli</i>	BL21-AI	F- <i>ompT</i> <i>hsdSB</i> (rB-, mB-) <i>gal dcm</i> <i>araB::T7RNAP-tetA</i>	Sigma-Aldrich, USA

Table 2.2 General Plasmids used in this Work

Plasmid	Size	Relevant Purpose	Selection	Reference/Source
pHis17		Over Expression and C terminal tagging of protein with 6x His Tag	Amp	Lowe, J (LMB, Cambridge)
pSKFII		Contains the <i>kan-sacB</i> cassette for <i>H. pylori</i> gene deletion	Kan	Ferrero, R (Hudson Institute)
pBSII		General cloning plasmid backbone used in <i>H. pylori</i> transformation	Amp	Labbate, M (UTS)
pKT25		BacTH - C terminal tagging of T25 reporter gene	Amp	Bottomly, A (UTS)
p25N		BacTH - N terminal tagging of T25 reporter gene reporter gene	Amp	Bottomly, A (UTS)

pUT18C	BacTH - C terminal tagging of T18 reporter gene	Kan	Bottomly, A (UTS)
pUT18	BacTH - N terminal tagging of T18 reporter gene	Kan	Bottomly, A (UTS)

2.2 BACTERIAL GROWTH MEDIA

Bacterial growth media used in this work is listed below in *Table 2.3*. Manufacturers are listed in *Table 2.4*.

Table 2.3 Bacterial Growth Media

Media	Composition*
Brucella Broth	28 g Brucella Broth base, 10% FBS in 1 L of MQW
LB	25 g LB Broth base in 1 L of MQW
Blood Agar	28 g Blood Agar Base No2, 8% defibrinated horse blood, 1.6 mL DENT in 1 L of MQW
LB Agar	25 g LB Broth base, 1.5% agar in 1 L of MQW
SOB	2 % (w/v) peptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl ₂ , 10 mM MgSO ₄ in MQW
SOC	2 % (w/v) Tryptone, 0.5 % (w/v) Yeast Extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl ₂ , 20 mM Glucose in MQW
X-Gal Agar	LB agar as above with 40 µg/ml X-gal (5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside), 0.5 mM IPTG (isopropyl-β-D-thiogalactopyranoside)

*All media were made up in ROW and sterilized by autoclaving or 0.2 µm filtration, or from sterilised components

2.3 CHEMICALS, REAGENTS AND SOLUTIONS

Most chemicals and reagents used in this work are outlined below in *Table 2.4*. They were of analytical reagent (AR) grade and obtained from the companies listed. Commonly used aqueous buffers and solutions in this work are listed below in *Table 2.5*.

Table 2.4 Commonly Used Chemicals in this Work

Chemical/Antibiotic	Source
Acetonitrile	Sigma-Aldrich, USA
Acrylamide	GE Healthcare, Sweden
Agarose	Bioline, Australia
Ampicillin	Genlantis, USA
Arabinose	Sigma-Aldrich, USA
BBL™ Brucella Broth Base	BD Biosciences, USA
BioLab Broad Range Protein Standards	New England BioLabs, USA
Bio-Rad Pre-Stained Protein Standards	Bio-Rad, USA
Bradford Assay Reagent	Sigma-Aldrich, USA
C ₇ BzO	Sigma-Aldrich, USA
Chemiluminescence Kit	Thermo-Fisher Scientific, USA
Coomassie Brilliant Blue R-250	Sigma-Aldrich, USA
Defibrinated Horse Blood	Thermo-Fisher Scientific, USA
Difco™ LB Broth Base (Luria-Bertani)	Thermo-Fisher Scientific, USA
FM4-64	Thermo-Fisher Scientific, USA
Foetal Bovine Serum (FBS)	Gibco, USA
Gel-Red	Biotium
Glycerol	Sigma-Aldrich, USA
GTP	Sigma-Aldrich, USA
<i>Helicobacter</i> Selective Supplement (DENT)	Thermo-Fisher Scientific, USA
Kanamycin	Sigma-Aldrich, USA
Lysozyme	Sigma-Aldrich, USA
Ni ²⁺ His-Trap HP Column	GE Healthcare, Sweden
Oxoid™ Agar (Bacteriological Agar No.1)	Thermo-Fisher Scientific, USA
Oxoid™ Blood Agar Base No.2	Thermo-Fisher Scientific, USA
Poly-L-lysine	Sigma-Aldrich, USA
Protease Inhibitor Cocktail	Roche, Switzerland
Protein Concentrator (MW 10,000) (2mL)	GE Healthcare, Sweden
Protein Concentrator (MW 10,000) (50mL)	GE Healthcare, Sweden
Sucrose	Sigma-Aldrich, USA
Sybr Green JumpStart Taq ReadyMix	Sigma-Aldrich, USA
Sytox Green	Thermo-Fisher Scientific, USA
Trifluoroacetic acid	Sigma-Aldrich, USA
Trypsin Sequencing Grade	Promega
Tween 20	Sigma-Aldrich, USA

Table 2.5 Aqueous Buffers and Solutions used in this Work

Buffer/Solution	Composition*
1 x PBS	150 mM Potassium Phosphate, 50 mM Tris-Cl, pH 7.4
1 x TBS	150 mM NaCl, 50 mM Tris-Cl, pH 7.4
2 x PBS	300 mM Potassium Phosphate, 50 mM Tris-Cl, pH 7.4
2 x TBS	300 mM NaCl, 50 mM Tris-Cl, pH 7.4
4 x SDS-PAGE loading buffer (SDS-LB)	250 mM Tris-Cl, 8% (w/v) SDS, 40% (v/v) Glycerol, 20% (v/v) 2-β Mercaptoethanol, 0.4% (w/v) Bromophenol Blue, pH 6.8
Antibody Buffer	1 x PBS, 0.1% (v/v) Tween-20, 2% (w/v) BSA
Blocking Solution	5% (w/v) Diploma skim milk powder in TBST
Cold Lysis Buffer	0.5 mM MgCl ₂ , Protease Inhibitors
Coomassie Destain	20% (v/v) Ethanol, 10% (v/v) Acetic acid
Coomassie Stain	50% (v/v) Ethanol, 10% (v/v) Acetic acid, 0.5% (w/v) Coomassie brilliant Blue R
Elution Buffer	20 mM Tris-HCl, 0.5 M NaCl, 1 M Imidazole, 1 mM DTT, pH 8.0
Elution Solvent	75% (v/v) Acetonitrile, 25% (v/v) MS Loading Solvent
GTE Buffer	50 mM Glucose, 10 mM EDTA, 20 mM Tris-Cl, pH 8.3
Liposome Buffer	50 mM Tris-Cl, 100 mM NaCl, 1 mM EDTA, 1 mM NaN ₃ , pH 7.5
M63	0.2% (w/v) (NH ₄) ₂ SO ₄ , 0.14% (w/v) KH ₂ PO ₄ , 0.5 mg FeSO ₄ .7H ₂ O, 1 mM MgSO ₄ .7H ₂ O, 10 mL of 20% Maltose, 2 mL of 0.05% Vitamin B1 (thiamin), 50 µg/mL Amp, 25 µg/mL Kan
MS Loading Solvent	2% (v/v) Acetonitrile, 0.2% (v/v) TFA
PM2	70 mM Na ₂ HPO ₄ .12H ₂ O, 30 mM NaH ₂ PO ₄ H ₂ O, 1 mM MgSO ₄ , 0.2 mM MnSO ₄ , 100 mM 2-β Mercaptoethanol, pH 7.0
Ponceau Stain	0.1% (w/v) Ponceau, 5% (v/v) Acetic acid
Reducing Solution	5 mM Tributylphosphine, 20 mM Acrylamide in NH ₄ HCO ₃
Refolding Buffer	20 mM Tris-HCl, 0.5 NaCl, 20 mM Imidazole, 1 mM DTT, pH 8.0
Resuspension Buffer	20 mM Tris-HCl, pH 8.0
SDS-PAGE Running Buffer	25 mM Tris, 190 mM Glycine, 0.1% (w/v) SDS, pH 8.3
Solubilisation Buffer	20 mM Tris-HCl, 0.5 M NaCl, 5 mM imidazole, 6 M Guanidine Hydrochloride, 1 mM DTT, pH 8.0
Sonication Buffer	2 M Urea, 20 mM Tris-HCl, 0.5 M NaCl, 2% Triton X-100, pH 8.0
Sucrose Buffer	20% sucrose, 1 mM EDTA, 20 mM Tris-Cl, pH 7.6, Protease Inhibitors
TAE	40 mM Tris, 20 mM Acetate, 1 mM EDTA, pH 8.6
TBE	89 mM Tris, 89 mM Boric acid, 2.5 mM EDTA, pH 8.3
TBST	150 mM NaCl, 50 mM Tris-Cl 0.01% (v/v) Tween 20, pH 7.4
Transfer Solution	50 mM Tris, 384 mM Glycine, 20% (v/v) Methanol

Transformation Buffer	250 mM KCl, 15 mM CaCl ₂ , 10 mM PIPES, 55 mM MgCl ₂
Trypsin Solution	12.5 ng/μL Trypsin in NH ₄ HCO ₃
UTC7	7 M Urea, 2M Thiourea, 1% (w/v) C ₇ BzO
Washing Buffer	20 mM Tris-HCl, 0.5 NaCl, 20 mM Imidazole, 6 M Urea, 1 mM DTT, pH 8.0

*All solutions were made up in ROW and sterilized by autoclaving or 0.2 μm filtration, or from sterilised components

2.4 BACTERIAL GROWTH CONDITIONS

2.4.1 NORMAL GROWTH CONDITIONS OF *H. PYLORI*

A frozen glycerol stock (300 μL) of the required *H. pylori* strain (*Table 2.1*) was thawed at RT and 50 μL was used to inoculate Blood agar (*Table 2.3*) as a patch. Plates were left for 48 h in a microaerophilic environment (CampyGen 2.5L) at 37 °C. This patch growth was gently lifted using a sterile plastic loop and inoculated into 25 cm² culture flasks containing 5mL of Brucella broth (*Table 2.3*), left for a further 18-24 h with shaking (120 rpm) in a microaerophilic environment (CampyGen 2.5L) at 37 °C.

2.4.2 NORMAL GROWTH CONDITIONS OF *E. COLI*

A frozen glycerol stock scrapping of the necessary *E. coli* strain was used to inoculate LB Agar (*Table 2.3*) (with appropriate antibiotics) and incubated overnight at 37°C. A single colony was then used to inoculate LB Broth (*Table 2.3*) (with appropriate antibiotics) for a further incubation overnight at 37°C with shaking (150 rpm).

2.4.3 PREPARATION AND TRANSFORMATION OF *E. COLI*

The *E. coli* strain was prepared as outlined in *Section 2.4.2*. The culture (20 μL) was streaked out on LB agar (*Table 2.3*) and grown overnight at 37°C. A single colony was used to inoculate 50 mL of SOB media (*Table 2.3*) and once again left to grow overnight at 37°C with shaking (150 rpm). At mid-log phase (OD_{600nm} 0.4-0.6) the culture was centrifuged at 3200 x g at 4°C for 5 min and then resuspended in transformation buffer

(Table 2.5) (2 mL). The culture was centrifuged at 3200 x g at 4 °C for 5 min and the pellet was re-suspended in transformation buffer and 7.5 % DMSO (v/v). Aliquots were stored at -80 °C.

For the transformation, 20 µL of competent *E. coli* cells were transformed with 0.5 µL of the plasmid of choice using heat shock at 42°C for 1 min. After 2 min incubation back on ice, 980 µL of SOC (Table 2.3) was added and the cells were incubated at 37°C with shaking for 1 h. The cells were then plated onto LB-selective antibiotic agar (Table 2.3) and allowed to incubate overnight at 37°C. Negative controls, containing component cells that were heat shocked with no additional DNA, were treated identically.

2.5 GENERAL METHODS USED IN THIS WORK

A number of general methods were utilised throughout this work and these are outlined below, with any specific changes noted where appropriate.

2.5.1 SDS-PAGE ANALYSIS

Throughout this work, all protein samples were analysed using SDS-PAGE using the following ‘in house’ protocol. Resolving gels (12%) were prepared using the components in Table 2.6, where APS and TEMED were added immediately before loading the gel into the casting apparatus (Bio-Rad Mini-Protean 3). Once the resolving gel solidified, the stacking gel (4%) was prepared (Table 2.6). The stacking gel was applied to the top of the resolving gel and a plastic comb was immediately inserted into the gel to create wells.

To each sample to be loaded into the gel, 4 x SDS-LB (Table 2.5) was added to a final concentration 1 x and 10 µL of each sample loaded into the polyacrylamide gel. Along with the samples, 10 µL of Bio-Rad Precision Plus broad range unstained standards was loaded onto the gel immersed in Running Buffer (Table 2.5). SDS-PAGE was run for 40 min at a maximum 300 V with a constant current of 30 mA using the Bio-Rad PowerPac Basic system. The gel was stained using Coomassie Blue stain (Table 2.5) for 30 min and then destained using Coomassie destain (Table 2.5) before being scanned using an Epson Perfection 4870 Photo scanner and visualized using Epson Scan software.

Table 2.6 SDS –PAGE Components for ‘in-house’ Gel

Resolving gel (12%) (for 10 ml)	Stacking Gel (4%) (for 5 ml)
3 ml Acrylamide solution (40%)	500 µl Acrylamide solution (40%)
4.4 ml MQW	3.2 ml MQW
2.5 ml 1.5 M Tris-Cl (pH 8.8), 0.4% SDS	1.25 ml 0.5 M Tris-Cl (pH 6.8), 0.4% SDS
100 µl Ammonium persulphate (10% - frozen stock)	50 µl Ammonium persulphate (10% - frozen stock)
10 µl TEMED	5 µl TEMED

2.5.2 DNA ELECTROPHORESIS

Depending on the size of the DNA fragment, the protocol was slightly altered. In general, a 1% or 2% TBE agarose (Bioline) (*Table 2.5*) gel with 0.5 µL Gel Red (Biotium) was cast. DNA loading buffer (New England BioLabs) to a final concentration of 1 x was added to each sample which were then loaded into the gel (if concentration known approximately 50 ng of DNA was loaded). DNA electrophoresis was run at 100V for 60 min and visualised via exposure to short-wavelength UV light (254 nm) using a transilluminator (Ingenius3; Syngene) and recorded as a digital image using a charge-coupled device (CCD) camera (Synoptics CAM-FLXCM; Syngene) linked to GeneSys molecular imaging software, version 1.5.0.0 (Syngene).

Alternatively, when performing gel purification of DNA fragments, a 0.75% TAE agarose (Bioline) gel with 0.5 µL Gel Red (Biotium) was cast. DNA loading buffer (New England BioLabs) to a final concentration of 1x was added to each sample which were then loaded into the gel (if concentration known approximately 50 ng of DNA was loaded). DNA electrophoresis was run at 60 V for 2 h and visualised via exposure to short-wavelength UV light (254 nm) using a transilluminator (Ingenius3; Syngene) and recorded as a digital image using a charge-coupled device (CCD) camera (Synoptics CAM-FLXCM; Syngene) linked to GeneSys molecular imaging software, version 1.5.0.0 (Syngene). Once the bands of insert were excised from the agarose gel, they were purified and cleaned up using the Bioline Isolate II PCR clean up kit as per the manufacturer's instructions. The final concentration of all samples was determined by the Nano Drop 100 Spectrophotometer at 260 nm.

2.5.3 PREPARATION OF gDNA AND PLASMID DNA

gDNA and plasmid DNA was prepared using the Bioline Isolate II genomic DNA (bacteria) and plasmid (high-copy and low-copy) preparation kits respectively (Bioline Isolate II, Australia) as per the manufacturer's instructions with the following alternations; elution was performed at 70°C and in two steps with a final volume of 50 µL. The final concentration of all samples was determined by the Nano Drop 100 Spectrophotometer at 260 nm.

2.5.4 POLYMERASE CHAIN REACTION (PCR)

Throughout this work, PCR was used to amplify genes of interest for downstream plasmid construction as well as to check for correct gene deletions. Single-stranded oligonucleotide primers for PCR were manufactured by Integrated DNA Technologies (IDT) and supplied in a lyophilised form. Primers were dissolved in MQW to a final concentration of 100 µM and stored at -20°C. Each reaction was prepared as per Phusion DNA polymerase manufacturer instructions (New England Biolabs). In summary, each reaction (50 µL) contained approximately 50 ng of gDNA template (prepared as per *Section 2.5.3*), 1 mM of reaction specific primers, 10 mM dNTPs, Phusion HF buffer, 1 M Betaine, MQW and Phusion DNA polymerase enzyme. All PCR reactions were subjected to the following general temperature cycle using a thermo cycler (Mastercycler; Eppendorf); 2 min at 98°C (for initial denaturation of the template), followed by 30 cycles consisting of 98°C for 30 sec (template denaturation), X°C for 1 min (primer annealing) and 72°C for Y sec (extension), followed by a further 10 min at 72°C for final extension (where X = predicted oligonucleotide annealing temperature minus ~5°C, and Y = 1 min per kb to be amplified). The PCR reaction products were analysed via agarose gel electrophoresis (see *section 2.5.2*) Each PCR reaction was cleaned up using the Bioline Isolate II PCR clean up kit as per the manufacturer's instructions. The final concentration of all samples was determined by the Nano Drop 100 Spectrophotometer at 260 nm.

2.5.5 SANGER SEQUENCING

Throughout this work, all plasmids created and used, had their PCR region (insert) confirmed by Sanger sequencing. 500 – 1000 ng of plasmid DNA purified as outlined in *Section 2.5.3*, along with appropriate primers, at 10 pmol/μL, were sent to the Australian Genome Research Facility (AGRF), where sequencing reactions were performed. Sequencing data was analysed using Finch TV software version 1.40 (QT at TrollTech) and Jalview software version 2.8.2 (181).

2.5.6 GENERAL CLONING PROTOCOLS

A number of plasmids were constructed during this work using standard cloning protocols for an insert and a plasmid backbone of choice. Speaking generally, *H. pylori* was grown as outlined in *Section 2.4.1* and gDNA purified (*Section 2.5.3*). The purified gDNA was then used as the template DNA in a PCR reaction (*Section 2.5.4*) to form the insert (PCR product). Both the insert and plasmid backbone (~1 μg) were digested with the appropriate enzymes (New England BioLabs), with each specific enzyme noted in the appropriate thesis chapters. Digestion reactions were cleaned up using the Bioline Isolate II PCR clean up kit as per the manufacturer's instructions. The final concentration of digestions was determined by the Nano Drop 100 Spectrophotometer at 260 nm. The digested plasmid (vector) was then alkaline phosphatase (New England BioLabs) treated for 1 h and cleaned up and concentration determined as per previously described. A ligation reaction was then set up for each insert with the digested vector, using T4 ligase (New England Bio Labs) in a 3:1 molar ratio insert:vector with 50 ng of vector. The ligation was left overnight at 16°C. Once again, the ligation was cleaned up as before, and transformed into *E. coli* DH5α as outlined in *Section 2.4.3*. Single colonies were grown as stated in *Section 2.4.2* and plasmids recovered (*Section 2.5.3*) for confirmation by PCR (*Section 2.5.4*) or digestion using appropriate enzymes. All confirmed plasmids had their sequences verified by Sanger sequencing (*Section 2.5.5*)

CHAPTER III

CHARACTERISATION OF DYNAMIN-LIKE PROTEINS (DLPS) IN *HELICOBACTER PYLORI*

PREFACE

Previously in this thesis it has been noted that the homologs of the ETEC LeoCBA dynamins in *H. pylori* have not yet been investigated. Therefore, this chapter endeavours to establish their presence in bacteria, and more specifically the level of conservation in *H. pylori*. Through the use of bioinformatics, any hallmarks of a dynamin, such as the presence of a GTPase domain or lipid binding domains, will also be investigated. Finally, with no previous reports detailing whether these genes are expressed in *H. pylori*, their *in vivo* expression is explored.

3.1 INTRODUCTION

The overarching function for the Dynamin superfamily of proteins including both classical and dynamin-related proteins is membrane remodelling. Members of the superfamily such as classical dynamin contain a GTPase domain along with a number of functionally supporting domains, which includes; stalk domain, PH (pleckstrin homology) domain, GED (GTPase effector domain) and PRD (proline rich domain) domain. These defining features can vary in presence as well as homology amongst the different members of the family. For example, dynamin-like proteins do not contain a PRD domain but maintain the GTPase dependent lipid binding mechanism through the GTPase, stalk and GED domains (2). It is these slight differences that allow this superfamily to be clustered into several different sub families that include classical dynamins, dynamin-like proteins, mitofusins, OPA1, Mx proteins, atlastins and GBPs. These differences give rise to varying eukaryotic dynamin function within a cell, from endocytosis to mitochondrial and endoplasmic reticulum fusion to antiviral properties. (1, 2, 182).

Characterisation of the dynamin superfamily has been well established in eukaryotic cells and despite the discovery of their presence in bacteria over 10 years ago, very little is known about their function(s). Despite the limited literature on bacterial dynamins, they have been shown to be highly conserved in both Gram-positive and Gram-negative bacteria, with nearly 1500 sequences identified in nearly 1000 species (6). This conservation is indicative of them performing potentially fundamental biological roles in prokaryotes and thus warrants further investigation. There are several identified bacterial dynamins including DynA from *B. subtilis*, YjdA from *E. coli* and IniA from *M. tuberculosis* (introduced in *Chapter 1*).

The recent identification of the previously characterised *E. coli* virulence factor, LeoA as a bacterial dynamin highlights the potential importance of these genes in bacteria. LeoA has been implicated in the release of the heat-labile toxin (LT) from ETEC into the host environment through potential involvement in the release of outer membrane vesicles (OMVs) (5, 7). LeoA showed structural homology to other bacterial dynamins, with all key domains conserved; GTPase, middle, trunk and paddle/tip (5). The most intriguing

finding surrounding LeoA is the identification of two other dynamin-like proteins directly upstream in the same *tia* locus, *orf 2* and *orf 3*, which Michie *et al* termed *leoC* and *leoB* respectively (5). With most bacterial dynamins to date being identified in tandem, this isn't surprising, but it's the arrangement of these genes that is of interest (6). LeoC and LeoB taken together encode for a single, full-length dynamin but for unknown reasons, these genes have been split in two. LeoC encodes for the GTPase domain containing the P loop and the G1-switch binding motif, as well as a hydrophobic insert that is predicted to be a membrane anchor (5). LeoB encodes for the remaining DLP features including, the G2 switch binding motif, completing the GTPase domain, as well as encoding for a predicted membrane binding domain between the middle and GED domain (5). The reason for this split remains unknown but is intriguing considering neither gene product is expected to function as a proper dynamin without the other.

It has been previously noted, that the closest homolog to LeoA with only 30% sequence similarity, is found in *H. pylori*, specifically HP_0731 (99) and in this thesis, has been assigned the name of *dlp2*. The other open reading frames in the locus *orf 2* (LeoC) and *orf 3* (LeoB) also share homology to the *H. pylori* genes directly upstream, HP_0733 and HP_0732 (5) and have been designated *dlp1a* and *dlp1b* respectively in this work. At first glance, an interesting feature is the difference in this split point of the LeoC and LeoB homologs in *H. pylori*. The length of the proteins is reversed, with Dlp1a being 521 amino acids in length and Dlp1b being 118 amino acids in length, approximately opposite proportions to *E. coli* LeoC (206 amino acids) and LeoB (572 amino acids). Despite this, both proteins show a high level of conservation of the dynamin-like features when taken together.

Bacterial dynamins are highly understudied in the literature and further investigations into their presence, their structure and their function are needed to truly elucidate why bacteria have maintained them. Specifically, the conservation of dynamin-like proteins in *H. pylori* is highly intriguing and is the focus of this research.

The aims of this chapter were:

1. To investigate the conservation of the Leo-Type homologs in bacteria and to identify their position in the dynamin superfamily
2. To examine the conservation of bacterial dynamin-like proteins in *H. pylori* and to identify the dynamin-like features
3. To assess the *in vivo* expression and co-expression of the dynamin-like genes in *H. pylori*

3.2 MATERIAL AND METHODS

3.2.1 *COLLECTION OF SEQUENCES OF THE LEO TYPE DYNAMINS AND THE DYNAMIN SUPERFAMILY*

LeoC, LeoB and LeoA sequences from ETEC (NC_017633.1) defined by (5) were used as individual query sequences in a psi-BLASTp iterative search with a restricted bacterial phyla database. The top five strains with the highest query coverage and lowest e-value generate by BLAST from each bacterial phyla and/or class were used. A diverse range of Dynamin superfamily protein sequences of both bacterial and eukaryotic origin were retrieved from the NCBI database.

3.2.2 *PHYLOGENETIC ANALYSIS*

A multiple sequence alignment was built from the sequences collected for each analysis in *Section 3.2.1* and was performed by MUSCLE (183) and maximum likelihood phylogenetic analysis carried out in MEGA6 using default settings (184). In the case of LeoC and LeoB homologs, sequences were manually fused, end-to-end (LeoC followed by LeoB) to allow for an analysis of biological significance to be conducted. The maximum likelihood tree was bootstrapped with 100 replicates and the tree output was visualised in FigTree (version 1.3.1)

3.2.3 *H. PYLORI SCREEN FOR LEO-TYPE DLPS AND PHYLOGENETIC ANALYSIS*

The iterative search algorithm, JackHMMER (185) was utilised to search all *H. pylori* genomes for homologs of LeoCBA. LeoC, LeoB and LeoA protein sequences from ETEC (NC_017633.1) were used as individual query sequences and all hits with a p-value of 0.05 or less was used in the analysis. Mathematica 10.0 was used to analyse the hits from JackHMMER as well as analyse the hits. Further analysis of the genomes found to contain all three proteins in succession in the genome was performed in MEGA 6 (184). Phylogenetic analysis of the protein sequences was performed as outlined in *Section 3.2.2*.

3.2.4 DLP STRUCTURAL FEATURES

H. pylori strain 26695 dynamin like protein sequences, HP_0731 (Dlp2), HP_0732 (Dlp1b) and HP_0733 (Dlp1a) were retrieved from the NCBI database and used as query sequences in transmembrane domain predication software, TMHMM (186). These sequences were also used in the protein predication web-based software, Phyre2 (187). Models predicted by Phyre2 were analysed in The PyMol Molecular Graphics System, Version 1.74, Schrodinger, LLC.

3.2.5 RNA PURIFICATION AND REVERSE TRANSCRIPTION (RT)-PCR

A *H. pylori* 26695 cell pellet (10 ml of culture at OD_{600nm} = 0.6, or equivalent) was resuspended in cold 10 mM Na Citrate, 10 mM Na Acetate (pH 4.5), 2 mM EDTA (150 µL). Cells were then transferred into a tube with phenol (400 µL) and 10 mM Na Acetate (pH 4.5), 2% SDS (150 µL), pre-warmed to 65°C. Contents were vortexed and incubated at 65°C for 3-4 min, with shaking at 1300 rpm. Samples were then snap frozen in liquid nitrogen and spun for 5 min at 14 000 x g. Phenol (400 µL) was added to the aqueous phase and transferred to a new tube and spun again as above. This was repeated with phenol (200 µL) and chloroform (200 µL) and then chloroform only (400 µL). After the final spin, 2.5 volumes of 100% (v/v) ethanol, 1/10 sample volume of 3 M Na-acetate and 1 µg glycogen was added to the aqueous phase and incubated at 0°C overnight. Samples were then spun at 20 000 x g for 1 h at 4°C. The RNA pellet was washed in 70% (v/v) ethanol and then re-suspended in RNase free H₂O.

Reverse transcription of the RNA was performed using Superscript III (Invitrogen) as manufacturer's instructions. Each PCR reaction contained *H. pylori* cDNA, paired primers (Table 3.1) for each gene region and Sybr Green JumpStart Taq ReadyMix (Sigma-Aldrich). PCR was performed on a ThermoCycler (MasterCycler Pro, eppendorf) for 40 cycles and data recorded. Raw Ct values generated were analysed using GraphPad.

Table 3.1 Primers used in RT- PCR

Primer Name	Sequence
26695LeoA_R	GAGGTGGGCTTCGCTGAT
26695LeoA_F	GTTGACACCCCGGGACTATTC
26695LeoAB_R	CTTTTAAATTGTTCTAGCGTTTCA
26695LeoAB_F	AGCCCATGAAAGATTAAGATACAT
26695LeoB_R	TCATGCGTTTGTAATTGTCAACAG
26695LeoB_F	TCCCAACAAAGAAAAGAAGTGGA
26695LeoBC_R	AATAATCCTGCGGTCAGAGCA
26695LeoBC_F	ACTGATAGCGGTATTGATGGAACA
26695LeoC_R	AAACCGCCATTTTATGTCTTCA
26695LeoC_F	ATTTAACCAATTCAGCGTTCTACG

3.3 RESULTS

In order to confirm the presence of *H. pylori* dynamins, a bioinformatic analysis into other bacterial homologs as well as their place within the greater dynamin superfamily was necessary. Establishing their position in the dynamin superfamily and their domain features, may help bring to light any features that have been previously noted in other members. In turn this may help elucidate their function(s) in *H. pylori*. Annotation of functional domains in the *H. pylori* dynamins as well as investigating the predicted structure of these proteins may help highlight important structural features conserved in bacterial dynamins. Furthermore, HP_0731 (*dlp2*), HP_0732 (*dlp1b*) and HP_0733 (*dlp1a*), the bacterial dynamins in *H. pylori* strain 26695, have been annotated as hypothetical in the NCBI database with confirmation needed for their expression. This work was essential to the future work in this project as there was no previous detailed information about this class of proteins in *H. pylori*.

3.3.1 CONSERVATION OF THE DYNAMIN SUPERFAMILY

A preliminary investigation into the conservation of dynamins in bacteria was carried out to both elucidate the scope of their retention in the bacterial genome as well as where they sit within the greater dynamin superfamily. It has been stated in the literature that there are many bacterial species that encode dynamin genes in their genome but have not been annotated as such or studied any further (6).

3.3.1.1 Dynamin conservation in Bacteria

Exploring the scope of the bacterial dynamins allows a thorough understanding of the depth of conservation of what we have termed the Leo-Type dynamin family members in bacteria. A multiple sequence alignment was built from detectable homologs, using NCBI psi-BLASTp of the ETEC proteins LeoA, LeoB and LeoC in each of the bacterial phyla that showed the presence of dynamin. For the purpose of this alignment, homologs of LeoB and LeoC were fused and annotated as Dlp1 and homologs of LeoA were annotated as Dlp2. From this alignment, a bootstrapped maximum-likelihood (ML) consensus tree was generated (*Figure 3.1*). Bootstrapping assigns measures of accuracy and confidence

to the phylogenetic analysis with the branching of highest confidence seen in the consensus tree. This allowed observations to be made about clustering of dynamin sequences from a range of bacterial species with a good level of confidence.

The first notable observation is that there is a distinct split between the LeoA (Dlp2) homologs and the LeoC/LeoB (Dlp1) homologs. This indicated that these proteins, known to only show approximately 20% sequence similarity, form two sub families in the Leo-Type dynamin family that is conserved across a range of bacteria. From this initial search, it was seen that the highest conservation of the Leo-Type dynamin genes was in the Proteobacteria phylum; within the Proteobacteria, homologs were found in the classes of beta, gamma and epsilon. Particular conservation was seen in the classes of gamma and epsilon Proteobacteria, with species but not all strains of important human pathogenic bacteria like *Pseudomonas aeruginosa*, *Salmonella enterica* and *H. pylori* encoding dynamin-like proteins. Along with Proteobacteria homologs, there was detection of dynamin-like proteins in other classes of bacteria, including Firmicutes and Bacteroidetes.

3.3.1.2 Bacterial dynamins in the context of the greater superfamily

Previously, the Dynamin superfamily of proteins has been characterised as a diverse group of proteins, forming clear distinct clusters in a phylogenetic analysis. With the addition of bacterial DLPs, in particular the Leo-Type bacterial DLPs, these clusters don't vary dramatically, and the bacterial homologs form additional branches as well as clustering with a few of the already established eukaryotic DLP branches. To more closely examine the conservation and the relationship between the dynamin family members from the different domains of life, both bacterial and eukaryotic, with the addition of the Leo-Type homologs of this family, a maximum-likelihood tree was generated. A multiple sequence alignment of a diverse range of dynamin family members was built through an iterative MUSCLE alignment using MEGA. This type of repeating protein alignment has been shown to have a more accurate distance measure to assess the relatedness of two sequences compared to other progressive methods. The distance measure between two sequences is updated after each iteration (repeat) of the initial alignment, building arguably a more accurate alignment for use in phylogenetic analysis (183).

As can be seen in *Figure 3.2*, the eukaryotic DSPs form a distinct branch consisting of both classical dynamins as well as dynamin-like proteins (blue). Moving away from this branch, dynamin members implicated in viral replication, Mx1 and Mx2, form an individual branch in the same cluster. Previously it was reported that the first identified bacterial dynamin, BDLP1 from *Nostoc punctiforme*, is most closely related to the Mitofusins class of DLPs, which includes *Drosophila* mitochondrial Fzo homologs and *Arabidopsis* chloroplast FZL and this is designated in pink in *Figure 3.2* (4).

From this point, the tree forms two divergent branches, consisting of the two previously characterised sub families of Leo-Type bacterial dynamins (green), with homologs from the Proteobacteria including *Shewanella ANA3*, *Vibrio cholera* and *H. pylori*, clustering together for both Dlp1 and Dlp2. Closely related to the Dlp2 homologs is a sub-branch that is a little more complex, consisting of both bacterial and eukaryotic dynamins. In this branch, there are two uncharacterised and newly identified *H. pylori* dynamin-like proteins (HP_0965 and HP_0966) clustering with eukaryotic members Atlastin1 and GBP1. Finally, other bacterial dynamin members, such as IniA from *M. tuberculosis* and

CrnC from *E. coli* and *S. enterica*, form their own sub branch (red) along with newly identified bacterial dynamins from *Streptomyces venezuelae*, DynA and DynB. Both of these groups of dynamins have been shown to have alternative function to the traditionally ascribed dynamin function in bacteria, such as involvement in DNA replication and FtsZ ring stabilisation (93, 98). Interestingly, the closest *H. pylori* homologs of CrnC, HP0965 and HP0966, don't cluster into this branch. Further interpretation and interrogation of the sequences is needed to understand this difference. However, it is evident from this phylogenetic tree of the dynamin superfamily that the Leo-Type members form their own branch or subfamily under the umbrella of the dynamin superfamily.

3.3.1.3 *H. pylori* genomic screen for Dlp

Having established homology between the LeoCBA genes found in ETEC with genes that are found in 26695, the type strain of *H. pylori*, the conservation of these genes was then further investigated by screening all available *H. pylori* genomes using an automated search strategy. In the ETEC genomes, dynamins are only found in low abundance, with only 3% of the genomes showing evidence of a dynamin (5). However, in the *H. pylori* genome the situation is vastly different. It has been previously stated in the literature that the closest homologs to the ETEC LeoCBA proteins are found in *H. pylori* (5, 99). Therefore, a thorough genome wide JackHMMER search was conducted using ETEC LeoCBA sequences as the search term. This type of search differs from BLAST+ in that it uses a form of Hidden Markov Models (HMM), called a Profile HMM which is a statistical model of single sequences that capture position-specific information about how conserved each residue in the column of the alignment is. JackHMMER (psi-BLASTp-like) specifically, is an iterative (repetitive) that uses the initial search to then build a HMM alignment for the next search and so on, until no more significant homologs are detected for the alignment (185).

During the search, 260 *H. pylori* genomes were analysed, both open and closed, and of these, over 95% of the 260 *H. pylori* genomes screened showed the presence of at least one putative dynamin like protein with homology to the Leo-Type dynamins (*Figure 3.3*, indicated in green) (*see S2 for strain names*). It should be noted that this proportion may increase (or decrease) as genomic and proteomic sequences are further resolved in this species with many of these genomes not closed in the NCBI database. This is an extremely high conservation of the Leo-Type dynamins in *H. pylori*; emphasising their potential significance in this bacterium and why they formed the focus of this research.

Curiously, when *dlp1a* was not detected in a genome, neither was *dlp1b*; supporting the notion that the absence of one affects the presence of the other. This was seen in the *H. pylori* strains; Aklavik117, PZ086, 35A, B38, Gambia 94/24, UM038 and UM298. With *dlp1b* being important to the presence of *dlp1a*, it could indicate that the genome needs to encode a full-length dynamin for it to carry out function in the pathogen. However, the opposite doesn't hold true with *dlp1a* being found in a number of genomes (CPY1962, FD577, A-14, P-4, P-4c, P-4d, Puno135, Shi169, F32, PeCan4 and Sat464) despite the

absence of *dlp1b*. It could indicate that in these genomes the *dlp1a* sequence encodes the full length DLP and therefore warrant further interrogation. Either way, in *H. pylori*, as seen in ETEC, dynamins come in pairs. Of the 249 strains that were seen to detect a Leo-Type dynamin like protein, 136 strains (*see S2 for gene names*) showed these genes to be back-to-back in similar genetic locations, juxtaposed to similarly annotated genes. Upstream to the *dlp* locus is several conserved hypothetical and putative proteins while downstream of the locus there are similar transferase and synthase enzymes. The “back-to-back” nature of *H. pylori* dynamins is a previously seen trait of these genes with bacterial dynamins always seemingly coming as adjacent pairs (5). The conservation of these genes, in the back-to-back operon structure, may indicate that they are both needed to carry out the biological function of dynamin within *H. pylori*. This screen emphasised that the DlpS are highly conserved in *H. pylori* and warrant further investigation to understand why they have been retained by this important human pathogen throughout its evolution.

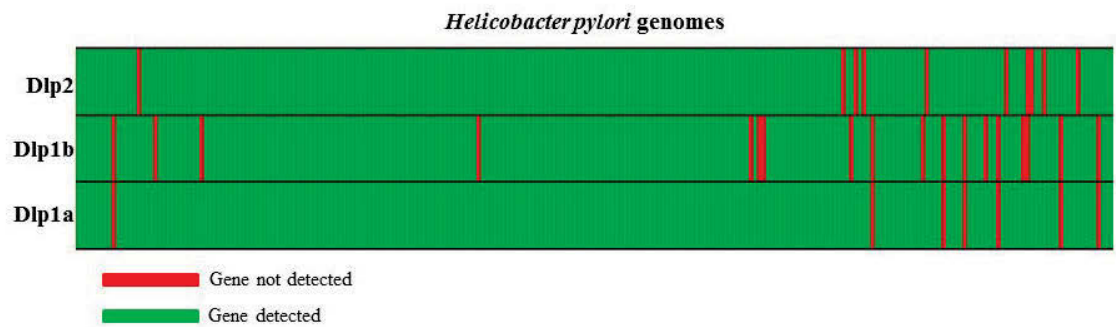


Figure 3.3 Genomic Screen of *H. pylori* for Leo-Type Dynamins

JackHMMER was used as an iterative search for homologs to ETEC leoCBA particularly in *Helicobacter pylori* genomes. Overall, 260 genomes both open and closed were screened for the presence of these genes. If a homolog was detected with a p -value of less than 0.05 it is represented by green, if no homolog was detected below this threshold it is represented by red. Each vertical bar represents a single genome.

3.3.2 DYNAMIN-LIKE PROTEINS IN *HELICOBACTER PYLORI*

The high conservation of dynamin like proteins in *H. pylori*, prompted to further work on these proteins. With the retention of these genes in this pathogen, it was hypothesised that they were functionally important for the varying strains, despite pathological differences. Further exploration of their dynamin-like features was considered essential to determine that these genes were genuinely of the dynamin superfamily.

3.3.2.1 *H. pylori* phylogenetic analysis of the Dlps

To further investigate the relationship between the different homologs of *H. pylori* dynamin-like proteins, Dlp1 and Dlp2, a boot-strapped maximum-likelihood phylogenetic tree was built from the multiple sequence alignment generated from the dynamins identified in the JackHmmer iterative HMM search (*Figure 3.3*). Only *H. pylori* strains that showed the presence of all three homologs, *dlp1a*, *dlp1b* and *dlp2* were used in this analysis (136 strains). The genes identified in these strains were translated into the encoded protein sequence and for accurate protein motif identification and ensuring a reliable phylogenetic tree, the sequences for Dlp1a and Dlp1b were joined (Dlp1) for the analysis.

Dlp1 and Dlp2 cluster distinctly on their own branches and even more so, cluster into distinct sub-clades within each branch (*Figure 3.4*). Dlp1 forms 4 sub-branches containing groups of *H. pylori* strains sharing similar pathology (pink). Dlp1 sequences from 26695, R055a and H-18 clustered on a single branch, with sequences from B8, HPAG1 and G27 on another but stemming from the same root. *H. pylori* J99 was on yet another branch. There was a rather small third cluster containing Dlp1 sequences from strains such as strain 908 and others. The final branch had two sub groups also, with a range of sequences including P13. Dlp2 did not show this same clear clustering with a more diverse spread of the sequences (blue) (*for the complete labelled phylogenetic tree of Dlp1 and Dlp2 see S4*).

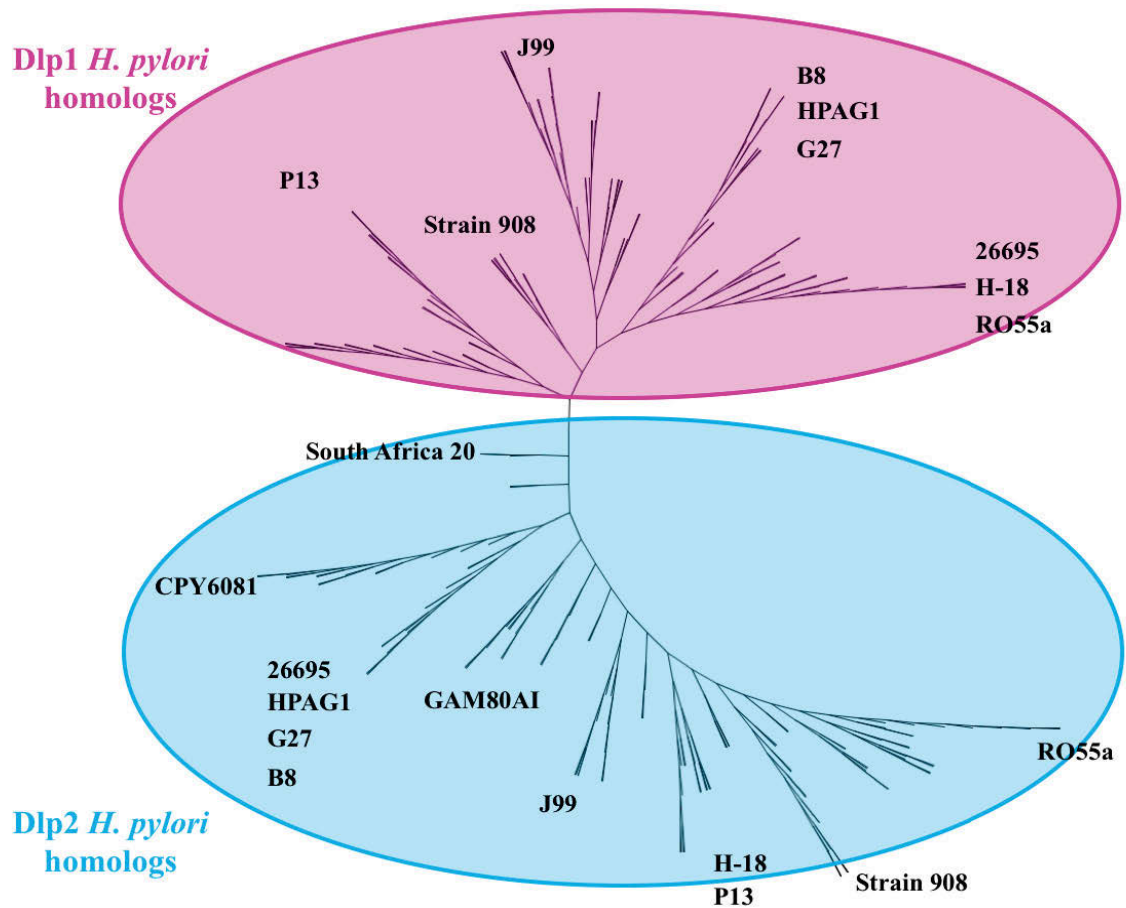


Figure 3.4 *H. pylori* DLPs form a Distinct Two Branch Phylogenetic Tree

A maximum likelihood tree was generated in MEGA6 from a multiple sequence alignment built from the protein sequences from the 136 *H. pylori* strains that had a homolog detected to the ETEC proteins *LeoC*, *LeoB* and *LeoA* (Figure 3.3). Dlp1 homologs cluster together (pink) and Dlp2 homologs cluster separately (blue). Representative strains for each of the major branches are labelled with a completely labelled tree available in S4.

3.3.2.2 Dlp operon in *H. pylori*

The *dlp* operon in *H. pylori* of the Leo-Type dynamins encodes Dlp1a, Dlp1b and Dlp2. Ten diverse *H. pylori* strains identified in *Figure 3.3*, to contain all three *dlp* genes back to back in their genomic locations, were further investigated in order to compare the overall structures of the *dlp* operons across these strains. As seen in *Figure 3.5*, the predicted encoded Dlp2 protein was consistently similar in length (approximately 570 amino acids) throughout the strains, with the exceptions of G27 and CPY6081 where it is shorter. Interestingly, the only evidence of a split or broken gene for Dlp2 was seen in B128.7.13, a gerbil-adapted *H. pylori* strain, where by the Dlp2 was split into three segments each showing homology to regions of the full length Dlp2. Interestingly, this split seems only genomic and not proteomic, with expression of Dlp2 confirmed at 66 kDa, the full length (see *Chapter IV*). One possible explanation for this is the existence of errors in the genome sequence for this strain and further work is needed to investigate this possibility.

From this analysis, the most interesting finding was the varying split point between Dlp1a and Dlp1b, which differed to what is seen for Dlp2. Varying lengths of Dlp1a and Dlp1b were seen consistently in the selected strains depicted in *Figure 3.5*. Despite this, a predicted transmembrane domain is conserved (see *S3 for hydrophobicity plots*) and located approximately 100 amino acids from the end of the protein (red). This conservation highlights that both genes would be needed for the function of the Leo-Type dynamins in *H. pylori*, with an essential feature in both; membrane binding, seen in Dlp1b and predicted GTPase domains in Dlp1a.

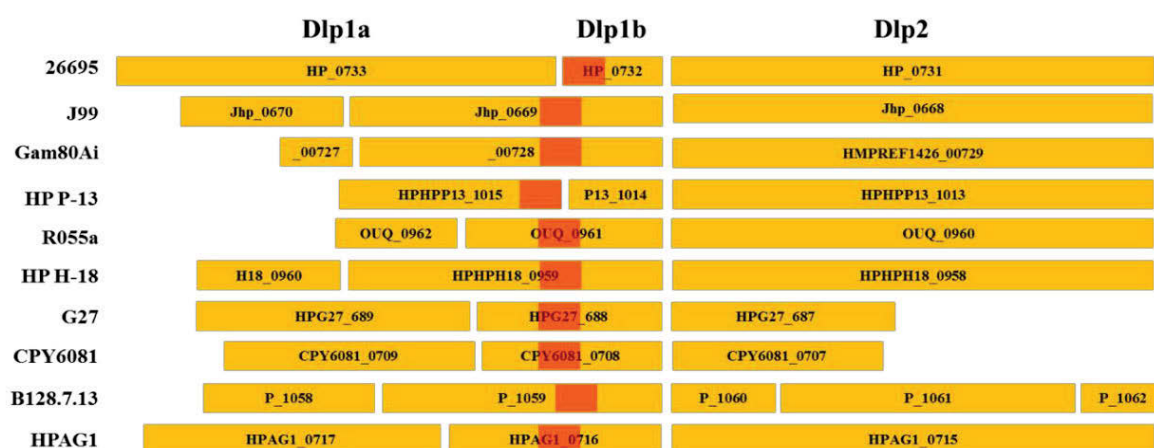


Figure 3.5 *H. pylori* dlp Operon Structure in Selected Strains

10 representative *H. pylori* dynamin protein operon structures (left – *Dlp1a* homologs, middle – *Dlp1b* homologs and right – *Dlp2* homologs) are pictured, labelled with their individual NCBI protein names. Using TMHMM (expasy) transmembrane domains were predicted in the *Dlp1a*/*Dlp1b* homologs (red). Length is proportional to protein amino acid length, but gaps are not to scale. There is a consistent variable split point between *Dlp1a* and *Dlp1b* homologs with *Dlp2* size relatively conserved except for G27, CPY6081 and B128.7.13 genomes.

The Dlp2s in *H. pylori* 26695 were further investigated by analysis of the *dlp* operon and the protein products of gene expression. Key dynamin superfamily features are seen in the 26695 *H. pylori* strain *dlp* operon as expected (Figure 3.6). Dlp2, is 573 amino acids in length with a predicted GTPase domain from amino acid 41 to amino acid 221. Dlp1a is 521 amino acids in length and also has a GTPase domain from amino acid 45 to amino acid 225. As seen in the multiple sequence alignment of a number of *H. pylori* Dlp1a and Dlp2 protein sequences, this GTPase domain shows high conservation of the G1 P-loop motif, GxxxGKT/S, the G2 Switch I motif, T and the G3 Switch II motif DxPG. Furthermore, there are two more regions of high homology that are highlighted within this GTPase region of the protein.

The final protein produced by this operon is Dlp1b, which is 118 amino acids in length and encodes a predicted transmembrane domain (TMHMM) from amino acid 19 to amino acid 29. Taken together, Dlp1a and Dlp1b contain all the features of a dynamin protein, with a GTPase domain for nucleotide hydrolysis and membrane binding domain for lipid interactions. From this it was evident that *H. pylori* dynamins contain the conserved dynamin superfamily features that have been shown previously to be essential to their biological function within a cell.

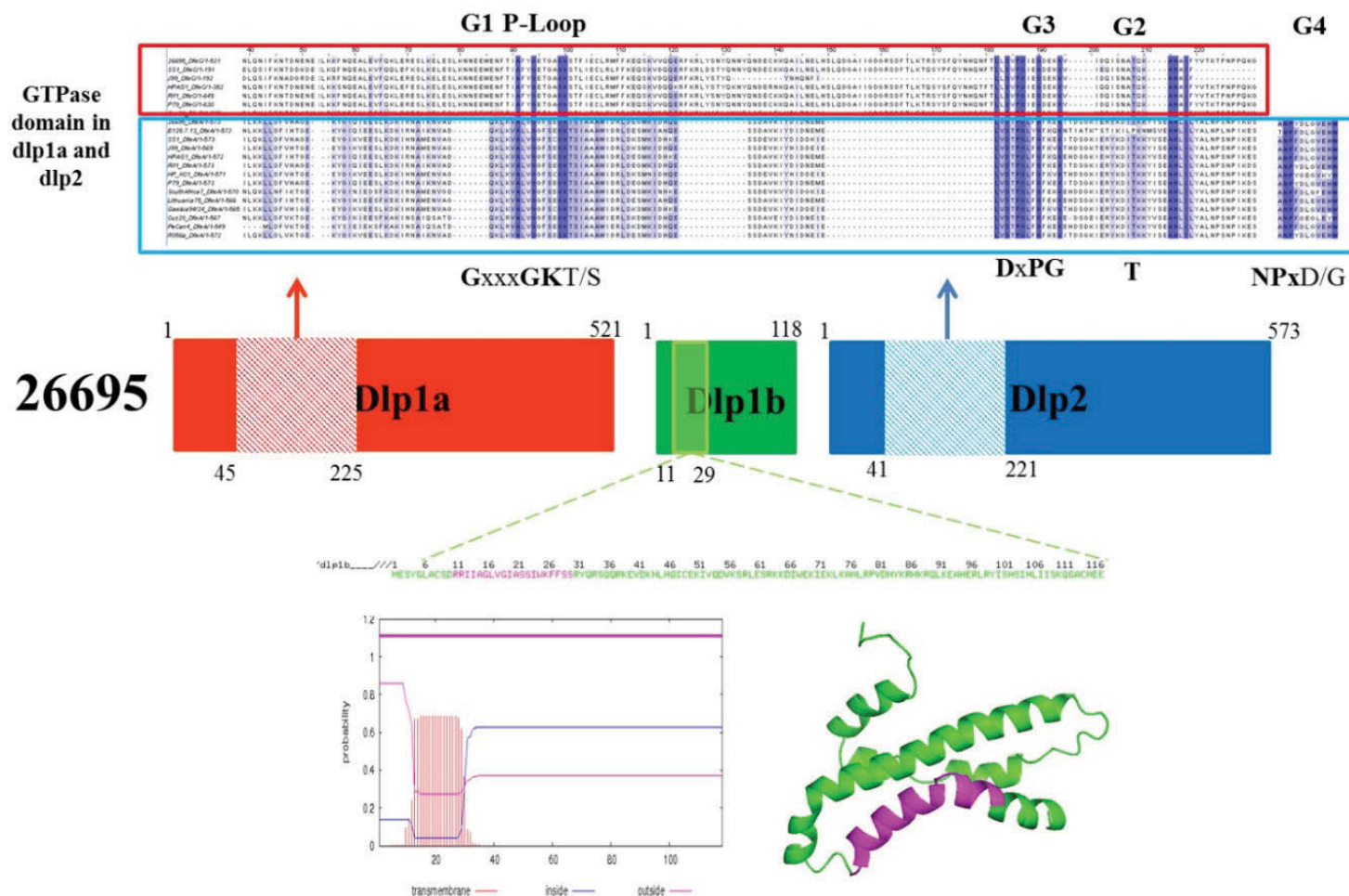


Figure 3.6 H. pylori 26695 dlp Operon Features

A closer look at the dynamin operon structure of the H. pylori strain 26695. Dlp1a (red) is 521 amino acids in length, with a GTPase domain (45-225 amino acid). Dlp1b (green) is 118 amino acids in length and has a predicted (TMHMM) transmembrane domain as seen by the hydrophobicity plot (bottom right). Additionally, this domain is mapped to the predicted Dlp1b structure generated by Phyre2. The domain is indicated in pink. Dlp2 (blue) is 573 amino acids in length, with a GTPase domain (41-221 amino acid). Both Dlp1a (red boxed sequences) and Dlp2 (blue boxed sequences) GTPase domains, show homology across a range of H. pylori dynamins (above), with the known binding motifs GxxGKT/S (P-loops) and G3 (Switch II) DxPG.

3.3.2.3 Predicted Dlp structural features

Structural predictions are extremely useful in highlighting structural homology and features that may not be evident with primary sequence homology, especially when the sequences are quite dissimilar, as in this case. *H. pylori* dynamins share only ~30% sequence homology with their closest homologs, LeoCBA in ETEC, but when the predicted structures were compared, the homology became clearer. For structural comparisons, the web-based prediction server, Phyre2 was utilised for Dlp2, Dlp1 as well as Dlp1b structural predictions (187). In summary, Phyre2 is a suite of tools available to predict and analyse protein structure, function and mutations. In this study, Phyre2 was used to predict the 3D structure of the Dlp from *H. pylori*, Dlp1 and Dlp2 (i.e. in analysis Dlp1a and Dlp1b were treated as a single protein, Dlp1).

Dlp2 shows a striking homology to LeoA from ETEC, as expected (*Figure 3.7*) (5). Phyre2 used one template from the protein database (pdb), LeoA (4AUR), to build the 3D model and the structural alignment can be seen in *Figure 3.7A*, where 92% of all residues were mapped with a confidence greater than 90%; with the reduction in confidence seen mainly towards the C-terminus of the protein. The predicted 3D structure of Dlp2, with the previously noted structural features of bacterial dynamins is seen in *Figure 3.7B*. The software predicted the existence of an N-terminal GTPase domain, a neck and trunk domain in the middle and then a tip domain at the C-terminus. The structure is colour coded according to the sequence (blue to red, N- to C- terminus). The most obvious difference between this predicted structure and its closest homolog LeoA,

from which this model is based, is the tip domain, with this structure seemingly predicted to have curved region back towards the GTPase domain. It will be interesting to determine whether this novel structure prediction for a DLP occurs in future structural studies. When the hydrophobicity of the protein was investigated, there was no large and striking patches evident from the structural prediction (*Figure 3.7C*) that might indicate a membrane binding domain. Rather, it seems that one face of the protein (left) is generally more hydrophobic compared to the other (right), with the hydrophobic residues (red) more concentrated towards the tip domain. This hydrophobicity on one side of the protein may allow the protein to come in close proximity to the membrane, or potential protein interaction partners(s) potentially needed for function.

Overall this structure prediction analysis was highly consistent with Dlp2 having a high structural homology to LeoA, and it appears to contain the structural features that are hallmarks of the dynamin superfamily.

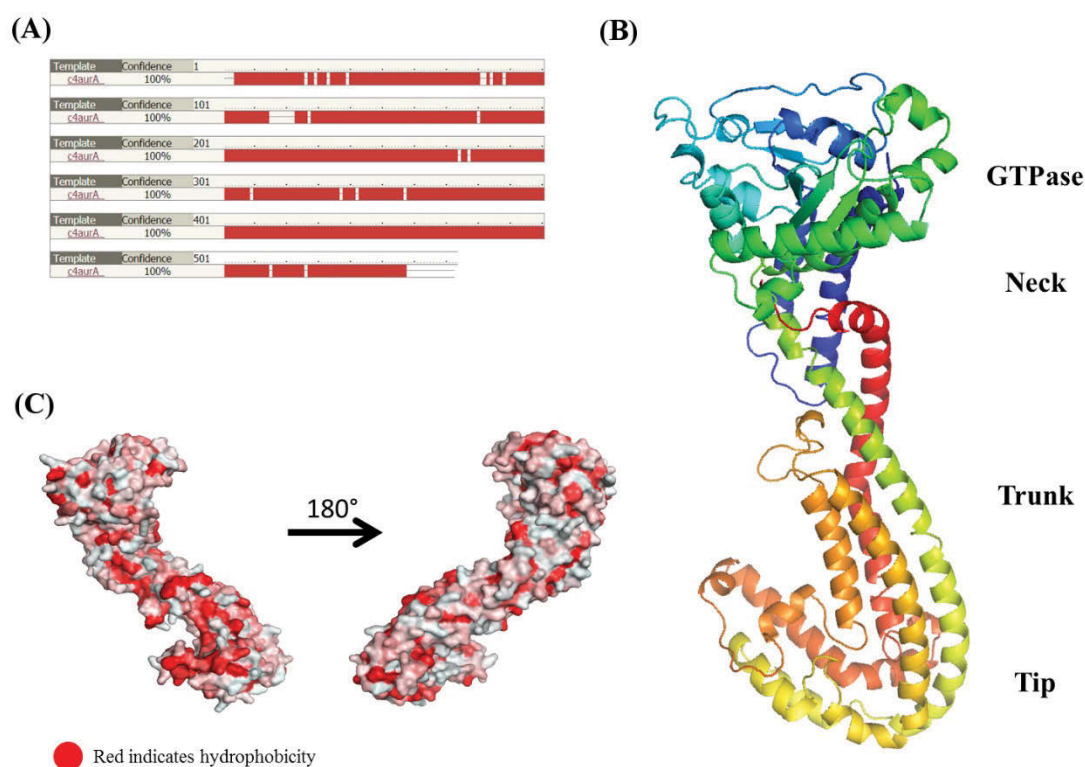


Figure 3.7 Structural Analysis of Dlp2 using Phyre2 Prediction

Phyre2, a web based protein structure prediction platform was used to create a model of Dlp2. (A) It used one template sequence, LeoA from ETEC from the protein database to model the residues, with 92% of the residues modelled with a greater than 90% confidence. Coverage across the full sequence is seen. (B) 3D predicted ribbon structure of Dlp2 colour coded N- to C- terminus in rainbow. This predicted structure shows all previously noted bacterial dynamin features, GTPase, neck, trunk and tip domains. (C) Hydrophobicity (red gradient) was mapped onto the predicted structure using PyMol.

With no previously characterised proteins with close sequence homology to Dlp1, it was intriguing to see the Phyre2 predicted resemblance to the previously-characterised bacterial dynamin BDLP1 (4). The full length Dlp was used to predict this model, with both Dlp1a and Dlp1b needed to create a full-length Dlp. Three templates were used by Phyre2 to build the predicted 3D model of Dlp1 and these were BDLP1 from *N. punctiforme* (14% sequence identity), LeoA from ETEC (16% sequence identity) and mitofusin-1 from *Homo sapiens* (10% sequence identity) (*Figure 3.8A*). The use of these three proteins to construct this model built more confidence into the predicted structure, with 99% of residues with a greater than 90% confidence in the structure and with coverage of the entire Dlp1 sequence. The predicted 3D structure of Dlp1, which contains previously noted structural features of bacterial dynamins is seen in *Figure 3.8B*. It shows an N-terminal GTPase domain, a neck and trunk domain in the middle and then a tip domain at the C-terminus. The structure was colour coded according to the sequence (blue to red). An important feature that was previously noted was a predicted transmembrane domain in Dlp1b (*Figure 3.6*). Therefore, the hydrophobicity was investigated for the Dlp1 full-length protein (*Figure 3.8C*). With red indicating hydrophobicity, a clear hotspot for these residues in the tip domain of the protein (black circle) was evident, consistent with the conserved location for membrane binding seen in both BDLP1 and LeoA (4, 5). When the predicted transmembrane was mapped onto the predicted 3D structure, the residues were located in the tip domain (highlighted pink) (*Figure 3.8D*).

Overall this analysis, confirmed that Dlp1, shows structural homology to previously characterised dynamins, and contains the structural features, such as a transmembrane region in the tip domain, to be termed a dynamin family member.

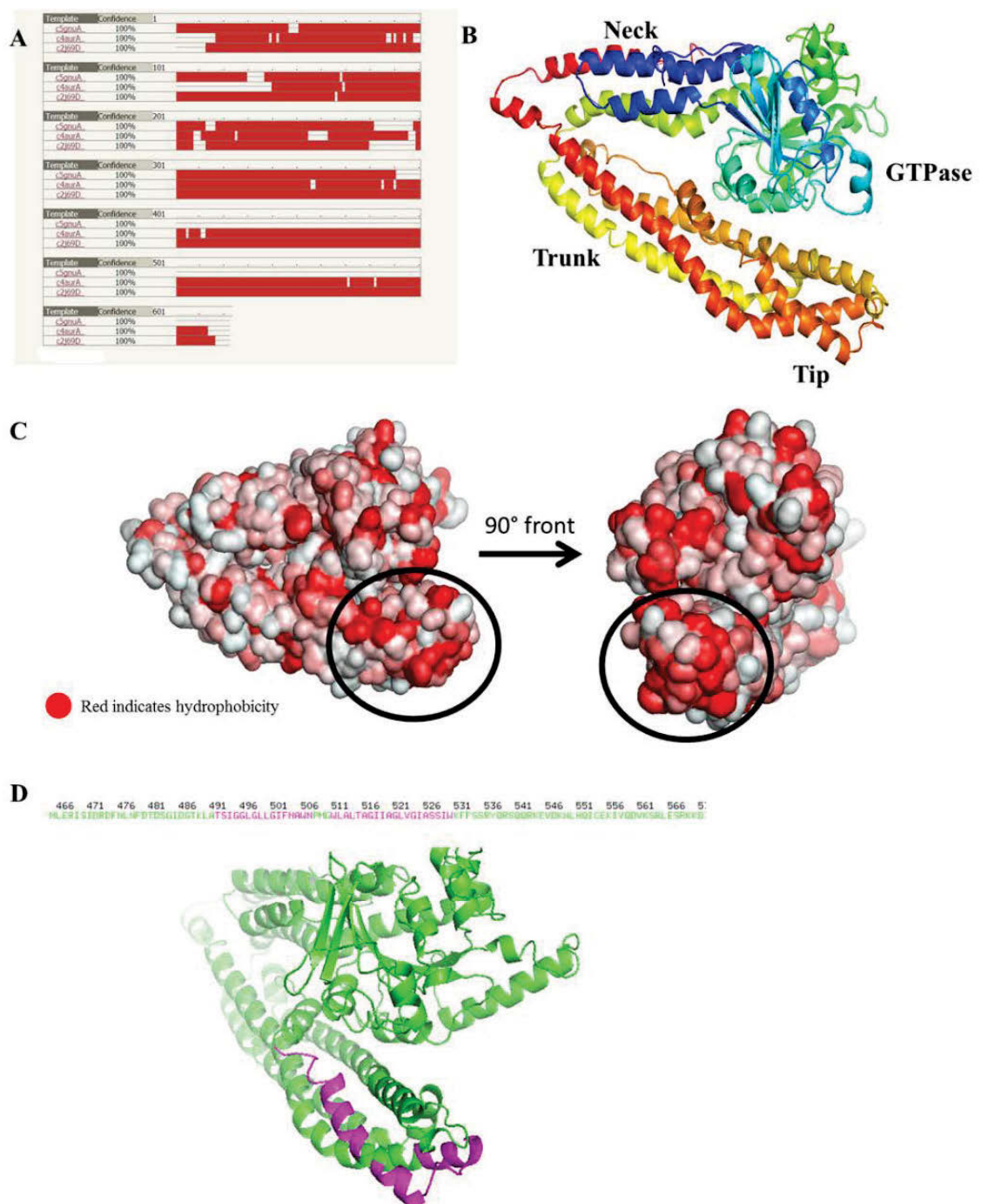


Figure 3.8 Structural Analysis of Dlp1 using Phyre2 Prediction

Phyre2, a web based protein structure prediction platform was used to create a model of Dlp1, where a full-length sequence was used as the query sequence. (A) Phyre2 identified three template sequences, BDLP1 from *N. punctiforme*, LeoA from ETEC and mitofusin-1 from Homo sapiens from the Protein Database to model the residues, with 99% of the residues modelled with a greater than 90% confidence. Coverage across the full sequence

is seen. (B) 3D predicted ribbon structure of Dlp1 colour coded N- to C- terminus in rainbow. This predicted structure shows all previously noted bacterial dynamin features, GTPase, neck, trunk and tip domains. (C) Hydrophobicity (red gradient) was mapped onto the predicted structure using PyMol, with a hotspot of these residues seen in the tip domain (black circle). (D) When the predicted transmembrane for Dlp1 is mapped onto the predicted 3D ribbon structure, the residues are located in the tip domain (highlighted pink).

3.3.3 EXPRESSION OF DLPS IN *HELICOBACTER PYLORI*

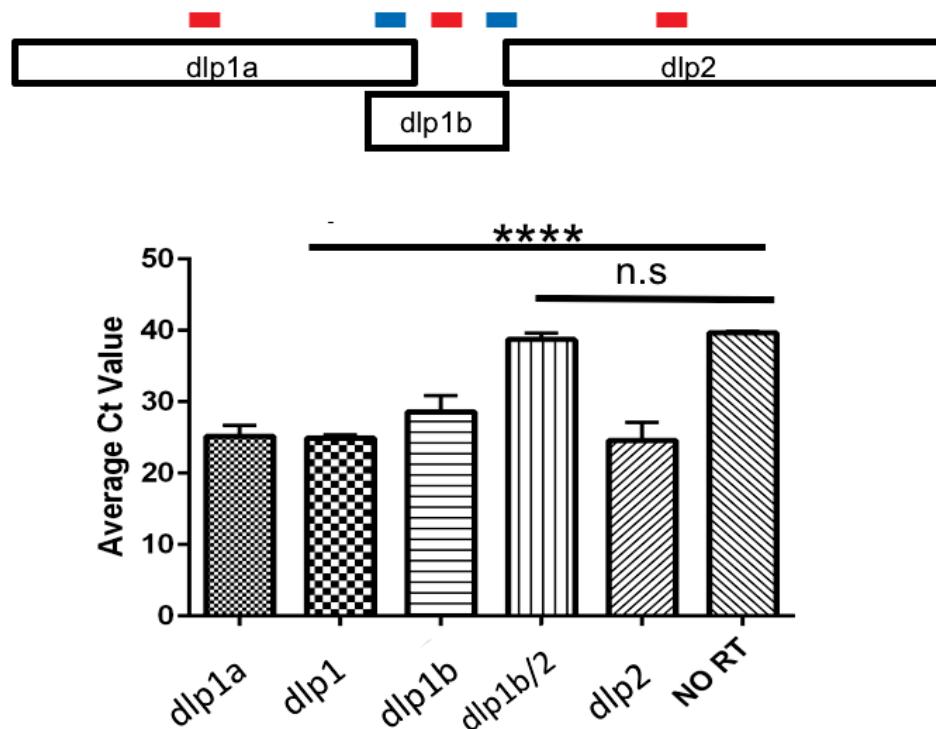
Bioinformatic characterisation highlighted that the *dlp* operon and the DlpS encoded within it have the conserved features seen in the Leo-Type dynamins as well as the greater dynamin superfamily. However, currently these genes are labelled as ‘hypothetical’ in the NCBI database, and their expression has not been explored in depth.

3.3.3.1 RT-PCR analysis of *dlp* expression

To examine the *in vivo* expression of the *dlp* genes in *H. pylori*, reverse-transcription PCR was used, to detect the presence of mRNA in specific regions throughout the *dlp* operon in *H. pylori* 26695. Primer sets for each of the individual genes were designed (red), *dlp1a*, *dlp1b* and *dlp2*, as well as the regions that span two neighbouring genes (blue), i.e. *dlp1a-dlp1b* and *dlp1b-dlp2* to ascertain whether these genes were co-transcribed or not (top panel Figure 3.9). RNA was purified from a late log phase *H. pylori* 26695 *H. pylori* culture and it was reverse transcribed to cDNA for PCR. There was no internal standard used in this initial experiment, as absolute quantification was not the goal but rather the establishment of the presence or absence of a product. As ‘real-time’ quantitative PCR was used to assess this, the cycle threshold (Ct) Ct values were obtained by monitoring the appearance of product over 40 cycles of amplification. The Ct value represents the cycle at which the product signal passes the background threshold, and the higher the Ct value, the lower amount of product and therefore lower expression of the gene of interest.

All three individual genes, *dlp1a*, *dlp1b* and *dlp2* showed a clear detectable level of expression, with Ct values averaging around 25 (Figure 3.9). An interesting observation

from this experiment was the potential co-expression of *dlp1a* and *dlp1b*, supporting the presence of the full length bacterial dynamin Dlp1. Compared to the *dlp1b-dlp2* region where the Ct value is indistinguishable from the negative control (No RT), the *dlp1a-dlp1b* spanning region showed a significant detection of mRNA expression ($p < 0.0001$) with a clear decrease in the Ct indicating the presence of a gene product. From this data, it was impossible to determine whether the genes were completely co-transcribed. However, it was noted that the *dlp1a* and *dlp1b* genes overlap slightly by approximately 25 nucleotides, providing more evidence for co-expression. The findings lending support to the notion that these fragments may function together or as the complete Dlp1 protein.



**** $p < 0.0001$

Figure 3.9 RT-PCR Analysis of dlp Operon Expression in 26695

Reverse transcription PCR was carried out on *H. pylori* 26695 RNA with *dlp* operon primer pairs (above) which were specific to each individual gene (blue) as well as spanning the overlapping regions of the operon (blue), examining co expression. Average Ct values were plotted for each primer pair. The negative control was a non-reverse transcribed (NO RT) sample. Pair-wise *t* tests were used to analyse any significant differences, with no difference seen between the dlp1b/dlp2 spanning region with the negative control, where by a significant difference (p -value < 0.0001) was seen between the dlp1 reaction with the negative control, indicating possible co expression of dlp1a and dlp1b. (N=5)

3.3.3.2 A DNA Sequence Anomaly in the Overlap Between *dlp1a* and *dlp1b*

In *H. pylori* 26695, *dlp1a* and *dlp1b* genes overlap, and it is in this region where a sequence anomaly was identified in this study. The first genome of 26695 (188) identified the sequence of this region to be as seen in *Figure 3.10*, highlighted in red, precisely GG. With this sequence, *dlp1a* and *dlp1b* remain out of frame and would form two separate protein products (*see S1 for nucleotide sequences*). This form of the operon has been termed truncated, with the dynamin protein seemingly truncated from this sequence. More recently, several different 26695 genomes have been published, in particular in 2013, when a new genome was published (189, 190) with many noted changes, one being in this overlapping region of *dlp1a-dlp1b*. This can be seen in *Figure 3.10*, highlighted in red, with a single nucleotide insertion of G. With this single insertion, the two genes are put into frame, to encode a full-length Dlp (*see S1 for nucleotide sequence*). Understandably, the simplest reason behind this might be a sequencing error in the original *H. pylori* 26695 genome compiled using Sanger sequencing in 1997, compared to potentially more reliable genomes using Illumina platform technology in 2013, but we have evidence of both forms of the sequence. This is interesting because, it would be thought that the Illumina platform technology would provide more reliable data, with the use of large numbers of overlapping reads of the same sequence to build the consensus compared to early Sanger sequencing which would of not.

Throughout this work, this region has been amplified, cloned and sequenced independently over 30 times and amongst these clones, both genome sequences (GG or GGG) were found. Of 31 independently sequenced PCR products, 6 have the GG sequence compared to 25 products showing the GGG, indicating there is an approximate 1: 4 ratio seen between the two sequences obtained from these clones (*Figure 3.11A*). This is highly intriguing as this single insertion, results in the conversion of an out of frame register of *dlp1a* and *dlp1b* to an in-frame full-length gene encoding Dlp1. Biologically speaking this takes the “split” dynamin with the GG sequence and turns it into a “full-length” dynamin with the GGG sequence. Interestingly, this sequence change does not alter the important membrane binding properties encoded in this region.

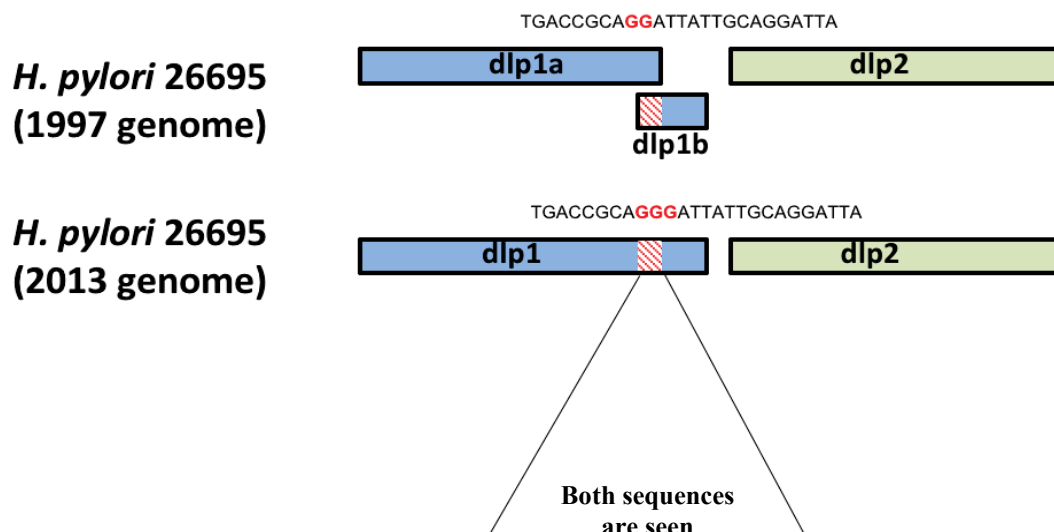
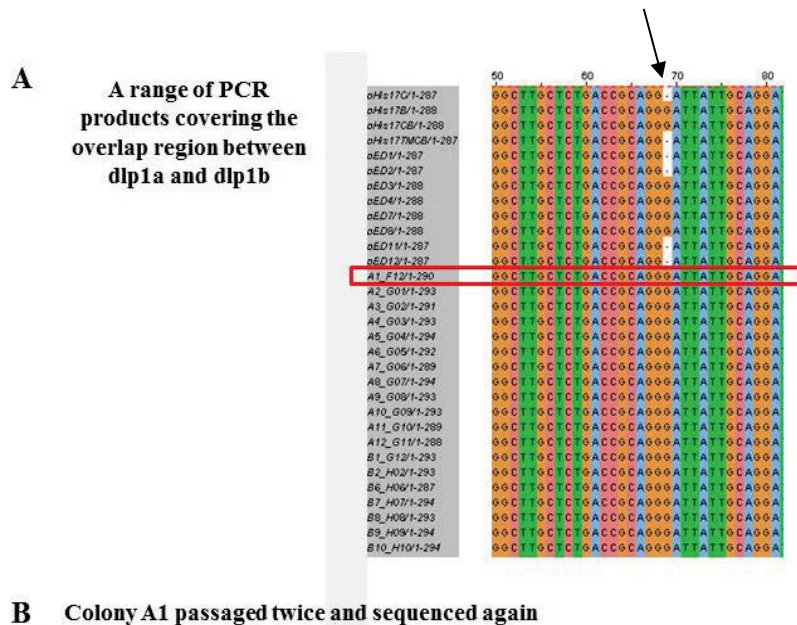


Figure 3.10 Overlapping Region between dlp1a and dlp1b Sequence Anomaly

The original sequence of the *H. pylori* genome 26695 (1997), showed the overlapping region between *dlp1a* and *dlp1b* (red lined box) to have the sequence seen above, with a GG present within the sequence. The same region in the newer 26695 genome sequence (2013), showed a single nucleotide insertion, G, into the sequence (red) now placing the open reading frames into frame forming one full length gene for transcription. Throughout the course of this study, both sequences have been observed (188, 189).

This could be the result of an evolutionary process during laboratory maintenance of *H. pylori* 26695, or it may be the first evidence of phase variation (regulated genotypic variation) controlling the switching of the *dlp* operon from truncated to a full-length dynamin. This warrants further investigation into the rate at which this phase variation is happening and whether it is triggered by an external factor, such as growth conditions or stress. For the purposes of this study, a strain expressing the full-length dynamin was isolated by selection and growing single colonies. After 2 passages of continuous outgrowth, the sequence of the *dlp* region in isolate A1 was found to contain the GGG variant (Figure 3.11B) and this isolate was used in a number of downstream experiments, such as localisation by biochemical fractionation of 26695 cells and antibody testing (Chapter IV).



B Colony A1 passed twice and sequenced again

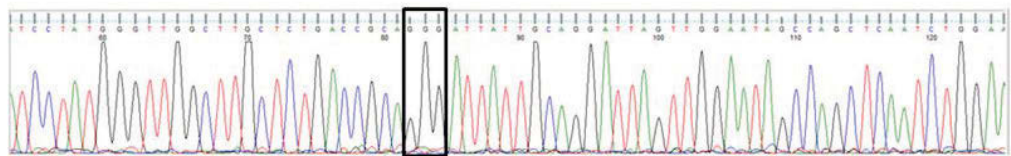


Figure 3.11 Sequencing Evidence of both dlp1a/dlp1b and dlp1

Over the course of this study, 31 separate PCR products of this region have been analysed and sequenced. (A) Alignment of these sequences shows that 6 out of 31 were confirmed to have the genome sequence containing the GG in this region, with all other products having the genome sequence containing the GGG (arrow). (B) One single colony was selected, grown for a further two passages and re-sequenced to confirm the presence of the full-length sequence (GGG) and would be used in subsequent studies.

In summary, this research has laid out the background and context needed to further investigate the dynamins from *H. pylori*. In particular, highlighting their conservation throughout the species, as well as their dynamin-like features, such as a GTPase domain and predicted transmembrane domains. Finally, confirmation of their expression in *H. pylori* provides evidence that they are in fact being used by the bacteria in some capacity. Interestingly, their function could potentially have sequence controlled expression with early evidence of phase variation seen in *dlp1a* and *dlp1b*, the first for bacterial dynamins.

3.4 DISCUSSION

The dynamin superfamily was long thought to be eukaryotic specific, consisting of several different members and carrying out a range of functions in nearly every cell of the body (2). This belief persisted until just over 10 years ago, when the crystal structure of BDLP1 from *N. punctiforme*, a cyanobacterium, was solved showing clear homology to human dynamin, and proving this view to be incorrect (4). This began a slow cascade of bacterial dynamins being identified and characterised such as DynA from *B. subtilis*, IniA from *M. tuberculosis* and LeoA from *ETEC*. LeoA specifically, a protein shown to be involved in vesicle release from *ETEC*, has been identified as a dynamin protein and has led to the establishment of a new subfamily of bacterial dynamins (5, 7), which are termed here Leo-Type dynamins. It has previously been noted that the closest homolog to LeoA is a hypothetical protein from *H. pylori*, HP_0731 and even though there are other bacterial LeoA homologs, no thorough search had been conducted (5, 99). Using a psiBLASTp search with LeoC, LeoB and LeoA as query sequences, each bacterial phylum was searched and, as seen in *Figure 3.1*, homologs were found across a range of phyla, in both Gram-positive and Gram-negative bacteria. Notably, many homologs were found in one particular Gram-negative phylum, the Proteobacteria. There are a number of different sub classes of the Proteobacteria including alpha-, beta-, gamma-, delta-, epsilon- and zeta-Proteobacteria. In this search, homologs of LeoCBA were mainly found in the gamma-proteobacteria and epsilon-Proteobacteria, which includes a wide variety of important human pathogens, such as *Escherichia*, *Salmonella*, *Vibrio*, *Helicobacter*, and many other notable genera. This conservation in these phyla indicates that the dynamins could be important to the survival of many bacterial pathogens.

For the building of the phylogenetic trees in this work, homologs of LeoC and LeoB were manually joined together. This was because previously it had been noted that only together do LeoC and LeoB form a full-length dynamin protein (5). Strength to this statement is highlighted by the distinct branching evidence in *Figure 3.1*, where homologs of LeoA (Dlp2) and LeoCB (Dlp1) form two separate branches with strong phyla grouping of the homologs with their respective phyla in each branch (*Figure 3.1* colour circles).

From this initial investigation of the Leo-Type dynamin members, it can be seen that the conservation of Leo-Type dynamins is spread throughout the bacterial domain. Future analysis that expands on the work presented here, as more sequences become available, may allow a greater understanding as to why certain bacteria have retained these proteins benefit their survival.

With the finding of a diverse range of bacteria containing these Leo-Type dynamins, their place in the greater dynamin superfamily needed to be investigated, to see if they clustered with any other known families of dynamins or whether they form their own sub branch. The dynamin superfamily is well-established to have a number of sub families; classical dynamins, dynamin-like proteins, Mx proteins, Mitofusins, GBPs and Atlastins (2). Previously it has been shown through alignment analysis using the isolated guanine-nucleotide-binding (G-) domains, that some bacterial DLPs such as DynA and BDLP fall into the mitofusin and atlastin class and are more distantly related to classical dynamins (6). *Figure 3.2* supports this view, with DynA1 and BDLP1, clustering with a number of Mitofusins (lower pink circle). With this previously identified relationship again evident in the analysis in *Figure 3.2*, a certain level of confidence was added to this observation.

It was seen that Leo-Type Dlp1 and Dlp2 homologs as well as two newly identified *H. pylori* dynamins, HP_0965 and HP_0966 were on the same divergent branch as GBP and Alastin (top pink circle). These two members of the dynamin superfamily have diverse roles; GBPs are an interferon-inducible family of DLPs involved in antimicrobial and antiviral responses of the cell, particularly in intracellular trafficking, and atlastins are involved in endoplasmic reticulum (ER) fusion processes (77, 79). In terms of the greater dynamin superfamily there is a distinct split between the “dynamin-like proteins” and “GBP-like proteins”, with the defining difference being a lack of a PH domain in the GBP-like proteins that instead modulate membrane dynamics through a transmembrane domain (76). It is therefore not surprising to find Leo-Type dynamins, clustering with these eukaryotic members, as they too have a transmembrane domain that is highly conserved rather than the classical dynamin PH domain (*Figure 3.5*). This branch of bacterial dynamins, with the addition of Leo-Type dynamins with their own distinct branches (Dlp1 and Dlp2), highlights the significant overall diversity in the dynamin superfamily.

H. pylori is an important and understudied human pathogen, with over 300 predicted proteins classified as hypothetical (95). In this study, three of these hypothetical proteins have been shown to have high homology to LeoCBA, not only in one strain but across 249 *H. pylori* strains (Figure 3.3). This previously unseen conservation of a bacterial dynamin across a wide variety of strains with a species highlights the potential importance of these genes to this pathogen. Closer examination of this conservation shows a pattern, with the absence of *dlp1a* bringing about the absence of *dlp1b*. Michie *et al*, stated that LeoC (*dlp1a*) and LeoB (*dlp1b*) were in fact two parts of a single fully functioning dynamin, with LeoC containing part of the G domain and neck, and LeoB containing the rest of the G domain and transmembrane binding domain (5). Therefore, it doesn't come as a surprise, to see the absence of *dlp1b* in the absence of *dlp1a*. However, it seems the reverse doesn't hold true. There were a number of instances where *dlp1b* was not detected in the search but *dlp1a* was. The high-throughput screening method used here should not be considered absolutely definitive in determining the presence or absence of a homolog. However, some strains of *H. pylori* that did not appear to have *dlp1b* might instead have a full-length Dlp1. Further research is needed to confirm this preliminary observation of the *dlp* region genomic structure in many of these strains. Importantly, 136 strains were found to have all three homologs, back-to-back in the genome (consecutive gene names) emphasising the strong tendency for dynamin genes to come in pairs, and adding indirect support to the motion that they may carry out their function as a possible hetero-multimeric complex (3, 5, 6). Phylogenetic analysis of these homologs as seen in Figure 3.4 brings to light that the LeoC and LeoB homologs (Dlp1) are quite distinct from the LeoA (Dlp2) homologs, each forming their own branches. Dlp2 protein sequences in *H. pylori* do not cluster clearly but Dlp1 homologs form 4 main sub groups, with representative strains labelled in Figure 3.4.

When the proteins encoded in the *dlp* operon of 10 *H. pylori* strains were further investigated (Figure 3.5), the Dlp2 protein length was generally conserved with some exceptions including strains G27 and CPY6081, which were slightly shorter and retained homology only to the N-terminal G-domain region of the 26695 Dlp2. There is no evidence if these proteins are still function in their particular strains, but it could be thought that because the GTPase domain is still present, GTP hydrolysis should still be able to be carried out; a necessity for dynamin. Of interest in this study was the varying split point noted between *dlp1a* and *dlp1b* throughout the strains, with no discernible

pattern evident. In contrast, the position of the transmembrane domain was far more consistent. In each of the strains screened, this transmembrane domain was approximately 100 amino acids from the C-terminal end of the protein. This conservation alludes to the split not being of importance to the protein and its downstream processing or regulation. An unusual split in the genes encoding Dlp2 in the *H. pylori* strain B128.7.13 was observed, where three separate genes were detected by whole genome sequencing (*Chapter V*). Despite the genes being split, a ClustalW alignment of the predicted B128.7.13 Dlp2 protein with its homolog in *H. pylori* 26695 shows clear homology across the full-length of protein. Furthermore, a single full-length protein product was detected in the B128.7.13 strain *in vivo* (*Chapter IV*). This suggests that the split in the dynamin encoding genes observed in the sequence data of some *H. pylori* genomes might not accurately represent the biological form of the expressed protein.

As this was the first extensive annotation of DlpS across many *H. pylori* strains, further bioinformatical characterisation was carried out on the 26695 *dlp* operon (*Figure 3.6*). GTP hydrolysis by GTPases requires crucial residues embedded in a conserved G-domain as sequence motifs G1–G5. Interestingly, in some of the recently identified GTPase domains, the motif order is circularly permuted, but this doesn't seem to be the case in dynamins (191). Predicted GTPase domains are evident in Dlp1a and Dlp2, with high conservation across a range of *H. pylori* strains at the key GTP binding residues in the G1 P loop, G2 (Switch I), the G3 (Switch II) regions and G4 binding motif in Dlp2. An interesting finding was the conservation of other residues in this GTP domain such as, E/K and AHxIx (*Figure 3.6*, highlighted purple top panel), which have not been previously noted for GTPase domains in dynamins (6). These are species specific residues for *H. pylori* dynamin GTPase domains with no homology seen to other members of the dynamin superfamily, emphasising the close homology of these proteins across *H. pylori* strains.

The final bioinformatic characterisation carried out on the *H. pylori* dynamins was the 3D structural predictions. A key structural feature of all dynamins, and in *H. pylori* 26695 this is most evident in Dlp1b, is a membrane binding domain. More specifically in the case of bacterial dynamins this is a predicted transmembrane domain. Therefore, a structural prediction was carried out to potentially highlight other important regions of structural homology that have been implicated in the function for dynamin previously (5,

29). Phyre2 is a reputable web-based server to predict 3D protein structures (187). It uses an iterative search to find templates of known structure, and then uses the templates to model the query sequence. Despite relative low sequence homology between ETEC LeoCBA and their homologs in *H. pylori*, their relationship became clearer when the structural predictions were done. Dlp2, as seen in *Figure 3.7*, closely resembles LeoA, with the elongated dynamin form. For the building of this model, one template, LeoA, was used from the protein database by Phyre2, with 92% of residues predicted with greater than 90% confidence. The known dynamin features of a GTPase, neck and trunk domain were evident in this predicted structure with one notable difference in the tip domain. There was a predicted bend to the tip domain, not previously seen in the LeoA structure, this may yet be an artefact of the prediction software, but the difference with LeoA may be resolved with a crystal structure of Dlp2.

Interestingly, one face of the Dlp2 protein showed a higher level of hydrophobicity (*Figure 3.7C*) and may indicate a location or interaction in the cell that requires membrane association; this is more important in the context of Dlp1. When Phyre2 was used to predict the 3D structure of Dlp1 three templates were used, LeoA, BDLP1 and mitofusin -1, with 99% of the residues predicted for Dlp1 done so with a confidence of greater than 90% (*Figure 3.8A*). The predicted structure shows a striking resemblance to BDLP1 from *N. punctiforme* with a 'folded' or 'closed' conformation when compared to LeoA which is termed 'open' (4). Again, all dynamin structural features were annotated; GTPase, neck and trunk domain all present in the predicted structure. One of the more important features of Dlp1 is the transmembrane domain, as seen in *Figure 3.8C-D*, it mapped perfectly to the tip domain where an increase in hydrophobic residues are clustered, the same as BDLP1(4). This could indicate a lipid binding mechanism reminiscent of BDLP1.

Taken together this preliminary bioinformatic analysis into the *H. pylori* dynamins, Dlp1 and Dlp2, highlights that they are highly conserved in this species and that they contain all the necessary features (GTPase domains and transmembrane domains) to be termed bacterial dynamins; more specifically Leo-Type dynamins.

With the establishment of these hypothetically annotated genes as bacterial dynamins, the expression of these genes needed to be confirmed to prove they are in fact not redundant in *H. pylori*. Previously, it has been shown that HP_0731, HP_0732 and HP_0733 form an operon with expression seen in dRNA-seq experiments where a novel approach to selectively identify native 5'PPP ends, a single-nucleotide resolution map of the primary transcriptome of *H. pylori* was generated (192). However, with this being the only evidence of expression in the literature, confirmation of their expression in a more targeted approach was needed; reverse transcription PCR was utilised for this. As seen in *Figure 3.9*, expression of the *dlp* operon was evident with cycle threshold (Ct) values between 25-30. Because the focus was the identification of the presence and/or absence of the *dlp* operon transcripts, and not quantification of the expression, no standards were used. From this study, the first evidence of co expression in the *dlp* operon was evident, with significant levels of transcripts being detected in the region that specifically spans *dlp1a* and *dlp1b*, which was not seen in the region between *dlp1b* and *dlp2* (*Figure 3.9*). This preliminary data was quite exciting and supported the previously outlined notion that *dlp1a* and *dlp1b* both contribute to a full-length dynamin. Further interrogation of this region of the genome, brought to light that the genes actually overlap slightly and in this overlap there was a single nucleotide indel. As shown in *Figure 3.10*, the first *H. pylori* 26695 genome was published in 1997 and had these two genes out of frame, forming two separate transcripts (188). In 2013 a new 26695 genome was sequenced using more updated technology and there was a single nucleotide insertion in this overlapping region (189, 190). These new data placed the genes into the same open reading frame, forming a single transcript; a full-length dynamin. Evidence would lean towards the 2013 *H. pylori* 26695 genome being the correct sequence, with a more controlled and robust sequencing technique used, but this is not the case. Throughout this work, this region has been amplified for cloning purposes numerous times independently and, as seen in *Figure 3.11*, the sequence detected in this overlap region differs by this single nucleotide on 6 occasions, at a 1: 4 ratio. This single nucleotide polymorphism (SNP) could be the first evidence of phase variation controlling gene expression in the *dlp* operon.

To establish long-term infections, bacterial pathogens like *H. pylori* must adapt their gene expression through sensing and responding to changing conditions or by the selection of dominant genotypic variations within the population. Hypervariable simple sequence repeats (SSRs), potentially like the GG/GGG seen in the *dlp* operon, are a major source

of this so-called phase variation and are often associated with genes involved in host interaction (193). Phase variation is mediated by slipped-strand mispairing leading to insertions or deletions (indels) within repetitive DNA tracts during replication. When an indel-modified tract arises in the coding region of the gene, these indels result in an altered reading frame (194). Phase variation is not a new concept in the *H. pylori* genome with a number of examples exhibiting this type of gene regulation (193, 195-200). However, the clear difference here is that a two-three base phase variation, is very short and therefore may be more an evolutionary switch occurring in the lab. Further investigation into this potential phase variation is essential in establishing if it's an active switch or if its evolution driven.

Establishing an understanding of bacterial dynamins in *H. pylori* has been the focus of this study, highlighting their conservation and their features as well as establishing their *in vivo* expression and potential control. However, deeper investigation will help elucidate the true conservation of the Leo-Type dynamins with a more comprehensive survey of the bacterial and archaeal phyla, which may provide more evidence of their ancestral function. Further characterisation and confirmation of the differences seen in the *H. pylori* *dlp* operon, both in conservation as well as structure, may help in understanding why this organism has retained these genes throughout its evolution. With this level of conservation, it could be hypothesised that the *dlp* operon is being positively selected for by *H. pylori* throughout its evolution and future studies will investigate this possibility. Bioinformatic tools for structure predications have their limitations but they provide invaluable information for the initial stages of this type of work establishing the main structural domains present. Crystal structures and other advanced techniques like single molecule cryo-electron microscopy will provide the necessary structural information to help guide future biochemical studies. Finally, the regulation of gene expression, especially in a human pathogen like *H. pylori* is key to its virulence and survival. Therefore, further elucidating the potential role that the genotypic variation observed here may have in the *dlp* operon expression, through an approach that allows for the quantification of the mRNA transcripts produced in the cell, such as RNAseq, is crucial to understanding the role that these potential membrane remodelling proteins have in this important human pathogen.

CHAPTER IV

INVESTIGATION INTO THE BIOCHEMICAL PROPERTIES OF DLP1 AND DLP2

PREFACE

In Chapter III of this thesis, it was demonstrated that in *H. pylori* there is a high level of conservation of dynamin-like proteins that have significant homology to the previously characterised ETEC LeoCBA genes and these have been termed Dlp1a, Dlp1b and Dlp2 being homologous to LeoC, LeoB and LeoA respectively (*Section 3.3.1*). Through the use of bioinformatics, these genes and subsequent protein products were shown to exhibit the important features of dynamin, such as a predicted transmembrane domain and structural homology to previously characterised bacterial dynamins (*Section 3.3.2*). Their hypothetical status in the genome was challenged with *in vivo* detection of mRNA transcripts for all genes, along with possible co-transcription evident for *dlp1a* and *dlp1b*. Furthermore, phase variation of *dlp1a* and *dlp1b* has been shown to be possible with the presence of a single nucleotide insertion that places *dlp1a* and *dlp1b* into frame forming one dynamin, *dlp1* (*Section 3.3.3*). With this preliminary bioinformatic characterisation of the presence of Dlp1s in *H. pylori* and their expression proven true, biochemical characterisation of the Dlp1s from *H. pylori* will be presented in the following chapter.

4.1 INTRODUCTION

Biological membranes are an essential component of cellular life, providing a semi-permeable barrier to ensure intracellular homeostasis is maintained and are thus designed to be incredibly dynamic to allow the necessary changes. Membrane remodelling is a key morphological change that must occur for several cellular processes to be carried out by the cell, including but not limited to endo- and exocytosis. The generation of membrane vesicles in these processes requires the recruitment of various proteins that will orchestrate the pressure and bending of the membrane (1). Large GTPase proteins have often been attributed to mediating these types of membrane dynamic events, most notably members of the Dynamin Superfamily of proteins (2, 201). Dynamin is an essential mechano-chemical enzyme that assembles into helical polymers at the neck of a budding vesicle where a GTP hydrolysis dependent conformation is required to pinch off the vesicle (1). The means by which this essential conformation changes occurs is still highly debated in the literature (33).

Despite being identified over 10 years ago, bacterial dynamins have largely been neglected, with very diverse and unclear functional phenotypes noted. Until now, several bacterial members of the dynamin superfamily had been identified, and only a few of these including, BDLP1, DynA and LeoA, have been studied *in vitro* (3-5, 7). Their *in vitro* properties have been established based on purified recombinant proteins and their ability to bind to different substrates such as nucleotide and lipids. Regardless of these bacterial dynamins showing significant differences in their sequence (< 30% homology), they maintain the fundamental features of the dynamin superfamily including; structural/domain homology, lipid binding mechanisms, self-interaction and GTPase activity, all of which contribute to their essential function in membrane dynamics.

With the solving of the structure of BDLP1, clear dynamin superfamily features were evident. It was able to be shown that BDLP1 has GTPase activity and is able to self-assemble and tubulate lipid bilayers, a feature reminiscent of eukaryotic dynamins (4, 91). With no known cellular function established for BDLP1, there is still a vast amount of research needed to be carried out on its biochemical properties including localisation to the membrane and a mechanism as to how it inserts and shapes the membrane in a GTP

dependent conformational change (91). However, since the discovery of BDLP1, there has been further research into other dynamins found in bacteria and of particular interest to this study are DynA and LeoA.

Despite not having a crystal structure DynA, a fused bacterial dynamin (two complete dynamin subunits) found in *B. subtilis* also has been shown to be a dynamin based upon its biochemical properties. DynA has a well-established GTPase activity that is dependent on both GTPase domains in each subunit being able to bind GTP to carry out hydrolysis (3). DynA can not only bind to liposomes but it can tether them together, indicating a possible function in membrane fusion which has been previously noted in eukaryotic dynamins, specifically Mitofusions and their role in mitochondria fusion (3, 100, 202). Furthermore, an interaction screen has been conducted using BacTH showing that the fused proteins as well as the individual subunits D1 and D2 have the ability to bind to each other, potentially showing evidence for polymerisation, a necessary feature for classical dynamin function (3, 100). Additionally, localisation studies have shown that a punctuated pattern at the membrane is seen for WT DynA that is dependent on both subunits being present, with only D1 showing binding to the membrane when studied individually (3).

The other bacterial dynamin of interest for this study is LeoA, a bacterial dynamin found in ETEC. LeoA was shown to be a bacterial dynamin through the solving of its crystal structure with clear similarities to the dynamin superfamily (5). Minimal research has been carried out on the biochemical properties of LeoA with no lipid binding studies conducted despite the presence of a tip domain similar to BDLP (5). However, preliminary localisation studies have shown a punctuated pattern, similar to DynA and BDLP (5). Interestingly, despite being proven to be a bacterial dynamin, there is ongoing contention in the literature as to its ability to exert GTPase activity. In earlier studies, it was shown to have dynamin-like levels of GTPase hydrolysis which were abolished when a GTPase active site mutation was incorporated (7). Recently however, it was reported to have no ability to hydrolyse GTP or to even bind to GTP at detectable levels (5). This allowed the authors to suggest that the former study was subject to contaminating GTP and hypothesise that LeoA may require its operon partners, LeoC and LeoB to effectively bind to and hydrolyse GTP, an interesting concept for dynamins yet to be proven. With limited ongoing research conducted on bacterial dynamins, further study of these

seemingly important proteins is needed to further elucidate what their function in bacteria is and how they carry it out biochemically.

The closest homologs to the ETEC dynamins LeoCBA are found in *H. pylori*, and are highly conserved throughout the species and have not been studied previously. Studying them in this organism, with their high conservation (Chapter III), allows a deeper understanding into their fundamental biology in this important human pathogen. A pathogen known to cause infection through membrane dynamic changes, such as through the release of toxin filled outer membrane vesicles, which are in turn important for host-pathogen interactions as well as membrane changes needed in the adaption to the harsh environment of the stomach. Establishing that the bacterial dynamin-like proteins in *H. pylori* have the capability to carry out a membrane-targeted function in a GTP-dependent fashion is an important step in characterising them as part of the dynamin superfamily and will be potentially applicable to other bacterial dynamins of this family. This understanding will also widen the general understanding of the biochemical features of dynamin superfamily.

The aims of this chapter were to;

- Express and purify Dlp1a-6H and Dlp2-6H with several downstream purposes including antibody production and lipid binding tests.
- Assess the ability of antibodies raised against Dlp1a-6H and Dlp2-6H to bind to the expressed proteins in *H. pylori* whole cell lysates
- Investigate the localisation of Dlp1 and Dlp2 in *H. pylori*
- Assess the ability of Dlp2 to bind to a lipid membrane in a GTP-dependent manner
- Investigate whether interactions occur between the different Dlps from *H. pylori*

4.2 MATERIAL AND METHODS

4.2.1 *pHis17*-DLP CONSTRUCTION

Using standard cloning methods as outlined in *Section 2.5.6*, the full open reading frames of Dlp1a and Dlp2 were cloned into pHis17 in such a way that each protein was tagged on the C-terminus by a poly-histidine tag (*Figure 4.1*). In summary, the primers listed below in *Table 4.1* were used in combination to amplify the gene of interest as well as add the necessary restriction sites NdeI and BamHI to each end. Both the PCR product and the pHis17 plasmid were digested with NdeI and BamHI, ligated overnight and transformed into DH5 α as outlined in *Section 2.4.3*. Transformants were selected on LB agar (*Table 2.3*) with ampicillin (100 μ g/ mL) and confirmed using PCR (*Section 2.5.4*) and digestion of the newly constructed plasmid. The sequence of the inserted PCR product was confirmed using Sanger Sequencing (*Section 2.5.5*).

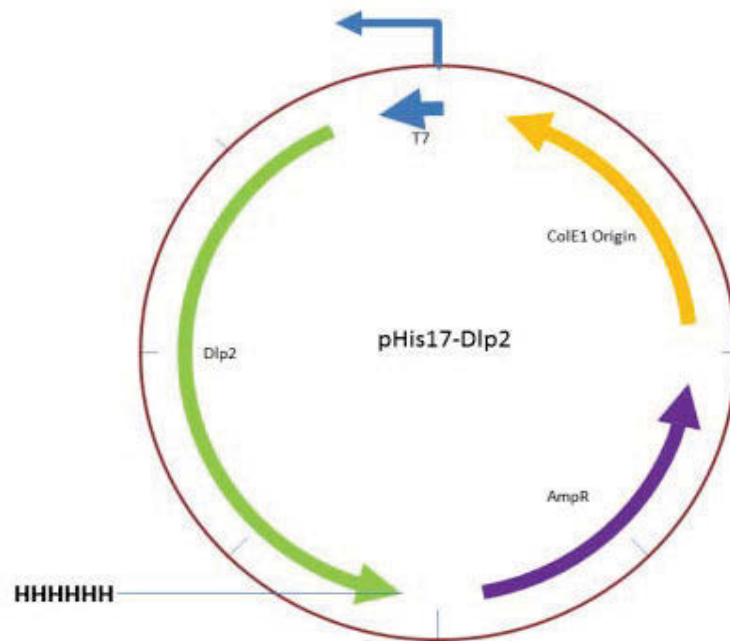


Figure 4.1 *pHis17 – Plasmid Map*

The dlp gene (green) was amplified and cloned into the pHis17 backbone, containing ampicillin resistance (purple) and a ColE1 origin of replication (yellow). This was done in such a way that when the T7 promotor (blue) is switched on, the expressed protein is tagged on the C-terminus with a six-histidine tag (labelled).

Table 4.1 Primers Used for Amplification of Dlps for pHis-17 Plasmid construction

Primer Name	Sequence
IS_LeoC_F	GGG CCC GGG GAT CCT CAT ATG AAA AAC ATT TAT CTT GAT GTG AAA GCC
IS_LeoA_F	GGG CCC GGG GAT CCT CAT ATG AAG AAT GAA ACG CTA GAA CAA TTT AAA
LeoA_R	GGG GTA CCG GAT CCG TTC ATG CCT CTT TCA GGC TC
LeoC_R	GGG GTA CCG GAT CCA TCC TGC AAT AAT CCT GCG GTC

4.2.2 OVER EXPRESSION AND PURIFICATION OF DLP2-6H

BL21-AI cells were transformed with pDlp2-His as previously outlined in *Section 2.4.3*. A colony was inoculated into a 5mL starter culture of LB (*Table 2.3*) with ampicillin (100 µg/mL) and incubated at 37°C with shaking (180 rpm) overnight. To upscale, two replicate flasks containing LB (*Table 2.3*) with ampicillin (100 µg/mL) (1.6 L), were inoculated with the overnight culture. The cultures were incubated at 37°C with shaking (180 rpm) and the OD_{600nm} monitored until mid-log phase (OD_{600nm} 0.4) was reached, after which the cultures were then placed in an incubator at 18°C with shaking (180 rpm), left for 15 min and a “pre-induction” sample of the culture (1 mL) taken. Expression of Dlp2-6H is under the control of an arabinose inducible promotor and thus the cultures were induced with 0.2 % (w/v) arabinose and left overnight at 18°C with shaking (180 rpm), after which the OD_{600nm} of the cultures was assessed and a “post induction” sample was collected. The cells were then centrifuged using a Hitachi R9A rotor at 7000 x g for 20 min at 4°C before proceeding immediately onto disruption using sonication on ice (20 x 20 sec intervals, 40% maximum output, VC 50T) followed by centrifugation at 13,000 x g for 30 min at 4°C.

4.2.2.1 His Trap Purification

The supernatant of the sonicated culture was isolated and Dlp2-6H was purified using an AKTA™ Explorer Fast Protein Liquid Chromatography (FPLC) system (GE Healthcare Life Sciences) at a flow rate of 1 mL/min. A 5 mL HisTrap HP column (GE Healthcare Life Sciences) containing immobilized Ni²⁺ was equilibrated with 2 x TBS (*Table 2.5*). The sample containing Dlp2-6H was circulated through the column using a peristaltic

pump for 1 h on ice to allow protein binding. The column was reattached to the FPLC system and washed with 2 x TBS (*Table 2.5*) to remove unbound or weakly bound protein. A gradient of 0 M-1 M imidazole, over the course of 10 column volumes (50 mL) was then performed. Fractions (2 mL) were collected and screened via SDS-PAGE and protein-containing fractions pooled (*Section 2.5.1*).

4.2.2.2 Ion Exchange Purification

Dlp2-6H has an isoelectric point of pH 5.92 and was therefore able to be further purified using ion exchange chromatography. The Dlp2-6H sample was prepared by dialysis overnight into 25 mM Tris-HCl (pH 8) and recovered by centrifugation (10 000 x g, 30 min, 4°C). Dlp2-6H was further purified using an AKTA™ Explorer Fast Protein Liquid Chromatography (FPLC) system (GE Healthcare Life Sciences) at a flow rate of 1 mL/min on a 5 mL HiTrap Q HP column (GE Healthcare Life Sciences) equilibrated with 25 mM Tris-HCl (pH 8.0). The sample containing Dlp2-6H was circulated through the column using a peristaltic pump for 1 h on ice to allow protein binding. The column was reattached to the FPLC system and washed with 25 mM Tris-HCl (pH 8.0) to remove unbound or weakly bound protein. A gradient of 0 M-1 M NaCl, over the course of 10 column volumes (50 mL) was then performed. Fractions (2 mL) were collected and screened via SDS-PAGE and protein-containing fractions pooled (*Section 2.5.1*) and run over the HisTrap HP column once again as outlined above in 4.2.2.1. Fractions collected were pooled, dialysed overnight into 2 x TBS (*Table 2.5*) and concentrated using protein concentrator columns (10,000 MW, GE Biosciences) at 3660 x g, 4°C in 30 min spins until the desired concentration was achieved. The final concentration was determined by the Nano Drop 100 Spectrophotometer at 280 nm at a distance of 10 mm and using the equation $A = \epsilon \cdot c \cdot l$ (where A is the absorbance (at 280nm), ϵ is the protein specific extinction coefficient, c is the concentration and l is the path length). The extinction coefficient, for Dlp2-6H, specifically was calculated to be 53 400, using ProtParam (203)

4.2.2.3 Gel Filtration Purification

Later in the project, affinity purified fractions containing Dlp2-6H were further purified using a gel filtration column, HiPrep 26/60 Sephacryl S-300 h at 4°C, on a AKTA Pure FPLC chromatography system with 2 x PBS (*Table 2.5*). The protein sample (5 mL) was loaded with a flow rate of 0.5 mL/min and 5 mL fractions were collected. Fractions containing protein as per the UV_{280nm} readings, were analysed by SDS-PAGE as described in *Section 2.5.1*.

4.2.3 OVER EXPRESSION AND PURIFICATION OF DLP1A-6H

The over expression of Dlp1a-6H was carried out as previously outlined in 4.2.2 with the pDlp1a-His plasmid.

4.2.3.1 Inclusion Body Purification

The inclusion body protocol was adapted from “*Application Note – Rapid and Efficient purification and refolding of a His Tagged recombinant protein produced in E. coli as inclusion bodies*”. The crude pellet from BL21-AI cells over-expressing Dlp1a-6H was resuspended in Resuspension Buffer (*Table 2.5*). The cells were then disrupted using sonication on ice (20 x 20 sec intervals, 60% maximum output, VC 50T) and centrifuged at 13,000 x g for 10 min at 18°C. The pellet was resuspended in cold Sonication Buffer (*Table 2.5*) (30 mL) and sonicated (15 x 20 sec intervals, 60% maximum output, VC 50T). The suspension was then spun at 13,000 x g for 10 min at 18°C. The pellet containing inclusion bodies was washed in Resuspension Buffer (*Table 2.5*), and stored at -80°C before use. As a measurement of purification efficiency, samples were collected from each supernatant and pellet resuspension step and assessed via SDS-PAGE (*Section 2.5.1*). The inclusion body pellet was suspended in Solubilisation Buffer (*Table 2.5*) (40 mL) and placed on a roller for 2 h. The sample was then centrifuged at 13,000 x g for 2 h at 4°C to remove any particles and the supernatant containing crude solubilised Dlp1a-6H collected.

The crude solubilised Dlp1a-6H was purified using an AKTA™ Explorer Fast Protein Liquid Chromatography (FPLC) system (GE Healthcare Life Sciences) at a flow rate of 1 mL/min. A 5 mL His Trap HP column (GE Healthcare Life Sciences) containing immobilized Ni²⁺ was equilibrated with Solubilisation Buffer (*Table 2.5*) (10 mL). The sample containing Dlp1a-6H was circulated through the column using a peristaltic pump for 1 h on ice to allow protein binding. The column was reattached to the FPLC system and washed with Solubilisation Buffer (*Table 2.5*) (10 mL) to remove unbound or weakly bound protein. Washing Buffer (*Table 2.5*) (10 mL) was then pumped through the column. Refolding of the bound protein to its native state was performed using a linear 6-0 M urea gradient (over 150 mL) beginning with Washing Buffer (6 M Urea) and finishing with Refolding Buffer (*Table 2.5*). Once the gradient had reached its endpoint, the system was washed with Refolding Buffer (*Table 2.5*). A linear gradient was then performed (over 20 mL) to elute the folded protein, starting with Refolding Buffer (*Table 2.5*) and ending with Elution Buffer (*Table 2.5*). Fractions (2 mL) were collected into tubes and screened via SDS-PAGE (*Section 2.5.1*) and protein-containing fractions pooled. The sample, containing purified Dlp1a-6H protein was dialysed into 2 x TBS (*Table 2.5*), quantified by a BSA standard curve on a SDS-PAGE gel (*Section 2.5.1*) and frozen at -80°C until use.

4.2.4 PRODUCTION OF DLP1A-6H AND DLP2-6H ANTIBODIES

Purified Dlp1a-6H and Dlp2-6H were sent to South Australia Health and Medical Research Institute (SAHMRI) for two rounds of immunization of a rabbit model for polyclonal antibody production. In summary, 0.5 mg and 2 mg of Dlp1a-6H and Dlp2-6H respectively were injected twice, 6 weeks apart, to ensure a high immune response to the antigen, into two rabbits. Test bleeds were checked after the first round (data not shown), to test for an immune response and to allow for any changes required for the second round to be made. The final polyclonal serum was then collected at 12 weeks post inoculation.

4.2.5 TESTING OF ANTIBODIES ON *H. PYLORI* WHOLE CELL LYSATE (WESTERN BLOT)

SDS-PAGE (Section 2.5.1) was carried out with samples of interest and Bio-Rad Pre-Stained Precision Plus Standard (10 μ L) were used as mass markers. For the Dlp2 blot, Dlp2-6H (~13 ng) and for the Dlp1 blot, Dlp1a-6H (~50 ng) and *E. coli* over expression cell lysate for Dlp1 were run as controls. A sheet of nitrocellulose membrane (Whatman Protran®, Sigma Aldrich) and the gel were then sandwiched between filter paper and sponge pads as per the instructions supplied with the transfer apparatus (Mini-Trans Blot® Cell, Bio-Rad). The sandwich cassette was placed in the transfer tank, filled with ice-cold Transfer Buffer (Table 2.5) and electrophoresed for 1 h at 100V. The membrane was removed and stained with Ponceau S (Table 2.5) to assess the transfer. The stain was decanted and the membrane destained with milli-Q water (MQW). The membrane was imaged as a record of transfer efficiency and incubated with blocking solution (Table 2.5) overnight at 4 °C. The blocking solution was discarded, and the membrane incubated with primary antibody from Section 4.2.4 (Table 4.2) for 60 min with slow rocking. The membrane was washed 3 times in TBST (Table 2.5) for 5 min. Secondary antibody (Table 4.2) was then added and incubated for 1 h with slow rocking. The membrane was washed 5 times in TBST (Table 2.5) for 5 min each wash. The HRP activity was developed using Pierce™ ECL Western Blotting Substrate (Thermo-Fischer) as per manufacturer's instructions, and chemiluminescence visualised using Amersham Imager 600 (GE Biosciences).

Table 4.2 Antibodies Used throughout this Work

Name	Concentration	Source
Dlp2 1° Antibody (Rabbit Dlp2 Anti Serum)	1:1000 dilution in TBST	This Study
Dlp1a 1° Antibody (Rabbit Dlp1a Anti-Serum)	1: 500 dilution in TBST	This Study
2° Antibody (Anti-rabbit IgG HRP conjugated 1 mg/mL)	1: 2000 dilution in TBST	Abcam.

4.2.6 SUBCELLULAR FRACTIONATION

Subcellular fractionation was carried out by my colleague Nina Sulkowski. *H. pylori* 26695 was grown as described in *Section 2.4.1*. Protease inhibitors (cOmplete™ Protein Inhibitor Cocktail, Roche) were incorporated into buffers before osmotic lysis and sonication. Cells were harvested at 4000 x g for 10 min at 4 °C. The pellet was washed twice in 1 x TBS (*Table 2.5*) and resuspended in Sucrose Buffer (*Table 2.5*). The cell suspension was centrifuged at 4000 x g for 15 min at 4°C. The pellet was resuspended in cold Lysis Buffer (*Table 2.5*) and placed on ice for 10 min. The cells were spun at 13,000 x g for 5 min at 4 °C and the supernatant (periplasmic fraction) collected. The pellet, containing periplasm-free cells was washed carefully in 1 x TBS (*Table 2.5*), resuspended in 10 mL of 1x TBS (*Table 2.5*) and lysed by sonication on ice (15 x 15 sec intervals, 40% maximum output [VC 50T, Sonic]). The crude extract was then subjected to ultracentrifugation at 250,000 x g for 1 h at 4 °C. The supernatant (soluble fraction) was collected and the pellet washed twice with 1 x TBS (*Table 2.5*). The supernatant was discarded and the pellet (membrane fraction) was resuspended in 100 µL of SDS-LB (*Table 2.5*). The soluble fraction was concentrated using a protein concentrator column (10,000 MW, GE Biosciences). Fractions were aliquoted and stored at -80°C before use. The protein concentration in each fraction was interpolated from a BSA standard curve using a short-run SDS-PAGE method. Briefly, a 4-12 % Criterion™ TGX Gel (Bio-Rad) was loaded with serially diluted BSA standards (0-1.0 mg/mL, 10 µL) and each fraction (10 µL). Electrophoresis was carried out for 5 min at 300 V. The gel was stained with Coomassie Blue and destained with 1% acetic acid. The gel was imaged using Typhoon FLA 3500 (GE Healthcare Life Sciences) at 50 µm pixel resolution and visualised using QuantityOne 1-D Analysis Software (Version 4.6.9 Basic, Bio-Rad). The protein concentration of each sample was then calculated on the basis of pixel density using QuantityOne as per manufacturer's instructions. Normalised samples were then analysed by western blot (*Section 4.2.5*).

4.2.7 IMMUNO-FLUORESCENCE MICROSCOPY (IFM)

H. pylori 26695 and B128.7.13 was grown to mid-log phase as described previously (Section 2.4.1). Cells (1 mL) were harvested by centrifugation at 5000 x g for 5 min. The pellet was then washed once in 1 x PBS (Table 2.5), fixed by suspension in 2 % PFA+ PBS (Table 2.5), and resuspended in 1 x PBS (1 mL) for 1 h. During the fixation period, a 15-well Multi-test slide (Premium grade, MP Biomedicals) was prepared by dispensing 0.1 % poly-D-Lysine (10 µL) into each well and incubating for 2 min. The liquid was then aspirated and wells washed twice with water and allowed to dry. The fixed cells were harvested by centrifugation and resuspended in GTE Buffer (Table 2.5) + Lysozyme (Sigma) at a range of concentrations (0, 1, 10 µg/mL). The cells (10 µL) were immediately added to the Multi-test wells and incubated for 3 min. The liquid was aspirated and wells washed 5 times with 1 x PBS (Table 2.5). Affinity purified primary antibodies (Table 4.2) were diluted in Antibody Buffer (Table 2.5) at varying concentrations (0, 1, 5, 10, 50 µg). The wells were incubated with primary antibody (Table 4.2) (10 µL) for 2.5 h in a humidified slide chamber box. The wells were then washed 5 times using 1 x PBS (Table 2.5) + 0.1% Tween 20. Secondary Antibody, Alexa-Fluor 488 Anti-Goat Anti Rabbit IgG (Invitrogen), was diluted in Antibody Buffer (Table 2.5), and 10 µL added per well. The slides were then incubated for a further 1.5 h in a slide chamber (in the dark). The wells were then washed 5 times with 1 x PBS (Table 2.5) + 0.1 % Tween 20, followed by 5 times in 1 x PBS (Table 2.5) only. The FM4-64 membrane stain (Sigma-Aldrich, USA) was diluted in 50% glycerol + 1 x PBS (Table 2.5) to a concentration of 4 µg. The membrane stain was added to each well and a clean coverslip (#1.5, 0.17 mm thickness) placed on top of the slide. The edges of the coverslip were sealed using nail polish. Microscopy was then carried out using a Zeiss (Axioplan 2 Model) fluorescence microscope, with a 1.4 NA 100x phase contrast objective lens, under oil immersion. The Zeiss FITC and Rhodamine fluorescence filter-sets with 500 ms and 40 ms exposures were used for visualisation of the secondary antibody and FM4-64 fluorescence, respectively. The AxioCam MRm camera was used for image collection. Images were processed and analysed using AxioVision software, version 4.8 (Zeiss).

4.2.8 LIPOSOME BINDING ASSAY

10 mg of total *E. coli* lipid extract (Avanti Polar Lipids) was dissolved in chloroform and air dried using N₂ gas, forming a thin layer of lipids on the bottom of a glass tube. Glass tubes were placed in a vacuum chamber overnight and stored for up to a week at 4°C. To form liposomes, the lipid film was hydrated to final concentration of 1mg/mL using filter sterilized liposome buffer (*Table 2.5*). Medium sized glass beads (4 in total) were added to the tube and vortexed for 15 min to re dissolve the lipids before 30 min incubation in a sonication bath was performed. To 10 µL of 1 mg/mL liposomes, 10 µM Dlp2-6H and 1 mM GTP was added and incubated with rolling at 4°C for 1 h. Liposomes were then stained with the membrane stain FM4-64 (Sigma-Aldrich, USA) and a clean coverslip (#1.5, 0.17 mm thickness) placed on top of the slide. Microscopy was then carried out using a Zeiss (Axioplan 2 Model) fluorescence microscope, with a 1.4 NA 100x phase contrast objective lens, under oil immersion. The Zeiss Rhodamine fluorescence filter-set with 50 ms exposure, was used for visualisation FM4-64 fluorescence. The AxioCam MRm camera was used for image collection. Images were processed and analysed using AxioVision software, version 4.8 (Zeiss).

4.2.9 CO-PELLETING AND CO-SEDIMENTATION ASSAY

For the co-pelleting assay, purified Dlp2-6H at the final concentrations of 15 µM and 30 µM was pre-spun at 57 000 x g in a Beckman TLA100 rotor for 30 min, supernatants transferred to new tubes and 2.5 mM GTP added. Next, the tubes were centrifuged at 57 000 x g at 20°C for 30 min. The supernatants were removed for analysis and the pellets washed with buffer containing the same components as the reaction buffer aside from proteins. Pellets were solubilised with SDS-LB (*Table 2.5*) (same volume as for the supernatants) and samples were analysed by SDS-PAGE (*Section 2.5.1*) (204).

For the co-sedimentation assay, liposomes were prepared using *E. coli* total lipid extract (Avanti Polar Lipids) using sonication as outlined in *Section 4.2.8*. Liposomes (1 mg/mL) were mixed with pre-spun Dlp2-6H at final concentrations of both 15 µM and 30 µM and centrifuged at 40 000 x g at 20°C for 15 min in a Beckman TLA100 rotor. The supernatants were removed for analysis, the pellets washed and solubilised with SDS –

LB (*Table 2.5*) (same volume as for the supernatants) and samples were analysed by SDS–PAGE (*Section 2.5.1*) (204).

4.2.10 BACTERIAL TWO HYBRID SYSTEM

Using standard cloning methods as outlined in *Section 2.5.6*, the full open reading frames of Dlp1, Dlp1a, Dlp1b, Dlp2, a truncated Δ TM Dlp1b and a truncated Δ TM Dlp1 were cloned into both a T25 and T18 tagged plasmid with the enzymes KpnI and BamHI. Dlp1, Dlp1a, Dlp2 and truncated Δ TM Dlp1 were tagged on the C-terminus using the vectors pUT18C and pKT25 (*Table 2.2*) and Dlp1b and truncated Δ TM Dlp1b were tagged on the N-terminus using the vectors pUT18 and p25N (*Table 2.2*.) A list of primers used for the amplification and sequencing confirmation of these open readings frames are listed below in *Table 4.3*.

Table 4.3 Primers Used for Amplification of Dlps for BacTH Plasmid construction

Name	Primer Sequence
LeoC_F	GGG CCC GGG GAT CCT CAT ATG AAA AAC ATT TAT CTT GAT GTG AAA GCC
LeoC_R	GGG CCC TGG GAT CCC GGT ACC CTA ATC CTG CAA TAA TCC TGC GGT
noTM_LeoB_F	GGG CCC GGG GAT CCT CAT ATG TCA AGA TAT CAA AGA TCC CAA CAA AG
LeoB_F	GGG CCC GGG GAT CCT CAT ATG GAA TCC TAT GGG TTG GCT
LeoB_R	GGG CCC TGG GAT CCC GGT ACC TCA TTC TTC ATT GCA TGC TCC TTG
LeoA_F	GGG CCC GGG GAT CCT CAT ATG AAG AAT GAA ACG CTA GAA CAA TTT AAA
LeoA_R	GGG CCC TGG GAT CCC GGT ACC TCA GTT CAT GCC TCT TTC AGG
OLE_LeoB_F	GCAGGATTATTGCAGGATTAGTTTCAAGATATCAAAGATC CCAACAAAG
OLE_LeoC_R	AAC TAA TCC TGC AAT AAT CCT GC
Seq-5'UT18C	GAA AAG CCT GTT CGA CGA TG
Seq-3'UT18C	ATC AGA GCA GAT TGT ACT GAG AGT G
Seq-5'KT25	ATT CGG TGA CCG ATT ACC TG
Seq-3'KT25	AGT CAC GTC GTT GTA AAA CGA CG
Seq-VRF83	TGT GGA ATT GTG AGC GGA TA
Seq-VRF84	GTA ACC AGC CTG ATG CGA TT

Seq – primers were used in confirmation by sanger sequencing of the plasmid

Ultra-competent *E. coli* BTH101 cells were co-transformed using electroporation, where cells were mixed with both a T25 and a T18 plasmid combination of interest (*Table 4.4*) and added to a pre-chilled 0.2 mm electroporation cuvette (Bio-Rad, USA). The cells were then electroporated using a PowerPac Basic (Bio-Rad, USA) electroporator at 25 μ F (capacitance), 2.5 kV (voltage) and 200 ohm (resistance). Each transformation was then allowed to grow and recover for 2 h at 37°C in SOC media (*Table 2.3*) with shaking. They were then plated onto indicator plates (*Table 2.3*) with the addition of Amp (100 μ g/ml) and Km (50 μ g/ml) for selection and X-Gal for indication of the interaction between the induced proteins. A double transformation of empty vectors served as negative control and a double transformation of pUT18C-zip and pKT25_zip served as a positive control. After 3 days of incubation at 30°C, colonies began to appear, and plates were photographed.

Table 4.4 Testing Panel for BacTH Assay

The protein of interest was tagged on the N- (dark colour) or C- terminus (light). Each transformation/ interaction pair was allocated a number 1-36 for identification purposes.

BacTH Testing Panel		T18 Tagged					
		Dlp1a	Dlp1b	Dlp2	Dlp1	Dlp1b Δ TM	Dlp1 Δ TM
T25 Tagged	Dlp1a	1	7	13	19	25	31
	Dlp1b	2	8	14	20	26	32
	Dlp2	3	9	15	21	27	33
	Dlp1	4	10	16	22	28	34
	Dlp1b Δ TM	5	11	17	23	29	35
	Dlp1 Δ TM	6	12	18	24	30	36

Bacterial cells to be assayed for β -galactosidase activity indicating positive transformants were grown in 5 mL of LB (*Table 2.3*) in the presence of both 0.5 mM IPTG (isopropyl β -D-thiogalactopyranoside) and the appropriate antibiotics, Amp (100 μ g/ml) and Km (50 μ g/ml) overnight. These cells were then diluted 1 in 5 with M63 medium (*Table 2.5*) and the OD_{600nm} recorded. Cells were then made permeable by adding 30 μ L of toluene and 30 μ L of 0.1% SDS. The tubes were then vortexed for 10 sec, lightly plugged with cotton and vigorously agitated in a shaker at 37°C for 40 min.

For the enzymatic reactions, aliquots of 0.1 to 0.5 mL of the permeabilised cells were added to 0.9 - 0.5 mL of PM2 buffer (*Table 2.5*) respectively and placed into a 28°C water bath for 5 min. A PM2 only control was also set up to serve as a blank for the assay. Each reaction was started by the addition of 0.25 mL of ONPG (o-nitrophenol-β-galactoside) substrate solution pre-equilibrated to 30°C. After a sufficient yellow colour had developed the reaction was stopped by the addition of 0.5 mL 1 M Na₂CO₃ and the OD_{420nm} recorded.

The enzymatic activities, A (in units /ml), were calculated for each sample according to the following formula and analysed in the software GraphPad:

$$A = 200 * . (OD_{420} - OD_{420 \text{ in control tube}}) / \text{min of incubation} . \text{dilution factor}$$

(One unit of β-galactosidase activity corresponds to 1 nmol of ONPG hydrolysed per min at 28°C)

*200 in the above formula is the inverse of the absorption coefficient of o-nitrophenol that is 0.005 per nmol/ml at pH 11.0 (i.e. after addition of Na₂CO₃).

4.2.11 GEL FILTRATION COMPLEX STUDIES

The ability of the expressed proteins to form complexes was conducted using a gel filtration column, HiPrep 16/60 Sephacryl S-200 h at 4°C, on a AKTA Pure FPLC chromatography system in 2 x TBS (*Table 2.5*) and 20 % glycerol. 5mL of affinity purified Dlp2-6H and Dlp1a-6H were individually run down the gel filtration column to establish the elution volume for each of them at 0.5 mL/min over 1.5 column volumes. Then Dlp2-6H and Dlp1a-6H were mixed together in a 1:1 ratio for 15 min at 4°C. The mixed protein sample (5 mL) was loaded and the same column run was conducted. Protein complex interactions were analysed by UV_{280nm} readings.

4.3 RESULTS

To investigate the biochemical properties of the dynamin-like proteins, purified proteins were required. Both Dlp1a and Dlp2 were expressed and purified with the final goal being the use of the expressed protein in the raising of antibodies against both proteins for *in vivo* studies. Also, both proteins were to be used in the initial studies into the lipid binding capabilities of these dynamin-like proteins. Therefore, the highest possible purification was required to ensure as minimal contamination as possible in future studies. The third protein produced by the *dlp* operon, Dlp1b was not purified due to insolubility and time restrictions.

4.3.1 PURIFICATION OF DLP2-6H

A T7 inducible plasmid, pHis17 was used for the expression and over production of Dlp2-6H, with the open reading frame of Dlp2 cloned in such a way that the protein was tagged with a poly-histidine tag (6H) on the C-terminus (*Figure 4.1*). This was to ensure that only full-length protein was purified by Ni²⁺-affinity chromatography. The plasmid was then transformed into the arabinose inducible *E. coli* expression strain BL21-AI (Invitrogen). Transformed BL21-A1 cells were successfully grown in up-scaled cultures and a sample collected at the mid-log phase (OD_{600nm} ~ 0.3) pre-induction for a whole cell lysate comparison after protein induction. After the addition of 0.2% (w/v) arabinose for protein induction the cells grown for a further 16 h at 18°C. SDS-PAGE results revealed successful over-expression of Dlp2-6H, as evident by a protein of high abundance at ~ 65 kDa in the soluble whole cell lysate of transformants post-induction (*Figure 4.2*). The position at which the band appeared in the gel correlated with the predicted size for Dlp2-6H and was absent from the cell lysate pre-induction.

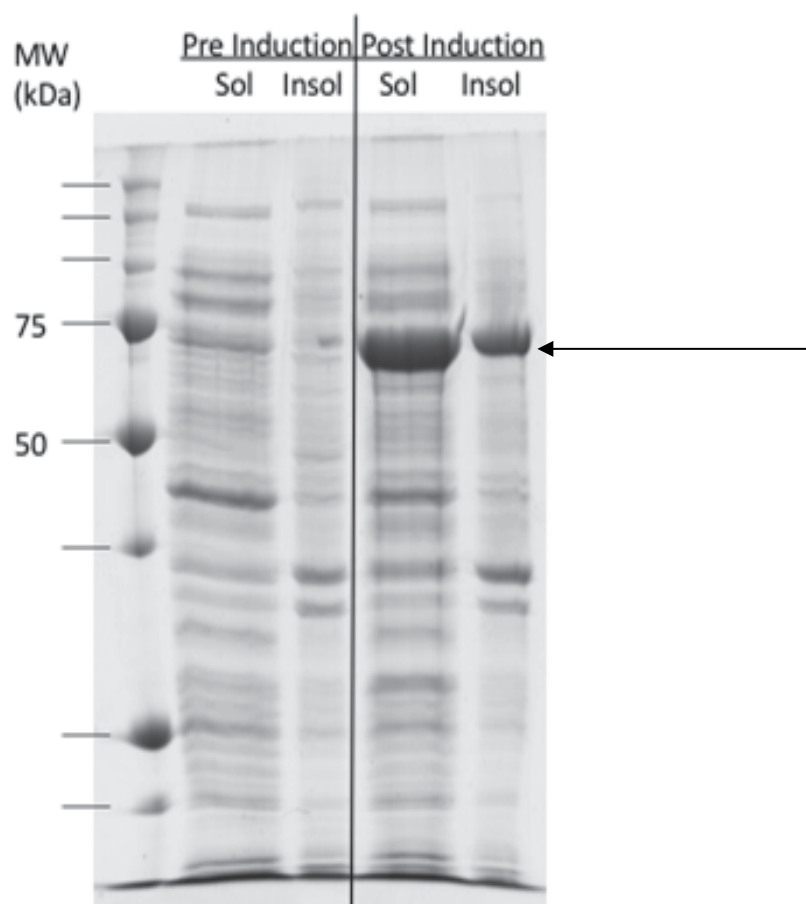


Figure 4.2 Dlp2-6H Over Expression

SDS-PAGE analysis of whole cell lysates from soluble (SOL) and insoluble (INSOL) fractions generated from representative transformed BL21-AI cultures. The gel demonstrated successful over-expression of soluble Dlp2-6H, as evident by a protein band of high abundance at ~ 65 kDa post-induction (arrow). This band was absent from the cell lysates pre-induction. Standards (Std); Protein Plus Standards (Bio-Rad).

The soluble fractions from replicate cultures were pooled together and loaded onto a 5 mL HisTrap HP column (GE Healthcare Systems). The flow through and the column washes were collected, to ensure that Dlp2-6H had bound to the column successfully before the elution steps. A gradient elution was performed into 1 M imidazole with fractions collected every 2 mL and analysed by SDS-PAGE. It can be seen in *Figure 4.3*, that Dlp2-6H was successfully bound to and purified from the His Trap column. There is evidence of some unbound Dlp2-6H in the flow through, but considering the very high amount of Dlp2-6H in the elution this was not considered important.

Due to the presence of impurities (e.g. *Figure 4.3* Lane E4-5), particularly at a lower molecular weight, a second column purification was performed using an ion-exchange column. The underlying principle of this type of purification is the binding of the molecule, in this case the protein of interest, to the column based on its charge. Dlp2-6H has a predicted isoelectric point of 5.65, with the hexahistidine tag taken into account, and therefore a strong anion exchange column with a slightly basic buffer (pH 8.0) was selected to promote binding of Dlp2-6H.

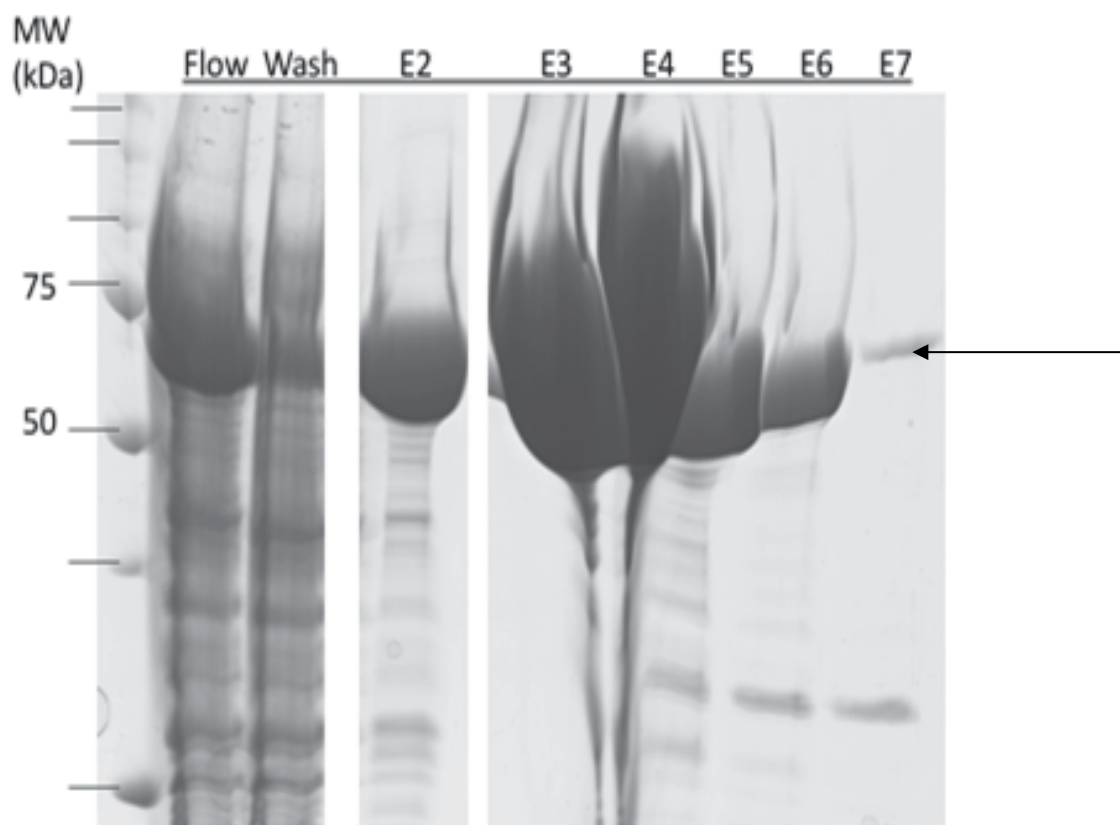


Figure 4.3 HisTrap Purification of Dlp-6H

SDS-PAGE analysis of the collected fractions from HisTrap purification; unbound flow through (Flow), column wash (Wash) and elution fractions (E2-E7) showed successful binding and purification of Dlp2-6H from the His Trap column with the presence of an overloaded protein band of ~65kDa (arrow), corresponding to the predicted MW of Dlp2. Standards (Std); Protein Plus Standards (Bio-Rad).

As it can be seen in *Figure 4.4*, the ion-exchange purification in the binding and elution of Dlp2-6H was successful but only provided a moderate additional level of purification to the protein, if any. This is evident in *Figure 4.4* by the presence of numerous impurities both above and below the band interest at the corresponding molecular weight of Dlp2-6H. This is most likely due to the column being vastly over loaded with evidence of Dlp2-6H in the flow through. These impurities seen could either be bound to the column directly or bound to the recombinant protein under these specific native conditions. Either way, the use of ion-exchange in future purification of Dlp2 was not performed due to the minimal improvement noted here.

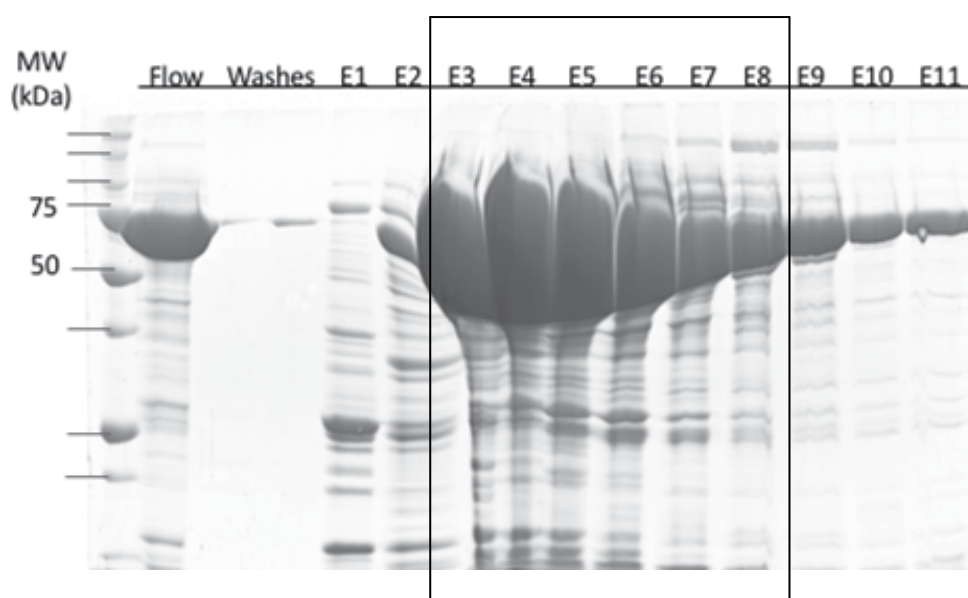


Figure 4.4 Ion Exchange Purification of Dlp2-6H

SDS-PAGE analysis of the collected fractions from ion exchange purification; unbound flow through (Flow), column wash (Washes) and elution fractions (E1-E11) showed successful binding and elution of Dlp2-6H from the Hi Trap HP column with the presence of an overloaded protein band of ~65 kDa, corresponding to the predicted MW of Dlp2-6H. E3-E8 fractions were pooled together for further purification (box). Standards (Std); Protein Plus Standards (Bio-Rad).

Due to the presence of impurities remaining in the Dlp2-6H sample, elution fractions E3-E8 (*Figure 4.4* boxed) were pooled together and run over an additional His Trap column, in the hope of improving the purification of Dlp2-6H. As seen in *Figure 4.5* this was successful with only a few additional bands below the molecular weight of Dlp2-6H in the overloaded SDS-PAGE gel. The purified Dlp2-6H fractions E3-E7 were pooled together and concentrated using a 10 kDa cut off Vivaspin concentrating column (GE Healthcare, Sweden). The final concentration of Dlp2-6H, calculated using the specific extinction coefficient 53 400 for Dlp2-6H, was found to be 12.84 mg/mL, with approximately 95% purity as seen in *Figure 4.6*. This preparation was considered a successful purification of Dlp2-6H and would be used in subsequent antibody production.

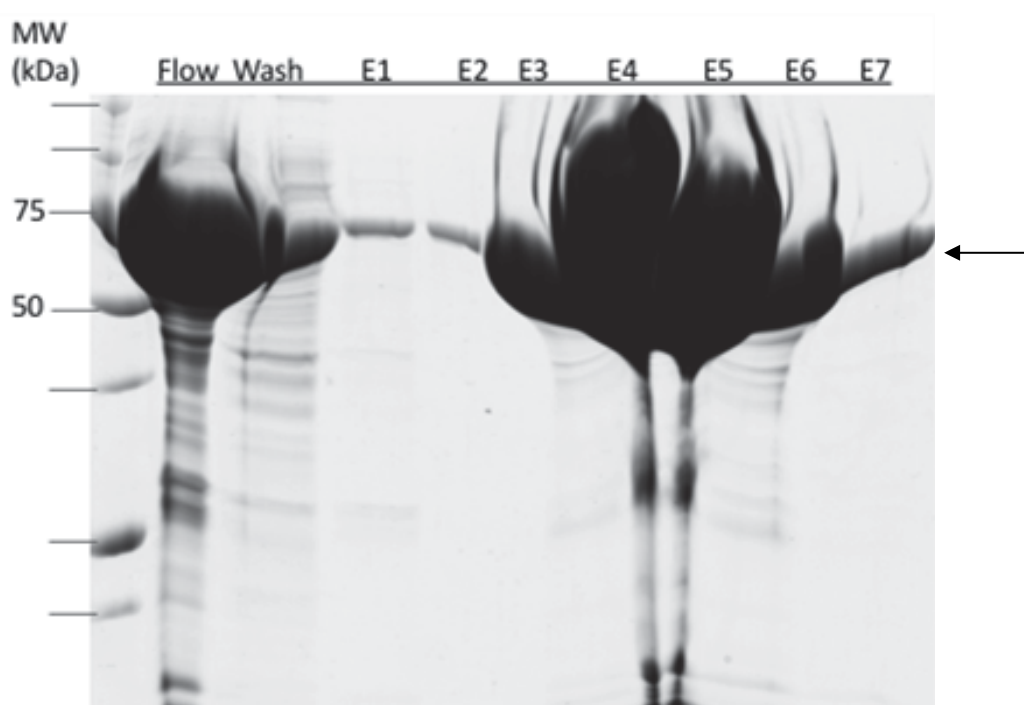


Figure 4.5 Further HisTrap Purification of Dlp2-6H

SDS-PAGE analysis of the collected fractions from HisTrap purification; unbound flow through (Flow), column wash (Wash) and elution fractions (E1-E7) showed successful binding and elution of Dlp2-6H from the His Trap column with the presence of an overloaded protein band of ~65 kDa, corresponding to the predicted MW of Dlp2-6H. Standards (Std); Protein Plus Standards (Bio-Rad).

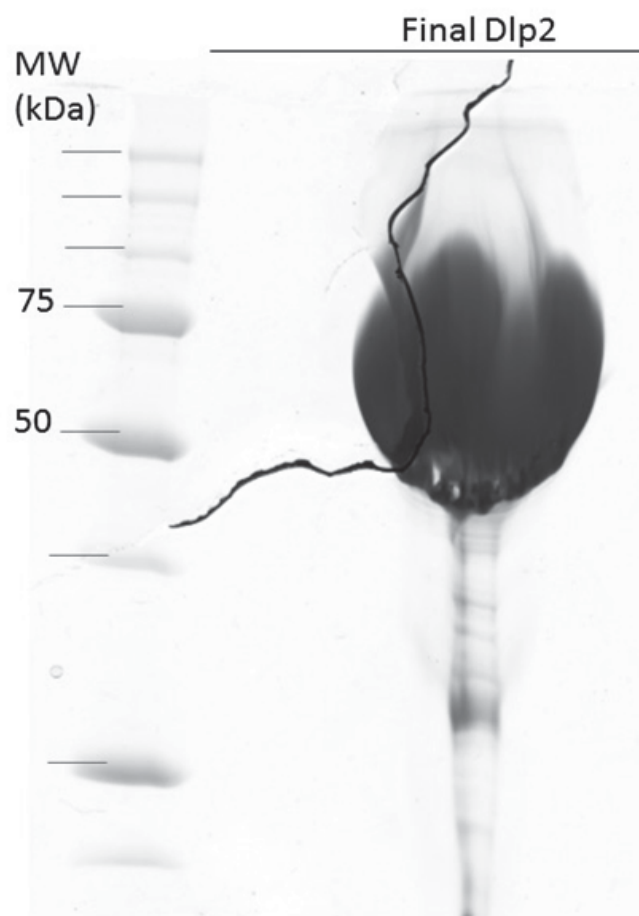


Figure 4.6 Final Purified Dlp2-6H

SDS-PAGE analysis of the final purified Dlp2-6H protein, even with overloading it can be seen that there are minimal contaminating proteins in this sample. Standards (Std); Protein Plus Standards (Bio-Rad).

Future use of Dlp2-6H in biochemical tests, such as GTPase and lipid binding assays required further purification and therefore a gel filtration column run was added to the protocol. This was in place of the ion-exchange purification step mentioned earlier which was concluded to not be helpful in the purification of this protein. The gel filtration purification as can be seen via SDS-PAGE in *Figure 4.7* was extremely successful in the removal of unwanted impurities in the protein sample making it as pure as possible for future uses.

Interestingly, in the gel filtration of Dlp2-6H evidence of potential dimerisation was seen (*Figure 4.7*). There were consistently two peaks in common between each replicate GF chromatogram for Dlp2-6H at 90-100mL and 160-200mL. Each line represents an individual run of Dlp2-6H on the GF column and the slight shift in the peak for one run (green line) can be attributed to a slightly different loading onto the column. The later peak, is the monomeric Dlp2-6H protein that was used for downstream assays with excellent purity (green box in insert *Figure 4.7*). The earlier peak in the run indicates protein of a higher molecular weight, but when these fractions were run on a reducing SDS-PAGE it was seen to also contain Dlp2-6H (red box in insert *Figure 4.7*).

With a reduction of approximately half in the elution time, this could potentially indicate a dimer complex for Dlp2-6H. This is the first indication of potential self-interaction of recombinant *H. pylori* Dlp2-6H, an important biochemical feature of all dynamin proteins. Future comparisons with molecular weight standards on a shorter gel filtration column and a non-reducing native PAGE gel will allow conclusive evidence of a protein-protein interaction for Dlp2.

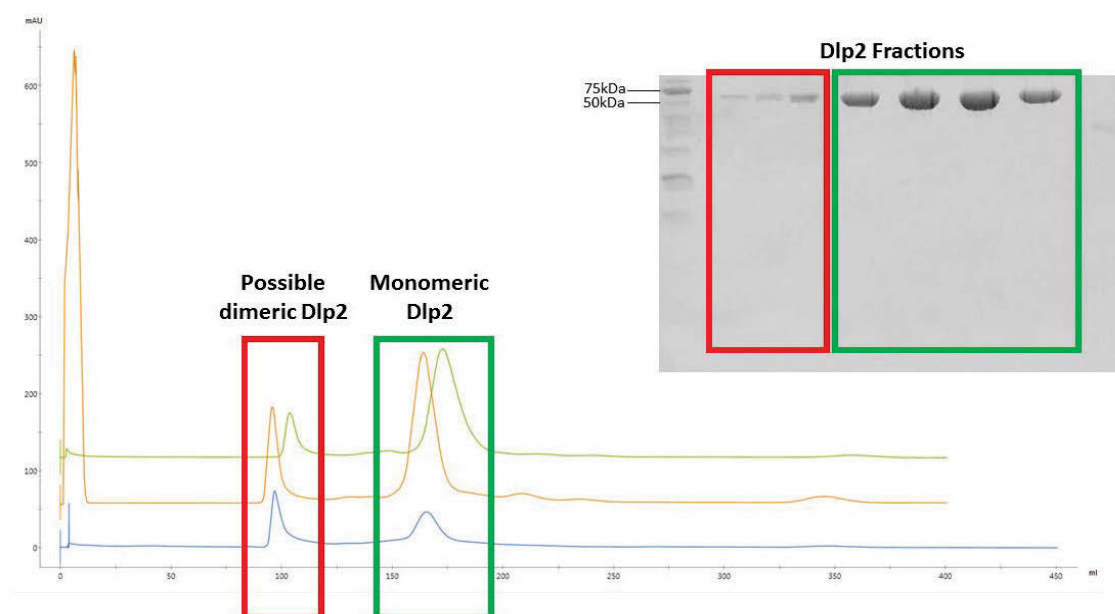


Figure 4.7 Possible Dlp2 Dimer Interactions

Gel filtration purification of the Dlp2-6H shows evidence of potential dimerization. The chromatograms of three replicate purifications highlight two peaks in common between each run. The earlier elution (around 100 mL) in the gel filtration specifies protein with a higher molecular weight and as the reducing SDS-PAGE gel shows (insert), these fractions contain Dlp2-6H (red box). This potentially indicates that Dlp2-6H is forming a higher order complex with itself during the run. The monomeric Dlp2-6H was repeatedly found to elute around 160mL (green box) and confirmed by SDS-PAGE.

4.3.2 PURIFICATION OF DLP1A-6H

The T7 expression system in the pHis17 plasmid, as previously described in *Section 4.3.1*, was once again utilized in the purification of another protein from the *Dlp* operon, Dlp1a. Dlp1a was cloned in the same manner as Dlp2, to ensure C-terminus tagging of the protein with a polyhistidine tail (-6H). Similarly, to Dlp2-6H transformed BL21-A1 cells with the pDlp1a-6H plasmid were successfully grown in up-scaled cultures and a sample was collected at the mid-log phase ($OD_{600nm} \sim 0.3$) pre-induction. After arabinose induction, the cells were allowed to continue to grow for a further 24 h. SDS-PAGE results revealed successful over-expression of Dlp1a-6H, as evident by a protein of high abundance at ~ 60 kDa in the insoluble post induction cell lysate (*Figure 4.8*).

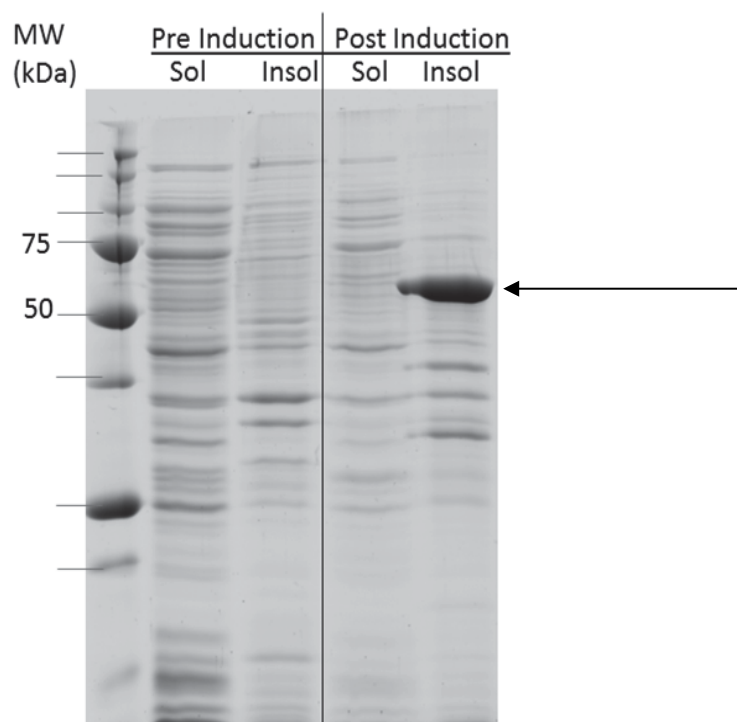


Figure 4.8 Dlp1a-6H Over Expression

SDS-PAGE analysis from whole cell lysates of soluble (SOL) and insoluble (INSOL) fractions generated from representative transformed BL21-AI cultures. The gel demonstrated successful over-expression of insoluble Dlp1a-6H, as evident by a protein band of high abundance at ~ 60 kDa post-induction (arrow). This band was absent from the cell lysates pre-induction. Standards (Std); Protein Plus Standards (Bio-Rad).

Due to the complete insolubility of Dlp1a-6H, as seen in *Figure 4.8* and the presence of clear inclusion bodies in the *E. coli* cells (data not shown), an inclusion body purification protocol was implemented (*Section 4.2.3.1*). Recombinant protein can accumulate intracellularly during over expression, become aggregated and insoluble; forming what is known as inclusion bodies. Inclusion bodies are made of misfolded, insoluble protein lacking any biological activity (205). The high buoyant density of inclusion bodies facilitates their easy separation from soluble *E. coli* proteins and cell debris by differential centrifugation in urea (205, 206). This protocol was highly successful and provided a significant initial purification of Dlp1a-6H with what seemed to be mostly degradation products in the sample (*Figure 4.9*).

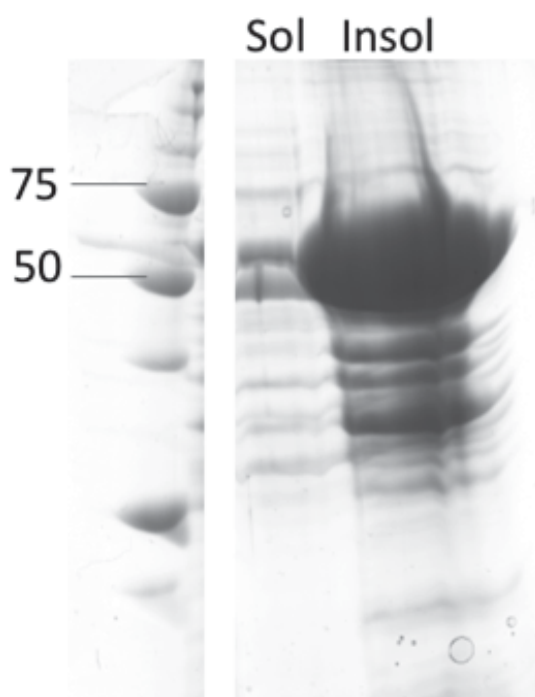


Figure 4.9 Final Inclusion Body Containing Dlp1a-6H

SDS-PAGE analysis of the final protein content in the insoluble (Insol) and soluble (Sol) fractions of the purified inclusion bodies. The gel clearly demonstrates the presence of the insoluble Dlp1a-6H with a band present at ~60 kDa. Standards (Std); Protein Plus Standards (Bio-Rad).

As mentioned, inclusion bodies are essentially aggregates of insoluble proteins, making the purification of the natively folded protein impossible. Therefore, the inclusion bodies containing Dlp1a-6H were solubilised in guanidine hydrochloride and the now denatured Dlp1a-6H protein was captured, purified and refolded using a His Trap column. Dlp1a-6H was slowly refolded on the column by applying a gradient of decreasing urea concentration aimed at returning the protein to its native conformation. The refolded protein was then washed and eluted.

This purification procedure was successful as seen in *Figure 4.10* with Dlp1a-6H observed in the elution fractions 2 - 7. These were pooled together for dialysis and concentration. It was also noted that there were a few products of lower molecular weight in the sample and these could be attributed to either potential degradation products of the protein throughout the purification process or consistent contaminants from the *E. coli* whole cell lysate (arrows).

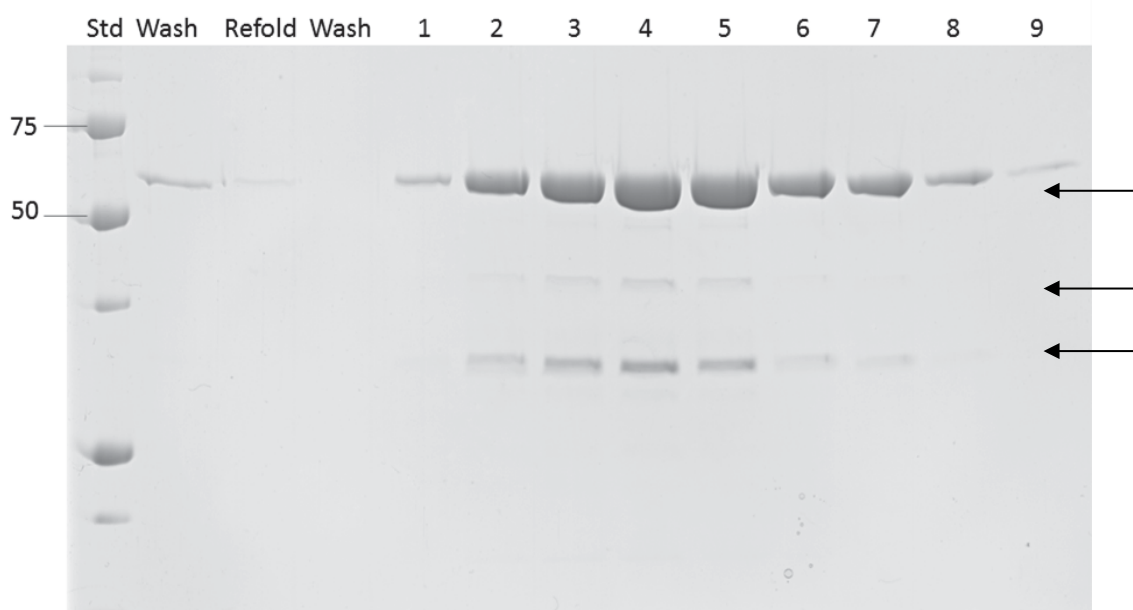


Figure 4.10 Refolding and Elution of Dlp1a-6H

SDS-PAGE analysis was performed on the collected fractions from the refolding and purification of Dlp1a-6H (top arrow). The gel demonstrated a successful purification of Dlp1a-6H from the inclusion bodies. Minimal protein was lost in the washing and refolding steps with a gradient of protein eluting in Fractions 2-7. Degradation and/or contaminating bands are indicated by middle and bottom arrow. Standards (Std); Protein Plus Standards (Bio-Rad).

Despite the apparent success of the refolding and elution procedure, Dlp1a-6H was found to readily precipitate in solution. This was particularly evident during dialysis, which was needed with the removal of imidazole required for downstream applications. Accordingly, SDS-PAGE with a BSA standard curve was used to estimate the concentration and yield of Dlp1a-6H generated from this protocol (*Figure 4.11*). It was estimated that approximately 50% of Dlp1a-6H had precipitated during dialysis, with the sample spun down to separate it into the soluble and insoluble protein. This seemingly equated to approximately 2mg/mL in both the soluble and insoluble fraction of the protein sample. Despite the partial insolubility of the Dlp1a-6H protein, the sample was retained for use for antibody production trials which should not be inhibited by partially-insoluble protein.

Further investigations into a more effective purification protocol for Dlp1a-6H considering its insolubility and unstable nature proved fruitful with the addition of 20% glycerol needed to ensure stability of the refolded protein during and after elution from the column (data not shown). This established addition to the purification protocol for Dlp1a-6H was then implemented in all subsequent preparations for biochemical assays and localisation studies.

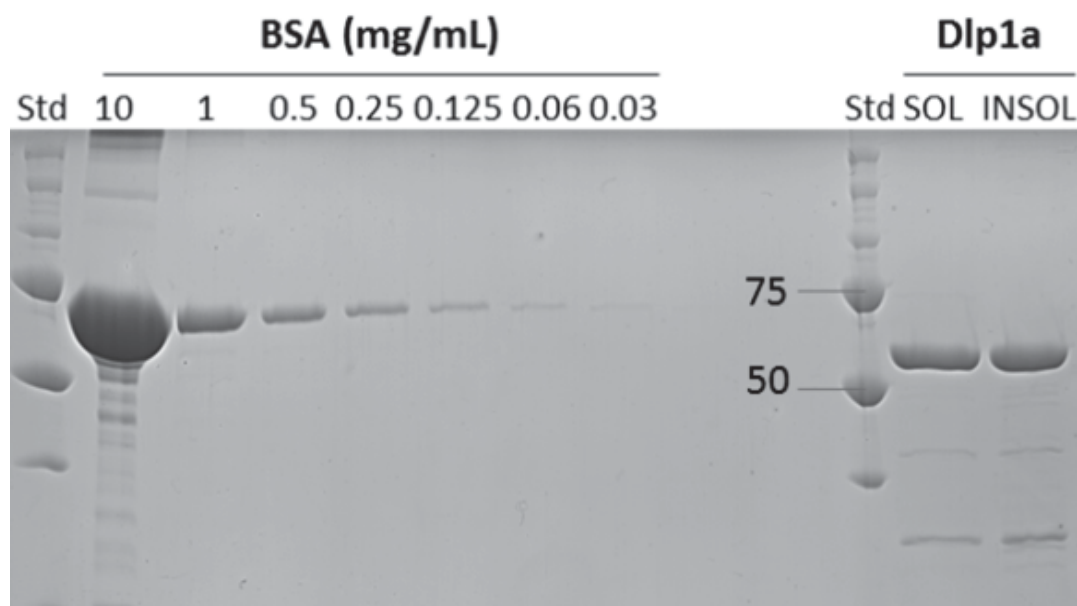


Figure 4.11 BSA Standard Curve for Estimation of Dlp1a-6H Concentration

A BSA standard curve from 10 mg/mL to 0.03 mg/mL was performed to estimate the concentration of Dlp1a-6H that was soluble and insoluble post purification and dialysis. It was estimated that the amount of Dlp1a-6H in the soluble (SOL) and insoluble (INSOL) fractions was 50:50 with approximately 2 mg/mL of Dlp1a-6H in each. Standards (Std); Protein Plus Standards (Bio-Rad).

4.3.3 IN VIVO EXPRESSION OF *DLP1A* AND *DLP2*

Polyclonal antibodies for both Dlp1a-6H and Dlp2-6H were produced by SAHMRI as described in *Section 4.2.4*. These were then tested for specificity, using Western blotting with WT *H. pylori* 26695 whole cell lysates separated by SDS-PAGE. Described later in *Chapter V*, a deletion strain was successfully constructed in *H. pylori* B128.7.13, a gerbil adapted strain, and was utilised in antibody testing and subsequent use of the antibodies in IFM experiments, despite slight differences in the amino acid sequence of the *dlp* operon (*described in section 5.3.3*).

It was clearly seen that the antiserum for Dlp2-6H (α -Dlp2) was specific, with a single major response band consistent with the expected size for Dlp2 in the *H. pylori* 26695 whole cell lysate, as seen in *Figure 4.12*. There was also a sharp band corresponding to the purified protein, Dlp2-6H, used as a size control in this experiment. It can also be seen that a clear band in the B128.7.13 WT whole cell lysate that corresponds to the correct size for Dlp2 is present and that there is a clear reduction in the signal seen at the same size in the whole cell lysate for the B128.7.13 Δ *dlp* clones. The signal that is still seen could be attributed to original hidden background signal for B128.7.13, as specificity is not as high when compared to 26695, with known slight differences between Dlp2 in the two strains.

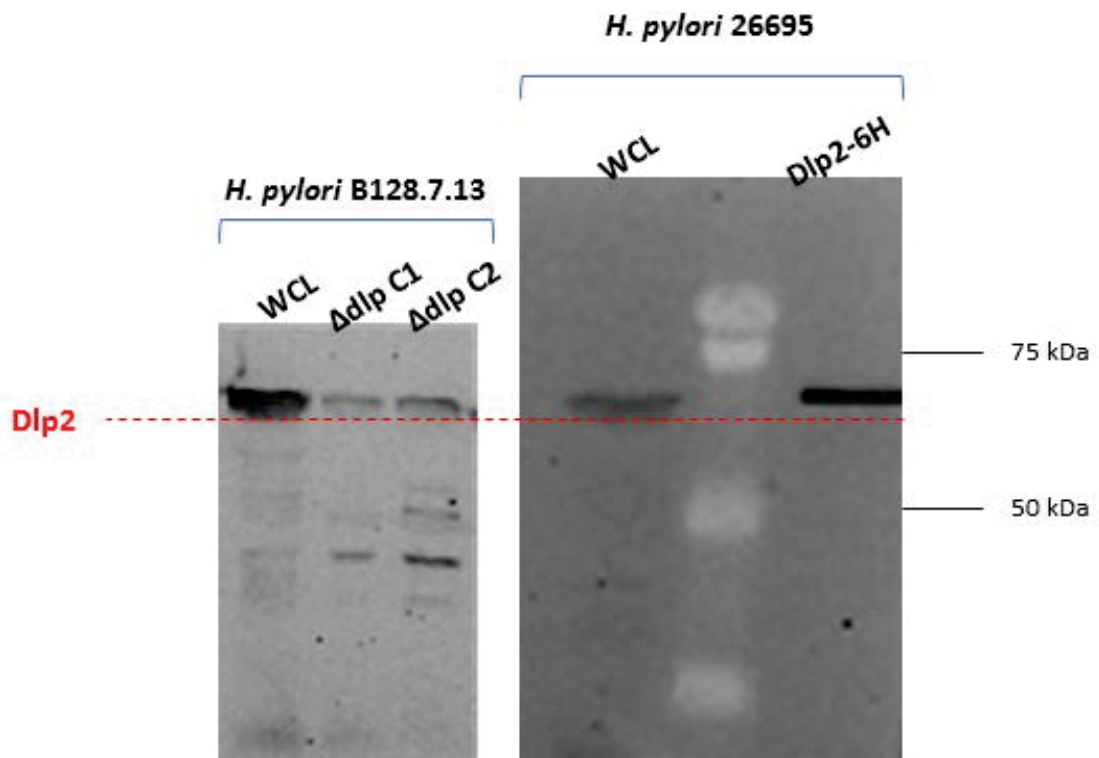


Figure 4.12 In vivo Expression of Dlp2 in both 26695 and B128.7.13

Specificity of the antiserum for Dlp2-6H was tested using western blotting. It was tested on *H. pylori* whole cell lysate (WCL) and the control of purified antigen (Dlp2-6H). A deletion strain of the *dlp* operon was successfully created in *H. pylori* B128.7.13 and therefore the antiserum was also tested on WCL of the Wild Type (WT) strain and individual clones of the deletion (Δ dlp C1 and Δ dlp C2). A clear band at approximately 66kDa (red line) is evident in the WCL of both 26695 and B128.7.13 that is comparable to the size control. There is clear reduction in signal seen in Δ dlp C1 and Δ dlp C2, with potentially only background signal being detected.

When antiserum for Dlp1a-6H (α -Dlp1a) was compared to α -Dlp2 it was clear that it was not as specific and most likely could be attributed to the immunogenicity of the antigen used to immunise the rabbits originally (1/4 of what was used for Dlp2-6H). In the testing of this antibody, two *H. pylori* 26695 extracts were used; a whole cell lysate prepared from the 'mixed' culture and a whole cell lysate prepared from a 'single' isolate of the strain, both previously described in *Chapter III*. It was hypothesised that the mixed extract, if phase variation was occurring in the strain, would show evidence of both the full length Dlp1 as well as the truncated Dlp1a *in vivo*, whereas the single isolate lysate would show a single band for only Dlp1 *in vivo*. With no available recombinant full length Dlp1 at the time of this study, an *E. coli* overexpression whole cell lysate was used as the size control, along with the purified Dlp1a-6H for comparison between the two lysates. Along with these, B128.7.13 WT and B128.7.13 Δ dlp C1 and C2 whole cell lysates were once again used to test specificity.

As seen in *Figure 4.13*, in the *H. pylori* mixed isolate whole cell lysate there are several bands detected using α -Dlp1a, two of which stand out when compared to the size controls. At ~60 kDa there is a clear band that corresponds nicely to the size of the recombinant Dlp1a (*Figure 4.13* green line). Then at ~75 kDa there is another band that agrees with the full length Dlp1 when compared to the over expressed size control (*Figure 4.13* blue line). This is the first indication of potentially both the full-length dynamin (Dlp1) and the truncated dynamin (Dlp1a) being expressed *in vivo*, consistent with the hypothesised phase variation expression of the *dlp* genes previously described in Chapter III. This is further supported in B128.7.13 where evidence of both proteins is once again present and then as expected both bands disappear in both the B128.7.13 Δ dlp clone whole cell lysates. This really highlights that this antibody is binding to expressed Dlp1 and Dlp1a *in vivo*, and it strengthens the hypothesis that both forms of the protein exist in this strain of *H. pylori*. However, the detection of a proteolytic breakdown product of similar size cannot be ruled out. Remarkably in the single isolate whole cell lysate, there is detection of only the full-length dynamin, Dlp1 (*Figure 4.13*). The disappearance of the band at ~60 kDa but the retention of the degradation or nonspecific binding products indicates that this lysate may only contain the product of expression of *dlp1*. Again, this further supports the hypothesis of phase variation to control the expression of the *dlp* genes, specifically *dlp1*, *dlp1a* and *dlp1b*. This is first report of *in vivo* detection of the Dlp1s in *H. pylori* whole cell extracts with Dlp2, full length Dlp1 and a truncated Dlp1a all being detected.

This is also the first *in vivo* evidence of potential phase variants of the *dlp* operon in *H. pylori*

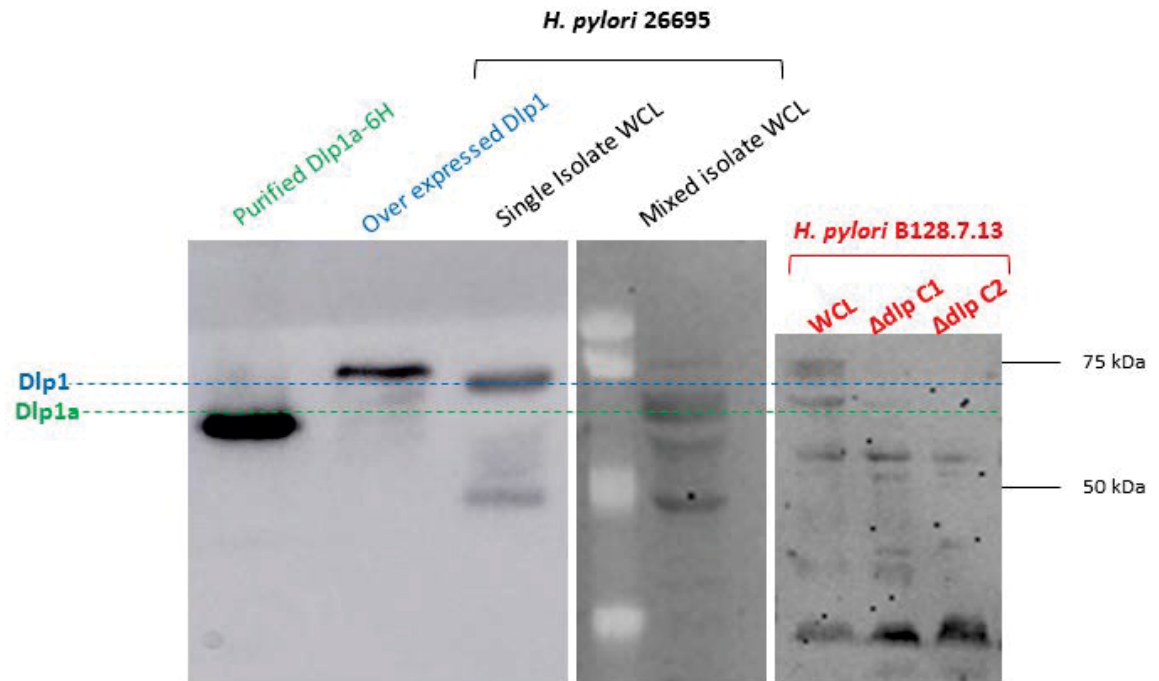


Figure 4.13 In vivo Expression of Dlp1a and Dlp1 in both 26695 and B128.7.13

Specificity of the antiserum for Dlp1a-6H was tested using western blotting. It was tested on two *H. pylori* 26695 whole cell lysates (WCL); mixed and a single isolate (previously described in Chapter III) as well as the control of purified antigens, Dlp1a-6H and *E. coli* WCL of overexpressed Dlp1-6H. A deletion strain of the *dlp* operon was successfully created in *H. pylori* B128.7.13 and therefore the antiserum was also tested on WCL of the Wild Type (WT) strain and individual clones of the deletion (Δ dlp C1 and Δ dlp C2). A clear band at approximately 60 kDa (green line) is evident in the WCL of the mixed WCL of 26695 and the B128.7.13 WCL that is comparable to the size control. A clear band at approximately 75 kDa is evident in the WCL of the both the mixed and single isolate WCL of 26695 as well as B128.7.13 WCL, that is comparable to the size control. There is a clear absence of signal seen in Δ dlp C1 and Δ dlp C2 for both forms of the protein, with the retention of the nonspecific binding signals.

4.3.4 SUBCELLULAR LOCALISATION OF DYNAMIN-LIKE PROTEINS IN *H. PYLORI*

To begin the assessment of the sub-cellular localisation of Dlp1a and Dlp2 in *H. pylori*, subcellular fractionation and Immuno-Fluorescence Microscopy (IFM) were utilised. These were performed by my colleague Nina Sulkowski (Honours student, UTS) with *H. pylori* 26695 cells. Subcellular biochemical fractionation was carried out on whole *H. pylori* cells, separating them into their soluble and membrane fractions using sonication and high-speed centrifugation. This was then followed by western blotting to detect the presence of both Dlp1 and Dlp2 proteins in their respective normalised fractions. Each lysate for the fractions was normalised using BSA standard curve calculations as well as short run total protein analysis (outlined in *Section 4.2.6*), with the Ponceau S staining indicating a similar amount of total protein in each fraction loaded onto the gel for analysis (*data not shown*).

Previously in *Chapter III*, it has been shown that Dlp1 has a predicted transmembrane domain and preliminary results from the fractionation suggest that Dlp1 localises in the membrane fraction (*Figure 4.14A*). This is indicated by a band of corresponding size to the full length Dlp1 found in the membrane fraction, with a much lesser amount of protein detected in the soluble fraction. Comparatively, Dlp2 localisation in the same subcellular fractionation study appeared to be split between the soluble and the membrane fraction (*Figure 4.14B*). The majority of the Dlp2 protein seemingly localised to the soluble fraction, which made sense with no membrane target domain present in the protein unlike Dlp1. This split in localisation of Dlp2 could be indicative of a potential interaction at the membrane with Dlp1, with only Dlp1 having the predicted membrane-binding domain.

Future fractionation studies would require a panel of compartment controls to ensure fractionation specificity to allow more conclusions to be made.

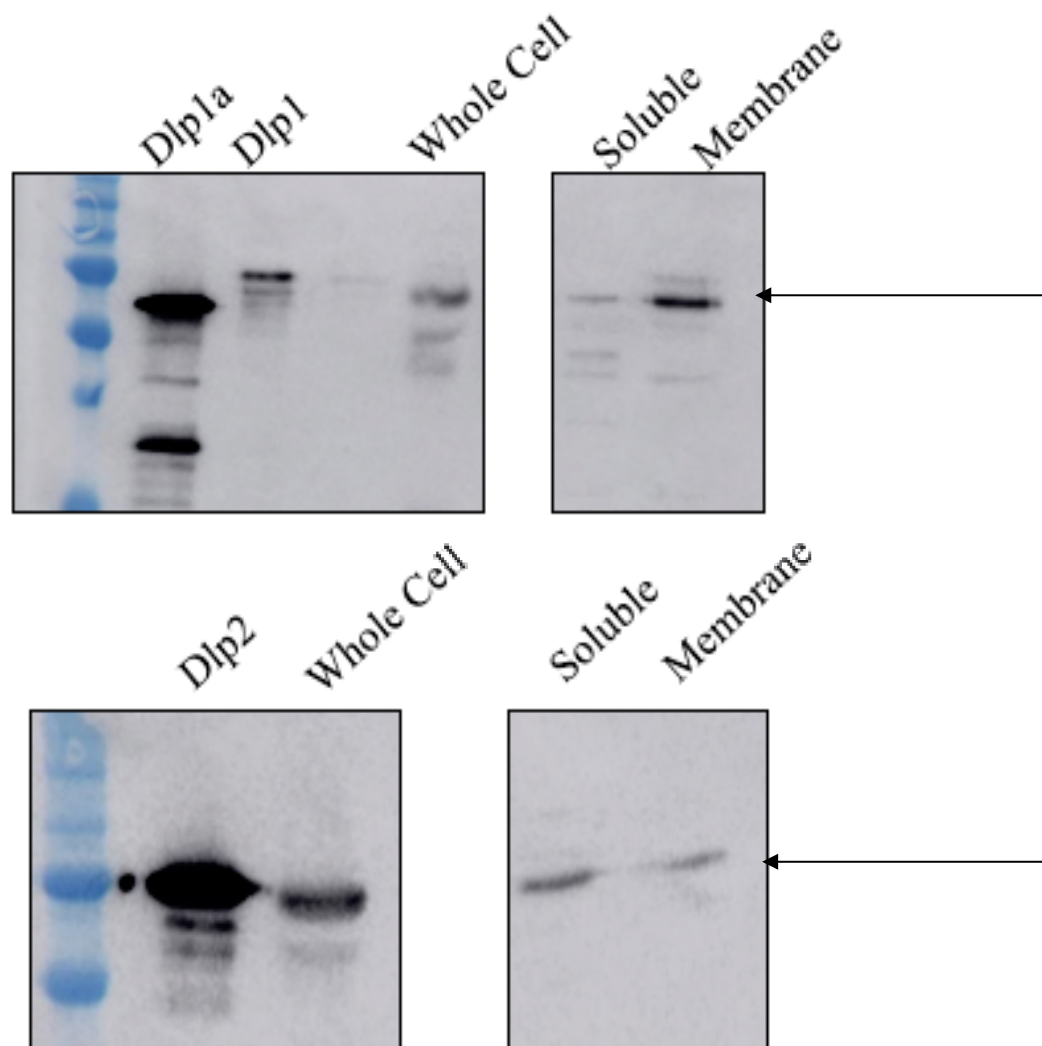


Figure 4.14 *H. pylori* Whole Cell Fractionation and Localisation of Dlp1 and Dlp2

Western blot analysis was performed on the 26695 preliminary fractionation samples both soluble and membrane fractions (which consists of all insoluble proteins) (right). Antigens and whole cell lysates were used as controls (left). (A) The Dlp1a blot showed bands corresponding to Dlp1a and Dlp1 in the controls as well as the whole cell lysate. Quite repeatedly it was shown to be in the “membrane” fraction with only faint detection in the soluble fraction (arrow). (B) The Dlp2 blot showed bands corresponding to Dlp2 in the antigen control as well as the whole cell lysate. Quite repeatedly it was shown to be split between the “membrane” fraction and the soluble fraction (arrow).

Following on from the fractionation studies, Immuno-fluorescence microscopy (IFM) was used to further investigate the localisation of both Dlp1 and Dlp2 in *H. pylori* 26695 cells. IFM requires the cells to be fixed and treated to allow for the antibodies to pass through the cellular membrane. A secondary antibody is introduced to the cells to bind to the primary antibody and then fluoresce to highlight the location of the protein of interest. Therefore, both α -Dlp2 and α -Dlp1a were further purified by affinity chromatography to ensure specificity to the protein of interest (data not shown).

Preliminary results for Dlp2 showed successful binding of the α -Dlp2 antibody to fixed *H. pylori* cells seen at an optimal concentration of 5 μ g antibody and 10 μ g lysozyme treatment. There seemed to be distinct foci present in a majority of the cells, which appears not only at the membrane but also dispersed throughout the length of the cells (*Figure 4.15D*). In regard to Dlp1a, despite showing successful binding to fixed *H. pylori* cells with an optimal concentration of 10 μ g antibody and 10 μ g lysozyme treatment, no distinct membrane localisation was observed as would be predicted (*Figure 4.15H*). However, upon closer examination of the cells, patchy dispersion of Dlp1a throughout the length of the cell can be seen with varying intensities between cells. This localisation pattern is quite reminiscent of other bacterial dynamins and could potentially indicate varying levels of expression for Dlp1 in the cell.

This fluorescence seen for both α -Dlp2 and α -Dlp1a is not due to non-specific binding of the secondary antibody alone, with no discernible levels of fluorescence detected in the absence of a primary antibody (*Figure 4.15L*). However, these preliminary localisation results cannot be validated without a deletion strain in 26695 of the *dlp* operon and it is not clear whether the Dlp/FITC signal seen was specific to the localisation of the *in vivo* expressed Dlps due to this.

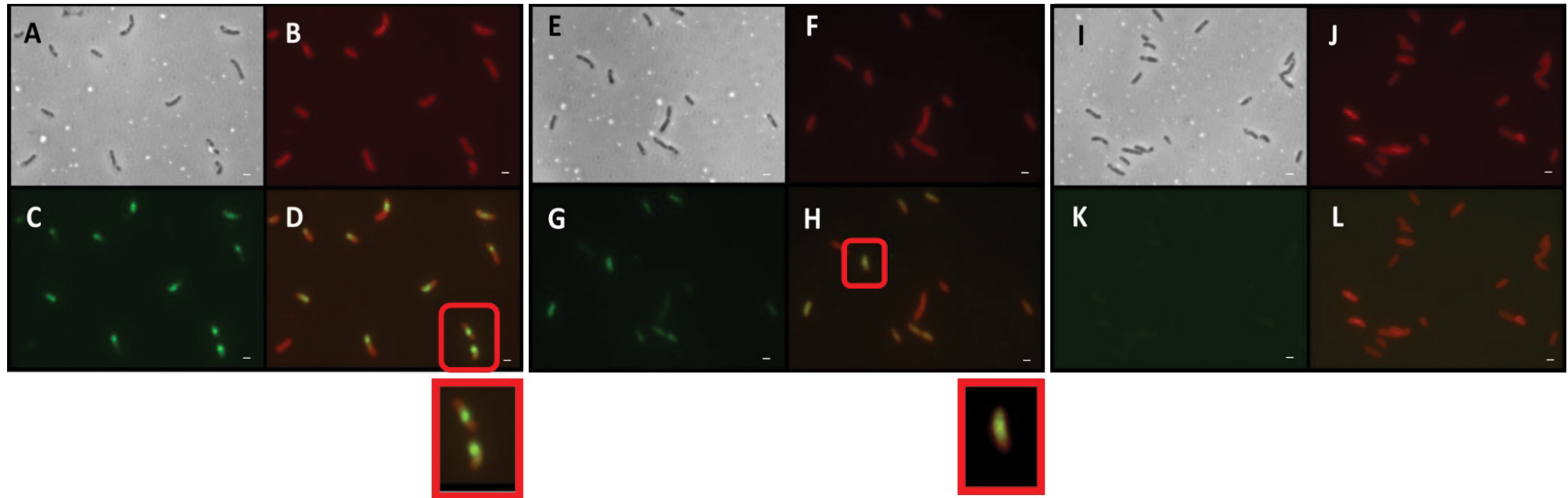


Figure 4.15 IFM of *H. pylori* 26695 for Localisation of both *Dlp1a* and *Dlp2*

IFM was carried out using an optimal lysozyme concentration of 10 μg , as determined by this study. Images a-d demonstrate staining with 5 μg purified anti-*Dlp2* polyclonal antibody. Images e-h demonstrate staining with 10 μg purified anti-*Dlp1a* polyclonal antibody. Images i-l (negative control) demonstrate that the secondary antibody does not bind non-specifically to cells. Phase microscopy was performed (images a, e, i) using the auto-exposure setting at 100x magnification. Imaging under FITC filter (c, g, k) was performed at 500 ms fixed exposure. FM4-64 staining was assessed using a rhodamine filter (b, f, j) at 40 ms fixed exposure. The FITC/FM4-64 composites (d, h, l) were constructed using Zeiss AxioVision software. Contrast settings were normalised manually between images using AxioVision software. Scale bar is 2 μm .

As previously noted, and described in *Chapter V*, a deletion strain was successfully constructed in *H. pylori* B128.7.13, a gerbil adapted strain. This was then utilised in preliminary IFM localisation experiments, despite slight differences in the amino acid sequence of the *dlp* operon already described (*Section 5.3.3*).

The specificity of the α -Dlp2, despite differences in the target protein sequence, is clearly seen in *Figure 4.16* where both B128.7.13 WT and Δdlp cells were treated with 1 μ g lysozyme and 1 μ g of α -Dlp2. In the treated Δdlp cells, a very slight FITC signal can be seen, indicating minimal nonspecific binding of the antibody to the cells in the absence of their target Dlp2 (*Figure 4.16E-H*). WT cells show a bright FITC signal, indicating binding of the antibody to its target Dlp2 in WT B128.7.13 (*Figure 4.16A-D*). The specificity seen here with the Δdlp strain allows conclusive remarks to be made about the localisation of Dlp2 in *H. pylori*. Upon closer examination of the localisation pattern for Dlp2, it is seemingly dispersed throughout the length of most cells imaged with no obvious pattern for localisation within the *H. pylori* cell (*Figure 4.16D insert*).

The α -Dlp1a specificity was not as clear as α -Dlp2, as seen in *Figure 4.16*. Both the WT and the Δdlp strains were treated with 1 μ g lysozyme but due to technical issues with fragility of the membrane in the Δdlp cells, WT cells treated with 1 μ g of α -Dlp1a and Δdlp cells treated with 50 μ g of α -Dlp1a. Despite the issues with the fragility of the membrane/cells in the Δdlp samples, a meaningful comparison was still possible between the two strains with a much higher amount of α -Dlp1a in the Δdlp cells acknowledging any fluorescence seen is at 50 times higher level of antibody present compared to what was added to WT. Keeping that in mind, the Δdlp cells showed a slight FITC signal, indicating potentially some level of nonspecific binding of the antibody to the cells in the absence of their target Dlp1 (*Figure 4.17E-H*). However, this low intensity nonspecific binding is observed at a 50 times higher level of antibody presence in the sample, therefore a difference could still clearly be noted between it and the WT FITC signal (*Figure 4.17A-D*). WT cells show a bright FITC signal, indicating that there was binding of the antibody to its target Dlp1. By comparing the overlay of the WT (*Figure 4.17D*) with the overlay for Δdlp cells (*Figure 4.17H*), a much clearer localisation in the WT cells is evident, with distinct foci present when compared to the haze of background signal seen in the Δdlp cells

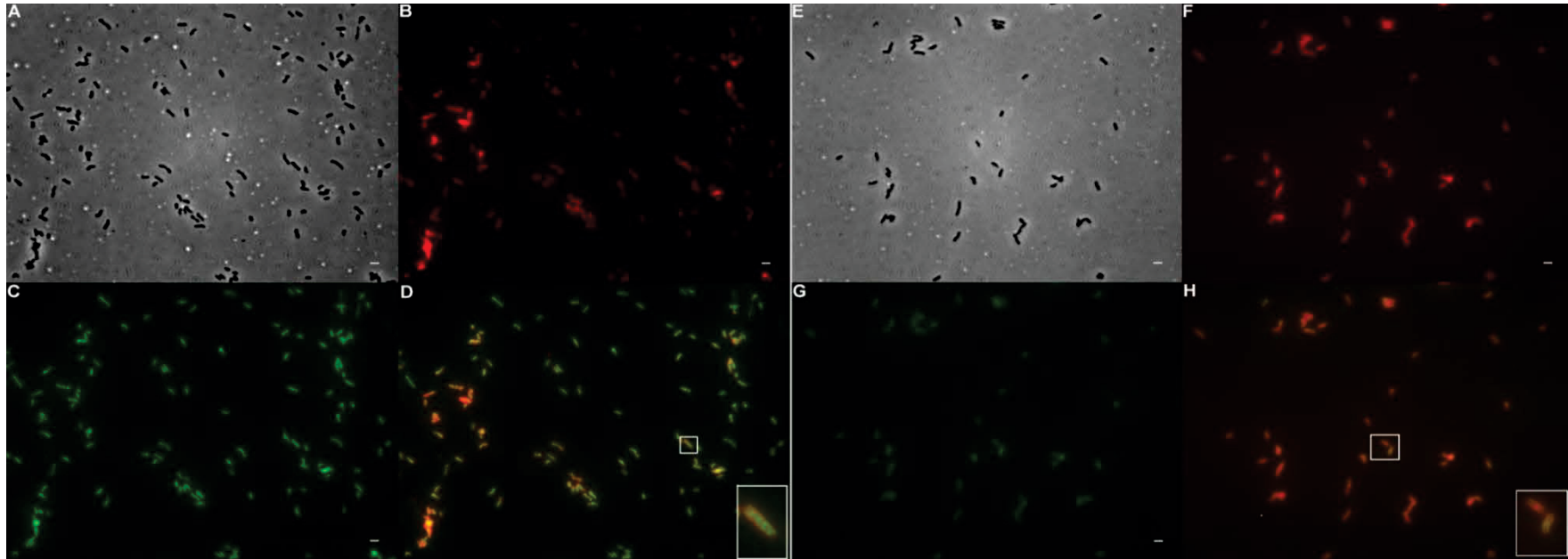


Figure 4.16 IFM of *H. pylori* B128.713 for in vivo Localisation of Dlp2

IFM was carried out using an optimal lysozyme concentration of 1 μ g, as determined by this study. Images A-D demonstrate staining with 1 μ g purified anti-Dlp2 polyclonal antibody on B128.7.13 WT cells. Images E-H demonstrate staining with 1 μ g purified anti-Dlp2 polyclonal antibody on B128.7.13 Δ dlp cells. Phase microscopy was performed (images A and E) using the auto-exposure setting at 100x magnification. Imaging under FITC filter (C and G) was performed at 500 ms fixed exposure. FM4-64 staining was assessed using a rhodamine filter (B and F) at 40 ms fixed exposure. The FITC/FM4-64 composites (D and H) were constructed using Zeiss AxioVision software. Contrast settings were normalised manually between images using AxioVision software. Scale bar is 2 μ m.

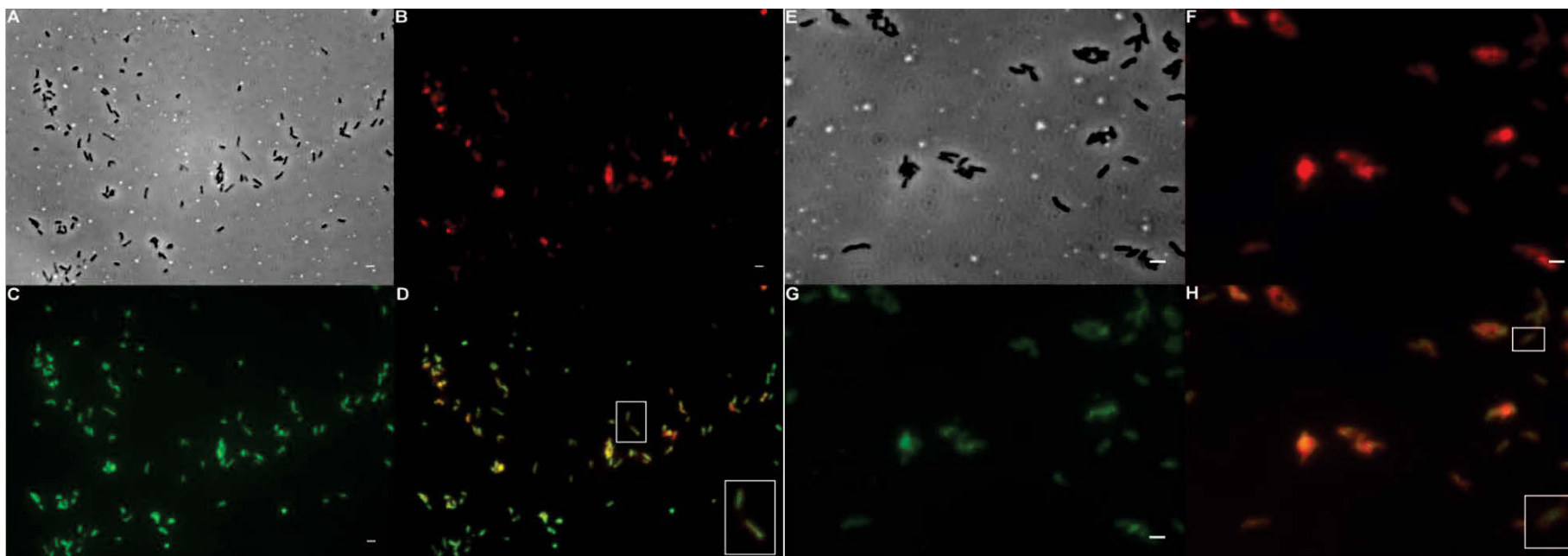


Figure 4.17 IFM of *H. pylori* B128.713 for in vivo Localisation of Dlp1a

IFM was carried out using an optimal lysozyme concentration of 1 μ g, as determined by this study. Images A-D demonstrates staining with 1 μ g purified anti-Dlp1a polyclonal antibody on B128.7.13 WT cells. Images E-H demonstrate staining with 50 μ g purified anti-Dlp2 polyclonal antibody on B128.7.13 Δ dlp cells. Phase microscopy was performed (images A and E) using the auto-exposure setting at 100x magnification. Imaging under FITC filter (C and G) was performed at 500 ms fixed exposure. FM4-64 staining was assessed using a rhodamine filter (B and F) at 40 ms fixed exposure. The FITC/FM4-64 composites (D and H) were constructed using Zeiss AxioVision software. Contrast settings were normalised manually between images using AxioVision software. Scale bar is 2 μ m.

4.3.5 LIPID INTERACTION STUDIES OF *H. PYLORI* DLPS

A key feature of the dynamin superfamily is their ability to interact with and manipulate the shape of lipid membranes. To mimic this *in vitro*, liposomes were formed from *E. coli* total lipid extracts (Avanti) and mixed with 10 μ M of Dlp2-6H and 1 mM GTP/Mg²⁺. Using the membrane dye, FM4-64, changes to the liposomes upon mixing with Dlp2-6H were observed as seen in *Figure 4.18*. The liposomes only control showed small particles of varying sizes, with a diameter of primarily less than 1 μ m. When the liposomes were mixed with a lipid binding protein, like BSA, which binds to fatty acids, there was no observable change to their appearance (*Figure 4.18B*). This indicates that any changes seen upon mixing the liposomes with Dlp2-6H are not non-specific changes due to the presence of a protein in general. As seen in *Figure 4.18D-I* substantial amounts of lipid aggregation is evident when the liposomes were mixed with Dlp2-6H and GTP.

The deformities were difficult to analyse quantitatively due to the high level of signal background in the liposome samples but qualitatively, the Dlp2-6H treated liposomes were observed to be much larger and highly irregular in shape (inserts *Figure 4.18D-F*), along with some very large apparent aggregates (*Figure 4.18 G-H*). Along with these individual liposome changes and clear aggregation of liposomes, it should be noted that not the entire population of liposomes were this way, with clear areas of population with minimal change to the liposomes as seen in *Figure 4.18I*. This aggregation is the first evidence of potential membrane fusion occurring in the presence of Dlp2-6H and its nucleotide GTP, although the exact cause of the apparent aggregation awaits further study.

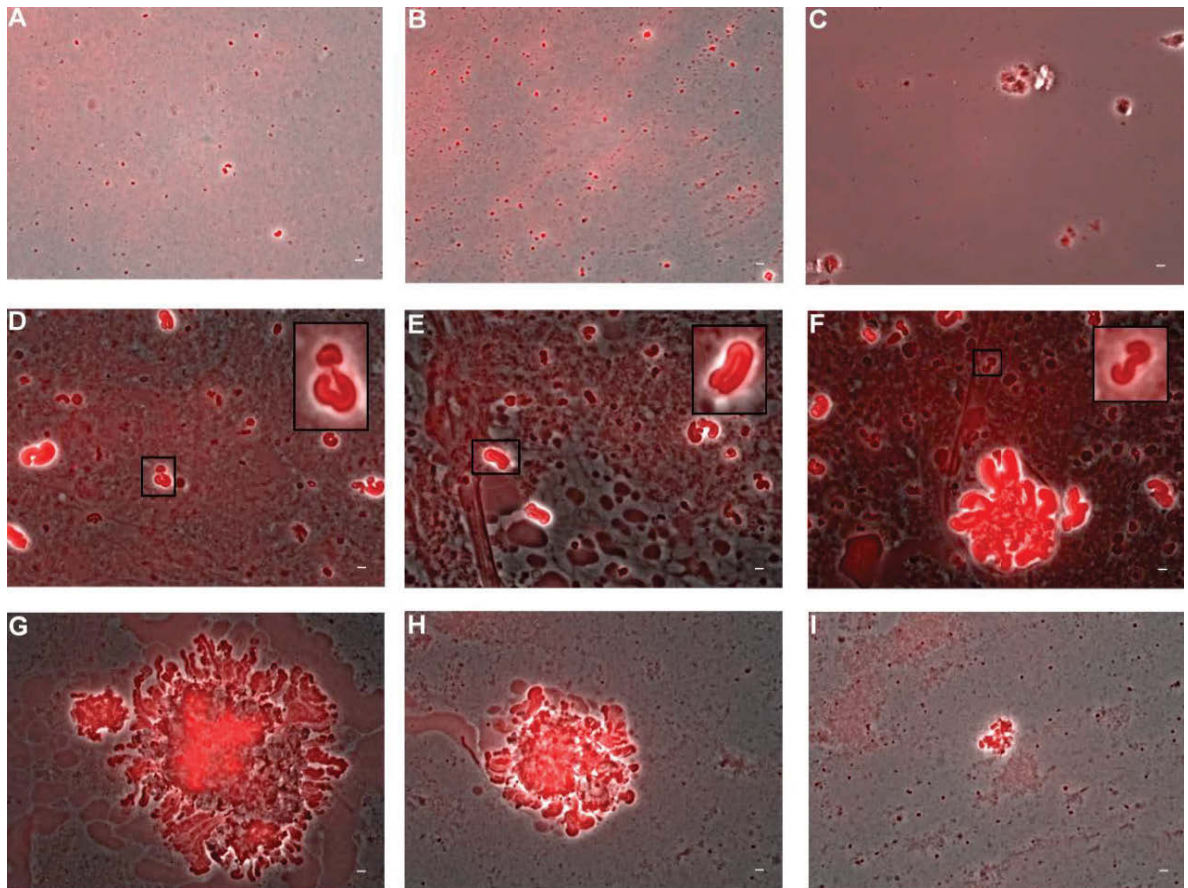


Figure 4.18 Liposome Deformations in the Presence of Dlp2-6H

Liposome and Dlp2-6H interactions were assessed using phase and fluorescence microscopy. All samples were stained with FM4-64 and mounted onto 1% agarose pads and imaged. Phase microscopy was performed on all images using the auto-exposure setting at 100x magnification. FM4-64 staining was assessed using a rhodamine filter at 50 ms fixed exposure. The Phase/FM4-64 composites were constructed using Zeiss AxioVision software. Contrast settings were normalised manually between images using AxioVision software (A) Liposome Only Control (B) Liposomes incubated for 1 h at 4°C with BSA control (C) Dlp2-6H only control (D-I) Liposomes incubated for 1 h at 4°C with Dlp2-6H and GTP. For all images, the scale bar is 2 μ m (bottom right hand corner of each panel)

Due to the observable aggregation under the microscope when Dlp2-6H and liposomes were mixed together, further investigation through a co-sedimentation assay was carried out. At two different concentrations, 15 μ M and 30 μ M, no discernible amount of Dlp2-6H was observed in the pellet fraction, i.e. associated with liposomes (*Figure 4.19A*), with minimal increase in the insoluble fraction when compared to the Dlp2 only control assay (below in *Figure 4.19B*). The high amount of Dlp2-6H still found in the soluble, non sedimented fraction, suggests that the protein carries out its potential membrane fusing function seen in *Figure 4.18*, in a more transient binding event that is most likely reversible or at least occurs at a low frequency as supported by the microscopy where unaffected liposomes are evident throughout the population.

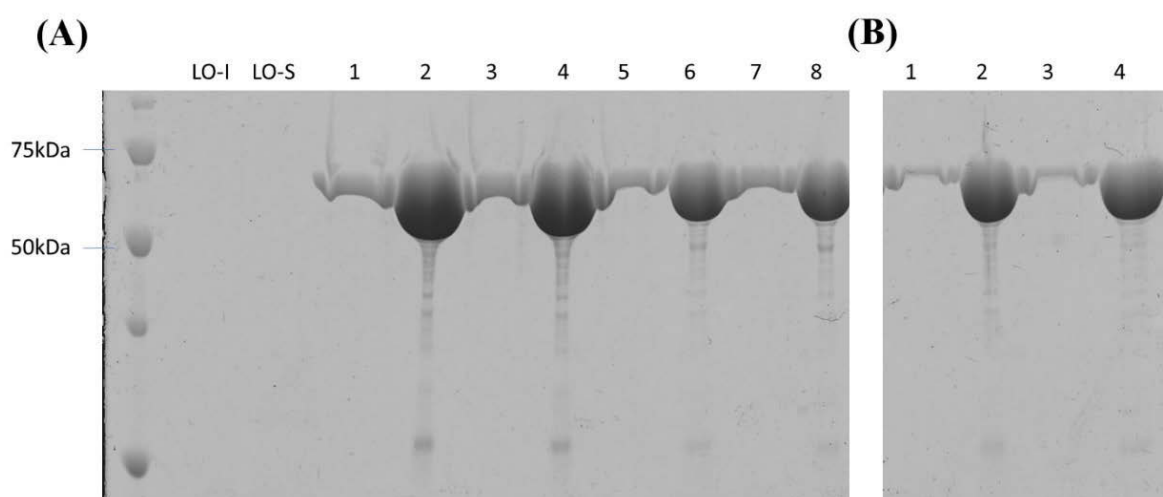


Figure 4.19 Co-Sedimentation of Liposomes with Dlp2-6H and Co-pelleting of Dlp2-6H
(A) SDS PAGE analysis revealed that at both 15 μ M (5-8) and 30 μ M (1-4) a potential weak interaction between Dlp2-6H and liposomes with the presence of the protein in the insoluble fraction both in the presence (1-2 and 5-6) and absence (3-4 and 7-8) of GTP. No protein contamination was seen in the liposome only insoluble (LO-I) and soluble (LO-S) fractions. (B) SDS PAGE analysis illustrates that Dlp2 does not polymerize and shift from the soluble fraction (2 and 4) to the insoluble fraction (1 and 3) with (1-2) or without (3-4) the presence GTP. Standards (Std); Protein Plus Standards (Bio-Rad).

4.3.6 INVESTIGATING POTENTIAL INTERACTIONS OF THE DLPS

It is well established that dynamins, both eukaryotic and bacterial, self-assemble as part of their normal function (6, 207). It is thought that due to the location of *dlp1* and *dlp2* (back-to-back) in the genome and the previous knowledge that dynamins do interact to carry out their function in a cell, a complex is hypothesised to form between Dlp1 and Dlp2. Therefore, a number of different approaches to studying this interaction were implemented. Firstly, to further investigate the ability of Dlp2-6H to self-polymerize as observed earlier in the gel filtration purification (*Figure 4.7*), a co-pelleting assay in the absence of liposomes was performed. Two concentrations of Dlp2-6H were incubated and spun at high speeds. As seen in *Figure 4.19B*, greater than 90% of the protein was observed in the soluble fraction, with the remaining less than 10% found in the insoluble fraction. With rigid washing steps used post spin to the pellet/ membrane fraction, the small pellet of protein may be self-polymerising Dlp2-6H. However, the lack/ low level of Dlp2-6H in the insoluble fraction does not mean that self-interaction is not possible, but it just suggests that there is no stable, long-chain polymerization occurring, causing the protein to become insoluble. Other factors required for potential Dlp2 polymerisation might also be missing from this assay such as lipids and/or Dlp1. The small amount of the protein sample self-interacting is supported by the seemingly lower abundance of the dimer interaction seen in *Figure 4.7* compared to the monomer.

Following on from this, it was thought that Dlp1 may be the interacting partner needed for the characteristic polymerization to occur. Therefore, to investigate this a Bacterial Two Hybrid (BacTH) interaction screen was set up in a checkerboard design, outlined in *Section 4.2.10*, where each protein of interest from the Dlp operon was tagged with both a T18 and T25 tag either on their C- (red) or N- (blue) terminus depending on predicted active sites i.e. Dlp1a, Dlp1 and Dlp2 were tagged on the C-terminus to ensure that the integrity of the GTPase domain remained intact. The BacTH technology requires co-expression of two hybrid proteins and their close proximity or interaction within the same recipient bacteria leading to cAMP synthesis and a positive phenotype in the recipient strain seen in the form of expression of the reporter gene, Lac. Each combination of plasmids outlined in *Section*

4.2.10 Table 4.4 was transformed into a non-reverting adenylate cyclase deficient (*cya*) *E. coli* reporter strain, BTH101, to detect the protein-protein interactions.

Each transformation was plated onto an indicator plate, LB-X-Gal, with complementation detected after 3 days and images taken (*Figure 4.20*). Despite all transformations seemingly being successful, with no plasmid instability noted, the results of complementation were inconclusive. Examining closely the LB-X-Gal plate, a slight difference can be seen between the positive and negative control transformations (red and green circled colonies), allowing small differences between the complementation to be observed. The strongest interaction on the plate was Dlp1 Δ TM with Dlp1a and the combination with the least complementation was Dlp1 Δ TM with Dlp2, potentially indicating a requirement for the TM domain for interaction, with Dlp1 and Dlp2 complementation showing a potentially positive reaction. However, additional testing is required to comment further on this.

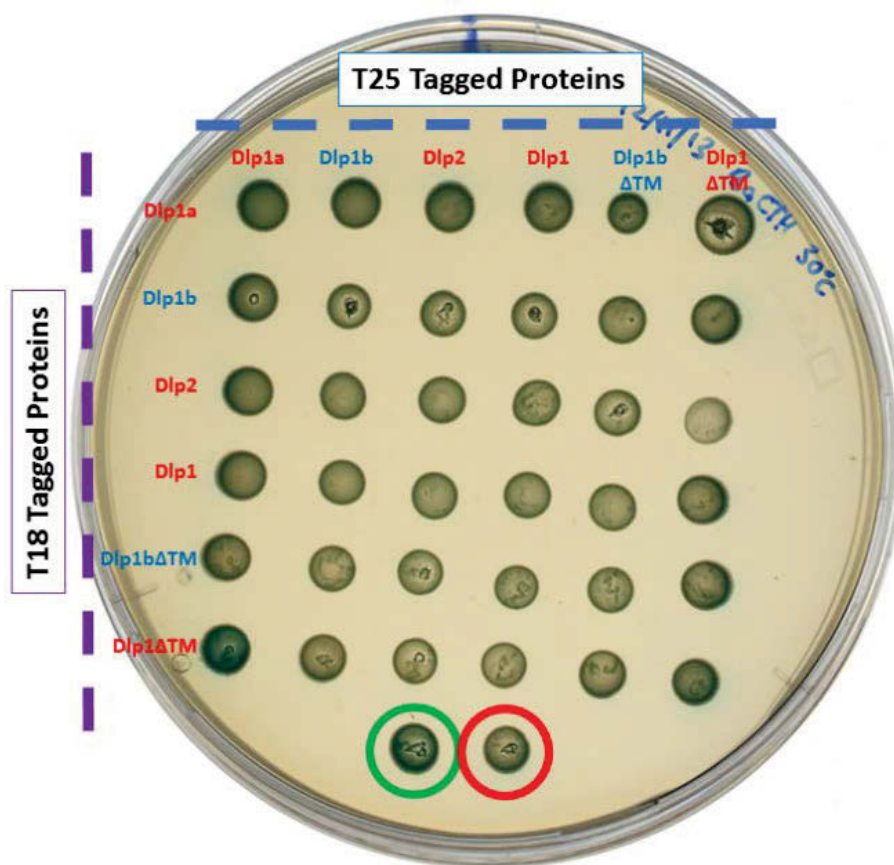


Figure 4.20 X-Gal Plate Results for BacTH Assay

A LB X-gal plate was used to detect complementation of the paired plasmids and after 3 days, an image was taken. In each transformation, a T25 (top) and a T18 (left) tagged protein expression plasmid was present. During expression, depending on plasmid construction the protein of interest could be tagged on the N (blue) - or C- terminus (red) depending on predicted functional domains. A positive control provided by BacTH assay kit (green circle) and an empty vector negative control (red circle) were also used. A blue colony indicates a positive interaction between the plasmids transformed and white (clear) colony indicating a null interaction. It can be seen that not much difference between the control positive (green) and negative (red) colonies is evident making it quite difficult to interpret the test interactions.

Due to the ambiguity of the indicator plates, a quantification approach was implemented to have a closer look at the potential interactions of the Dtps. The β -galactosidase assay provides a quantitative determination of the efficiency of the functional complementation between pairs of hybrid proteins. Using this assay, it was quite clear that in the BacTH interaction screen there was minimal to no interaction between the Dtps detected. The β -galactosidase assay showed that for ten selected pairs of hybrid proteins (numbers defined in *Table 4.4*), there was only slightly higher enzymatic activity to that of the negative controls of both empty vectors (EV) and BTH101 cells only (*Figure 4.21*). The positive control, which used a pair of -zip plasmids provided in the BacTH system, was excessively positive and, despite dilutions, was not a useful comparison for our possibly weak and/or transient interactions. These results were supported by the plating of the individual complementation transformations onto MacConkey plates, with the expected colour changes observed for the positive and negative controls. The test interactions (examples shown in *Figure 4.21*) showed an ambiguous colour change to neither confirm nor deny a positive or negative complementation.

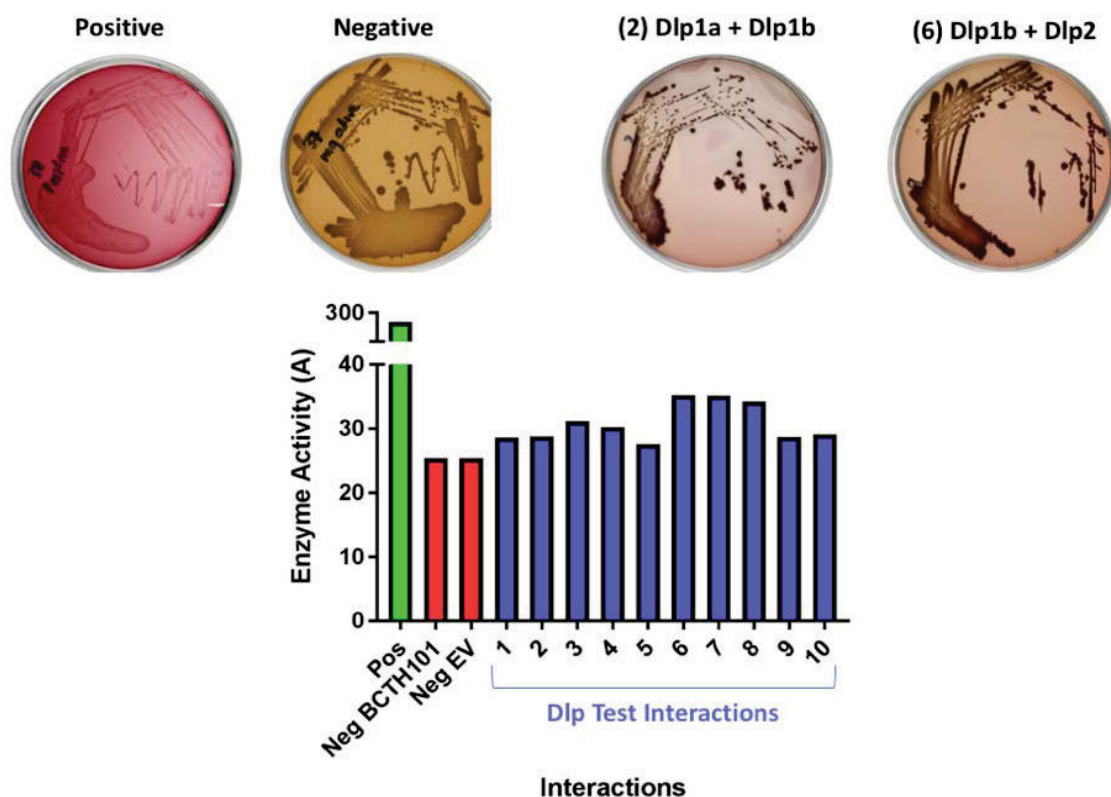


Figure 4.21 Quantification of Potential Dlp Interactions

A beta-galactosidase assay was used to quantify the level of enzymatic activity in the complementation transformations. Ten reactions were screened from the panel of 36 (blue), two negative controls of both empty vector transformation as well as cells only (red) were used and a positive control (green) complementation transformation was implemented using a pair of interacting Zip plasmids. The enzymatic activity was calculated as outlined in Section 4.2.10 and analysed in Graphpad. Each transformation was plated onto selective MacConkey agar plates, the positive transformation showed the appropriate pink colour change, the negative had the predicted yellow colour change and the test interactions had an ambiguous combination colour seen.

With the null result seen using the BacTH, a high throughput system of expression may not be sensitive enough to detect the type of interaction that may be occurring with the dynamins. Therefore, an *in vitro* approach is now being implemented to begin to look at this possible interaction. The use of gel filtration, allows the addition of the two recombinant proteins of interest, Dlp2-6H and Dlp1a-6H to be separated on a S300 column as both individual proteins (if they don't interact) and/or as a complex. As previously shown, Dlp2-6H during the gel filtration purification step, showed evidence of potential dimerisation during the run (*Figure 4.7*) and so it was thought that this approach would allow any interaction between these two proteins to be seen also. As an initial examination carried out by my colleague Nina Sulkowski, an equimolar ratio of Dlp2-6H and Dlp1a-6H was mixed and incubated for 1 hour before loading onto the S300 for a 16 h column run. As it can be seen in *Figure 4.22*, though preliminary, it seems that there is evidence of a potential interaction between Dlp1a-6H and Dlp2-6H with a repeatable elution of protein at 70 mL, just beyond the void peak. With two discernible peaks for the individual proteins Dlp2-6H and Dlp1a-6H seen at 110 mL and 130 mL respectively. Further investigation into this is definitely required, with repeated runs using higher resolution columns and SDS-PAGE analysis of the eluted proteins needed to confirm this interaction.

In summary, all these results taken together form the basis for developing an understanding of the biochemical properties of Dlps in *H. pylori*. Purified recombinant Dlp1a-6H and Dlp2-6H have been successfully used in investigating the *in vivo* expression and localisation of Dlps in *H. pylori*. Along with these results, they were used in establishing preliminary results for describing important lipid binding properties and self-interactions that are required for the normal function previously noted in the Dynamin superfamily.

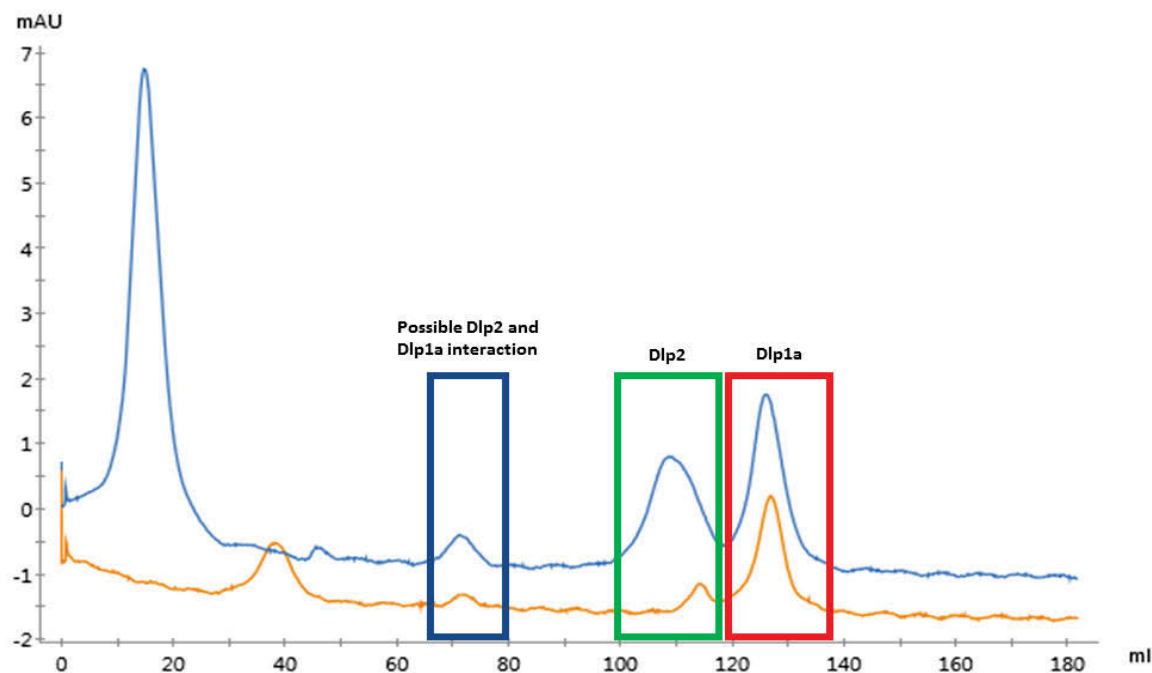


Figure 4.22 Preliminary Interaction between Dlp1a-6H and Dlp2-6H using Gel Filtration

Gel filtration was used to test for “natural” interactions between Dlp1a-6H and Dlp2-6H. The chromatograms for two replicate runs show evidence of three peaks in common. The earlier elution (around 70 mL) in the gel filtration specifies protein with a higher molecular weight, potential Dlp1a-6H and Dlp2-6H (blue box). The next elution at 110-120 mL is Dlp2-6H (green box), followed by Dlp1a-6H at 130 mL (red box). The earlier peaks are contaminating proteins as well as the void peaks of the runs.

4.4 DISCUSSION

The investigation of protein biochemical characteristics poses many difficulties surrounding both the experimentation and the results that are generated. Despite these difficulties, *in vitro* work allows a glimpse at a function, by linking, in this case, the hypothetical structure of a protein to its cellular function. In this set of experiments, answers to some of the fundamental questions surrounding the *H. pylori* dynamin biochemical properties were attempted; essential in classifying these proteins as members of the dynamin superfamily. This is thus the first such investigation into the presence of Dlp1a in *H. pylori* significantly adding more information to the bacterial dynamin field.

Dlp2 was successfully purified using traditional His-Tagged methods in its native form; however, Dlp1a posed a problem due to insolubility in its native purification. The seemingly absolute insolubility of Dlp1a was expected; with dynamins known to be membrane associated as well as a tentatively predicted transmembrane domain found at the C-terminus of the protein (2, 201)(Chapter III). The consistent presence of inclusion bodies in the over expression strain was a clear indication of the insolubility of Dlp1a. Inclusion bodies are by definition insoluble intracellular aggregates of proteins that are dynamic and conformationally diverse structures. They are formed by a sequence selective aggregation process that is likely driven by certain hot spots in the protein sequence, such as with eukaryotic dynamin members, the Mx proteins, having been highlighted to have this type of aggregation feature previously (208). We utilized the inclusion bodies to purify the protein quite early in the purification process, by isolating the inclusion bodies away from the whole cell lysate using a series of urea washes. Unfortunately, this requires the complete denaturation of the inclusion bodies to purify Dlp1a, and a slow gradient of decreasing urea was performed in an attempt to refold Dlp1a back into its native state. Due to the downstream applications, such as antibody production, the need for Dlp1a to be in a native conformation was not completely necessary. However, the use of this method to produce recombinant Dlp1a-6H for *in vitro* assays such as GTPase and lipid binding, confirming its conformation, will be necessary to ensure the correct biological function is assigned.

With the *H. pylori* Dlp purified, antibodies were able to be made to them. The use of antibodies in this study was multifactorial, with the overarching theme of their use being localisation. Firstly, *in vivo* detection of *H. pylori* Dlp was pivotal in establishing their presence in the cell. Previously it was shown that mRNA transcripts were detectable in 26695 for the *dlp* operon (Chapter III) and so this was further studied to see if the proteins are produced in *H. pylori*. This is the first report of Dlp being expressed *in vivo* in *H. pylori*. Currently as they sit in the NCBI database, they are considered hypothetical proteins with no protein expression data present and the only evidence of their presence in *H. pylori* being transcriptional data (192). From this the most interesting finding in the identification of their expression, was the presence of both the full length Dlp1 and the truncated Dlp1a (*Figure 4.13* last panel). The presence of both these bands in 26695 and B128.7.13 WT whole cell lysate was further evidence of some sort of phase variation controlling the expression of these genes. The confirmation of this was in the disappearance of both bands from the B128.7.13 Δ *dlp* clones (*Figure 4.13*). Also, the single clone whole cell lysate (*Figure 4.13*, middle panel) showed only the presence of the full length Dlp1, a result supported by sequencing data presented earlier (Chapter III). This potentially indicates that clonal selection is possible for the *dlp* operon, and that the switch between the two expressions is not a rapid transition but rather across multiple generations, greater than what was tested for in *Chapter III*. Further investigation, using western blot analysis into multiple generations of the single clone isolate could bring to light how and/or why the transition comes about.

Validating the efficacy and the specificity of these antibodies for *in vivo* Dlp was an important step in the use of them in investigating the localisation of Dlp in *H. pylori*. A key step in unravelling the function of a protein is identifying its location in the cell, as the protein will mostly be found where it is carrying out its function. Current literature surrounding *H. pylori* fractionation is restricted to proteomic studies of the membrane and secretome, due to these fractions often containing immunogenic proteins, which are of interest when identifying vaccine candidates (209-211). In this study though, the periplasm was of particular interest as the closest Dlp2 homolog, LeoA, was recently localised to this compartment in *E. coli* (5). LeoA is a dynamin-like protein requiring GTP to carry out its function therefore being localised in a compartment known to be nucleotide-free is quite

perplexing. However, due to the lack of compartmental controls to confirm a successful fractionation, isolation of the contents of the periplasm was not possible during this study. Ideally compartmental antibody controls are required to validate the sub-cellular fractionation but due to the lack of availability of these types of antibodies for *H. pylori* proteins, a new protocol using quantitative mass spectrometry is currently being developed in the lab to look at marker proteins in each compartment to validate the fractionation. Nevertheless, general soluble and membrane (containing mostly membrane proteins, with possibility of insoluble proteins) fractions were possible. With dynamins being membrane remodelling proteins and Dlp1 showing the presence of a highly predicted transmembrane domain, it was anticipated that Dlp1 would be membrane associated (1). The subcellular fractionation confirmed this, with a strong signal present in the membrane fraction using Western blotting (*Figure 4.14A*). On the other hand, the localisation of Dlp2 in *H. pylori* was variable; consistently observed in the soluble and in most membrane preparations (*Figure 4.14B*). From the fractionation of *H. pylori*, the presence of Dlp2 in the membrane fraction was intriguing, given it lacks a predicted lipid-binding domain. This may suggest that Dlp2 could localise to the membrane through interaction with a lipid- associated binding partner and therefore the exclusively membrane-associated Dlp1 could represent a promising binding candidate for Dlp2.

Localisation was further explored using IFM, a technique that unlike in other bacteria like *E. coli* and *B. subtilis*, has not been well established in *H. pylori*. The majority of *H. pylori* IFM studies have been performed on clinical specimens using antibodies that target the cell surface whilst studies that target intracellular proteins are scarce (212-216). The use of two strains, 26695 and B128.7.13, allowed the results to be strengthened through repetition and comparison especially with a deletion strain of the *dlp* operon available for B128.7.13. In 26695, Dlp1a despite having strong membrane association in the fractionation, showed a less convincing localisation via IFM, with a dispersed pattern throughout the length of the cell (*Figure 4.15*). Dlp2 also showed a dispersed pattern, but there were distinct foci in a sub population of the cells imaged that surprisingly was associated with the membrane, which supports the fractionation results (*Figure 4.15*). In B128.7.13, a successful deletion of the *dlp* operon has been created (discussed later in *Chapter V*) and was used in the IFM as a *dlp*

negative control, and highlighted the specificity of the fluorescence seen in the 26695 as well as B128.7.13. In B128.7.13, Dlp2 showed again a dispersed pattern throughout the length of the cell with punctuated foci evident in most cells and these foci are absent in the deletion strain, where only background fluorescence can be seen (*Figure 4.16 right panel*). These foci are slightly different to what was seen in 26695, with them existing along the width of the cell rather than along the length of the cell. Localisation of Dlp1a in B128.7.13 was slightly more difficult to discern as matching antibody concentrations were not available due to the fragility of the cells. Considering this, there was 50 times more antibody present in the deletion strain IFM, allowing us to recognise the specificity to Dlp1a where fluorescence above the background was still seen. A clear disruption to the fluorescence was seen in B128.7.13 Δ dlp cells compared to WT (*Figure 4.17, right panel*). Localisation of Dlp1a in B128.7.13 again was slightly different to 26695, with foci seen which were not seen in 26695, potentially indicating that localisation in the strains is different. The patchy localisation of both proteins throughout the length of the cell in both strains, indicates that there is no conserved region in the cell that dynamin functions, which could signify a more variable whole cell function for *H. pylori* dynamins similar to what is seen for DynA in *B. subtilis* (100). Further investigation using high-resolution microscopy, such as using 3D- structured illumination microscopy would allow a much more convincing look at the localisation of these proteins in *H. pylori*, with this study establishing the baseline conditions needed for IFM in both strains. Better still, a GFP-fusion of Dlp would allow a more native look at the localisation of the proteins *in vivo*, without having to introduce antibodies to the cells.

A conserved feature of all dynamins, both eukaryotic and bacterial, is their ability to bind to and manipulate lipids *in vitro* in a GTP dependent manner (6, 33). To begin to look at the dynamin features of *H. pylori* Dlps, lipid binding abilities were investigated through a sedimentation assay (*Figure 4.19A*) and lipid-binding microscopy (*Figure 4.18*) both in the presence of hydrolysable GTP. When purified Dlp2-6H was mixed with pre-formed liposomes the result was rather striking as visualised using microscopy. The previously uniform liposomes aggregated, providing clear evidence of potential membrane fusion occurring. This type of membrane fusion has been seen previously in the function of DynA (3). The opposite process of membrane fission would show evidence of the liposomes being

pinched, which is not seen in these images. Despite the changes to the lipids seen in the microscopy, the sedimentation assay showed minimal evidence to support this, with only a small amount of Dlp2-6H binding to the lipid fraction. This could indicate that this process is transient, where by Dlp2-6H associates for a short period of time, to carry out the membrane changes seen in the microscopy and disassociates from the membrane soon after. Eukaryotic dynamin is known to have a transient association with a lipid membrane, providing enough time to constrict the neck of a budding vesicle in a GTP dependent manor but once the process is complete and GTP is hydrolysed to GDP, it disassociates (207). The transient binding is supported by the entire population not being seen to form the aggregates with clear fields of view showing minimal binding of the protein to the liposomes (*Figure 4.18I*). Maximum binding of Dlp2-6H was potentially not achieved due to the use of *E. coli* total lipid extract rather than *H. pylori* specific lipid extract, which has a slightly different composition. A key difference seen is the presence of unique components in the *H. pylori* membrane called cholesterol glucosides which account for approximately 25% w/v of the total lipid extract from *H. pylori* (217). Therefore, future work should investigate the ability of the *H. pylori* dynamins to bind to *H. pylori* lipid extract specifically. Similarly, there are a number of factors that could be attributed to this type of liposome phenotype such as disruption to the surface tension of the liposomes by the addition of the protein to the sample. Therefore, to ensure that Dlp2-6H is interacting with the liposomes directly, fluorescently labelling the protein with an antibody or fluorescent tags would allow co-localisation to be determined between the protein and the liposome.

To carry out its function in a cell, dynamin usually polymerises or interacts with a partner to manipulate the membrane (218). Self-interaction is noted throughout the dynamin family, with several different members of both the eukaryotic and bacterial domains showing evidence of polymerisation (3, 98, 218). No polymerisation of Dlp2-6H was seen in the co pelleting assay (*Figure 4.19B*) but this isn't definitive enough to say that Dlp2-6H doesn't self-interact, however it does rule out long chain polymerisation. Other possibilities for self-interaction are dimerisation and trimerisation and would potentially not be seen in this type of assay but would be detected very clearly by gel filtration. As seen in *Figure 4.7*, during gel filtration purification, consistently present in an early peak in the run, at approximately

100 mL, Dlp2-6H was found. This volume does not correspond with the size of monomeric Dlp2-6H (180 mL) and could indicate self-interaction, in the form of a dimer occurring for Dlp2-6H. This is the first evidence of self-interactions occurring in *H. pylori* dynamins, a key fundamental feature for all dynamins.

Bacterial dynamins seemingly always come in pairs in the genome, with a few cases of particular interest to this study; of either a fused dynamin, DynA or a split dynamin LeoC/LeoB and obviously in *H. pylori*, the split dynamin is present in the form of Dlp1a and Dlp1b (3, 5). In regards to the split dynamin, one would think that the individual parts would need to interact in some way to carry out their dynamin function with each part containing necessary domains for full function of a dynamin protein. Therefore, a BacTH system was implemented to identify any interactions between the *H. pylori* dynamins and to investigate if they formed a complex. BacTH has been utilised previously in identifying interaction partners of DynA both with itself as well as other proteins. Burmann *et al*, created a library from the *B. subtilis* genome in the BacTH plasmids and screened it against the full-length DynA (103). Ideally, a system like this could be utilised like this in the future, but initially the focus was on self-interaction and interaction between the Dlp's in *H. pylori*. However, no strong interaction was detected between the dynamins in this testing system (*Figure 4.20 and Figure 4.21*). A potential interaction of Dlp1 with itself in the absence of the transmembrane domain was the strongest interaction, along with a potential interaction between Dlp1 without the transmembrane domain and Dlp2. Though, when further tested in a quantitative beta-galactosidase assay these interactions along with the others tested showed minimal enzymatic activity above the negative control (*Figure 4.21*). It is important to note the two-hybrid screening has many limitations including, the addition of quite a large tag to the proteins of interest that may disrupt important domains needed for interaction. With no known structure of *H. pylori* dynamins, the plasmid was constructed based on homology to ETEC proteins, with proteins tagged away from important functional domains like the GTPase domain (Dlp1a, Dlp1 and Dlp2) and the transmembrane domain (Dlp1b). Despite these efforts, the BacTH was not sensitive enough to detect any interaction between *H. pylori* dynamins. However, due to the limitations of BacTH, belief in the interaction between Dlp1 and Dlp2 remained. An initial look examination of the interaction between Dlp1a and Dlp2 has begun

with the implementation of gel filtration as a means to test for this weak interaction and has shown promising results. With an equimolar mix of Dlp2-6H and Dlp1a-6H, a discernible peak that may correspond to a native complex forming between the proteins was evident during a gel filtration run, with this peak well before the established volumes for elution for the monomeric proteins (*Figure 4.22*). These are however, preliminary results and need further confirmation via Native PAGE analysis to show that Dlp1a and Dlp2 do in fact interact.

In summary, an initial investigation into the biochemical features of the *H. pylori* dynamins, Dlp1a and Dlp2 has been conducted and showed a number of promising results. Firstly, this is the first study to show *in vivo* expression of Dlp2 as well as both forms of Dlp1, full and truncated (Dlp1a). Furthermore, Dlp1a was localised to the membrane of *H. pylori* cells, with Dlp2 showing a split localisation between the membrane and cytoplasm, the first indication of a membrane localised interaction. In future research, higher resolution microscopy and/or GFP fusions would allow a more detailed look into the co-localisation at the membrane of Dlp1a and Dlp2. Lipid binding features as well as self-interaction have been established for Dlp2, two fundamental features of dynamin proteins. In addition, a preliminary examination into the possibility of Dlp1a and Dlp2 interacting to form a complex potentially for function showed encouraging results. Future studies into this potential interaction could include more direct binding assays such as thermophoresis as well as less direct assays such as affinity chromatography. This biochemical study into the properties of the *H. pylori* dynamins will continue to add a deeper understanding to the research into bacterial dynamins, where multifunctionality despite similar underlying biochemical properties remaining is beginning to be established just like the eukaryotic members of the superfamily.

CHAPTER V

WHAT IS THE FUNCTION OF DYNAMIN-LIKE PROTEINS (DLPS) IN *HELICOBACTER PYLORI*?

PREFACE

With the demonstration of the conservation and expression of Dlp1 in *H. pylori*, their biochemical properties were investigated in Chapter IV. Purified recombinant Dlp1a (*Section 4.3.2*) and Dlp2 (*Section 4.3.1*) were utilised for a number of downstream assays. Firstly, they were used for antibody production allowing *in vivo* protein expression to be confirmed (*Section 4.3.3*) as well as sub cellular localisation to be explored via fractionation and immune-fluorescence microscopy. It was found the Dlp1 localised repeatedly to the membrane and Dlp2 was split between the membrane and soluble fractions with both proteins showing no distinct localisation pattern, despite the antibodies proven to be specific (*Section 4.3.4*). Moving on from this, to provide a more direct look at the ability of the proteins to interact with a membrane, *in vitro* liposome binding assays were able to show that Dlp2 could in fact change the shapes of liposomes in the presence of GTP (*Section 4.3.5*). Finally, with interactions known to be important for dynamin function, the potential interaction between the *H. pylori* Dlp1 was explored firstly using a BacTH system, where no discernible interactions were detected (*Section 4.3.6*), and then using more ‘native’ or ‘direct’ gel filtration methods (*Section 4.3.6*). In these preliminary gel filtration studies, evidence of potential interactions both between Dlp2 and itself (*Section 4.3.1*) as well as Dlp1a were seen, with follow-up needed. With this understanding of the biochemical characteristics of Dlp1 from *H. pylori*, their function was then explored, through the creation of a Dlp1 knock out strain and downstream phenotypes tested, the focus of Chapter V.

5.1 INTRODUCTION

The dynamin superfamily of proteins plays a fundamental role in membrane dynamics within several cellular compartments of the eukaryotic cell. These include the cell membrane, at the mitochondria and the golgi apparatus (2). Classical dynamins utilise the energy from GTP hydrolysis to sever membrane vesicles from the cell surface by catalysing membrane fission during the processes of endo- and exocytosis, although this is not the only function they have in the eukaryotic cell (33). Other members of the superfamily, such as Dynamin-like proteins, have a diverse range of functions although comparatively less is known about them. These functions include fission and fusion of the mitochondrial membrane during division by DRP1 and mitofusins, fusion of the endoplasmic reticulum by Atlastin, and antiviral activity exhibited by Myxovirus resistance (Mx) proteins (2). It is evident that whilst members of the dynamin superfamily differ in their functions, there is a common, overarching theme of membrane remodelling.

Much of what is known about the dynamin superfamily is restricted to eukarya alone, with the function of bacterial dynamins remaining highly elusive. This is despite the identification of over 1400 dynamin-like proteins in approximately 1000 species of bacteria (6). The study of bacterial dynamins has been limited to structural analysis with minimal research in regards to their biological function(s) within the cell. The first bacterial dynamin to be identified, BDLP, has been shown to bind and modulate lipid dynamics in *N. punctiforme*, but its specific function has not yet been identified (4, 91). There have also been investigations into the functions of other bacterial DLPs outlined in Chapter I, such as DynA in *B. subtilis* and LeoA in *E. coli* (5, 100).

DynA, a fused dynamin-like protein, has been implicated in the maintenance of membrane integrity in the presence of membrane targeted stresses, such as phage attacks or antibiotics like the bacteriocin peptide, Nisin (100). It is proposed that in unstressed cells, DynA is mobile along the cytoplasmic side of the membrane. Upon membrane pore formation however, DynA is recruited to the site of damage where it appears to polymerise into static

assemblies and repair the gaps in the membrane, possibly by membrane fusion (100). LeoA on the other hand, has been implicated in toxin delivery in ETEC.

It is well established that ETEC utilises vesicles for the delivery of heat labile (LT) toxin to the host during infection (99, 219). The role of LeoA in this process has been supported by studies which showed a significant reduction in the release of the toxin into the supernatant of Δ LeoA strains and a clear 50% reduction in vesicle formation by the cell (5, 7). Taken together these studies suggest LeoA plays a role in vesicle release and perhaps formation. Both proteins appear to differ greatly in function, with DynA showing a key role in the maintenance of membrane integrity when under challenge from the environment and LeoA seemingly involved in the release of vesicles during infection (5, 100). This variability in function is somewhat reminiscent of the eukaryotic dynamin family, which carry out diverse functions in the cell. It is thus clear that bacterial DLPs, although underappreciated, perform diverse roles in the modulation of membrane dynamics, much like eukaryotic members of the dynamin superfamily.

H. pylori as stated previously, is an understudied and underappreciated pathogen of the human gastrointestinal tract. It is a Gram-negative pathogen that colonises the gastric epithelium and induces an inflammatory response by the release of toxins such as VacA, notably by the release of outer membrane vesicles (OMVs) (173). The mechanism behind their release remains unknown. Additionally, *H. pylori* possesses the unique ability to survive the harsh, acidic environment of the stomach (where pH levels can reach as low as 2). To overcome these stressors, *H. pylori* has adapted several internal pathways with an essential component to this defence being a robust and protective membrane (122). With the diverse and still largely unknown functions noted for bacterial DLPs, the study of *H. pylori* DLPs could aid in furthering our understanding of this superfamily of proteins. Given the role of DynA in the maintenance of membrane integrity and the role of LeoA in the release of membrane vesicles, it is compelling to explore the function of these proteins in *H. pylori*. Further insights into DLP function in *H. pylori* may provide important insights into the pathogenic mechanisms orchestrated by the organism, which have allowed it to successfully

colonise over half the world's population, as well as help in the development of relevant therapeutic interventions.

The aims of this chapter were to;

1. Construct Δdlp and $\Delta::dlp$ strains in *H. pylori* strains 26695, B128.7.13 and SS1
2. Investigate the potential phenotype of the Δdlp strain to help elucidate the function of dynamin in *H. pylori*.

5.2 MATERIALS AND METHODS

5.2.1 PLASMID CONSTRUCTION

Using traditional cloning techniques (*Section 2.5.6*) pBSII-SK (*Table 2.2*) was utilized as the plasmid backbone for insertion of two flanking regions (~750bp in length) of the *dlp* operon deletion of interest (outlined in *Section 5.3.1*, *Figure 5.2*). Each flank of the region of interest was amplified using site specific primers (*Table 5.1*) using PCR (*Section 2.5.4*). Each flank was designed with a region of overlap to allow for an additional insertion of the *kan-sacB* cassette (220). Therefore, to join the two flanks together, equimolar ratio of both flanks was mixed together in a PCR reaction (without primers) and an overlap extension PCR was performed for 30 cycles. The full-length insertion (joined flanks, ~1.5kb) was gel purified as outlined in *Section 2.5.2*. The purified flank insert and pBSII were digested with BamHI and SalI and ligated together. The ligation was transformed into DH5 α as outlined in *Section 2.4.3* and clones were selected for using ampicillin on LB agar (*Table 2.3*). Initial plasmid construction was confirmed by PCR (*Section 2.5.4*) and plasmid sequence was confirmed by Sanger sequencing (*Section 2.5.5*).

The cassette (*kan-sacB*) was housed in a plasmid backbone, pSKFII provided by Richard Ferrero (*Table 2.2*). The cassette was purified from pSKII through digestion with SmaI and XhoI. The cassette was gel purified away from the rest of the plasmid after digestion as outlined in *Section 2.5.2*. To then generate two blunt ends, the cassette was treated with Klenow fragment (New England BioLabs). The pBSII plasmid was digested with SmaI, and a blunt ended quick ligation (New England BioLabs) was performed with the pre-treated cassette at room temperature (RT) for 30 min. The ligation was transformed into DH5 α as outlined in *Section 2.4.3* and clones were selected for using 100 μ g/mL ampicillin on LB agar (*Table 2.3*). Initial plasmid construction was confirmed by PCR (*Section 2.5.4*) and plasmid sequence confirmed by Sanger sequencing (*Section 2.5.5*). A final general plasmid map can be seen in *Figure 5.1*.

Table 5.1 Primers Used for Amplification of Flanking Region of dlp for pBSII Plasmid construction

Primer Name	Sequence
26695C_US_F	TTTCCCGGATCCCCCATCCAGCACATCAGCGACTCCA
26695C_US_R	CCCGGGTCTGGTTGTAGATGCAGGCTCGAGGCTTTCACATCA AGATAAATGTTTTTCAT
26695B_DS_F	CTCGAGCCTGCATCTACAACCAGACCCGGGAATAATATCAA AACAAGGAGCATGCAATGA
26695B_DS_R	TTTCCCGTCGACGCGACAATACTCAAACCCTCTTTTTCT
A_US_F2	TTTCCCGGATCCCTTGAGAGAAAAAATGTATGCGCATATTAA
A_US_R2	CCCGGGTCTGGTTGTAGATGCAGGCTCGAGAAATTGTTCTAG CGTTTCATTCTTCAT
A_DS_F2	CTCGAGCCTGCATCTACAACCAGACCCGGGGTTGAGCCTGA AAGAGGCATGAACTGA
A_DS_R2	TTTCCCGTCGACTAGGGATCAAACCTCTCATCAATATCTTTAG
TMB_US_F	TTTCCCGGATCCGGACTGATTTTGATAAAAACAAAC
TMB_DS_R	TTTCCCGTCGACGAGCGCGTATAAAATGAGGTG
TMB_US_R	CCCGGGTCTGGTTGTAGATGCAGGCTCGAGGTCAGAGCAAG CCAACCCATA
TMB_DS_F	CTCGAGCCTGCATCTACAACCAGACCCGGGAGTTCAAGATAT CAAAGATCCCAAC
B128C_US_F	GGGCCCCGGATCCTTTTAAAAAACTGATAATAGAAATAGAA A
B128C_US_R	CCCGGGTCTGGTTGTAGATGCAGGCTCGAGATTTGATGCCCC ATGTAGTGTTTGCCTAAA
B128A_DS_F	CTCGAGCCTGCATCTACAACCAGACCCGGGCCCCCGCTTATC GGTCTTGCTTTTGAAGTGTGG
B128A_DS_R	GGGCCCCGTCGACTATTTCAAAATCTCTGTGAGATTCTGCTTG
SS1C_US_F2	GGGCCCCGGATCCAAAGATGCGGCGCAACTCCAG
SS1C_US_R2	CCCGGGTCTGGTTGTAGATGCAGGCTCGAGTCTTTTCTACGC CAGAATAAATGTTT
SS1A_DS_F2	CTCGAGCCTGCATCTACAACCAGACCCGGGTTGAGCCTGAA AGAGGAATGAACTGA
SS1A_DS_R2	GGGCCCCGTCGACAAACTCTCATCAATCTCTTTAGAAACC



Figure 5.1 General Plasmid Map for all Deletion Constructs

pBluescript II SK- forms the backbone of the plasmid with a *f1* origin under the T7 promoter with Ampicillin resistance. Also present in the plasmid backbone is the *lac* promoter and operon along with the T3 promoter. *H. pylori* upstream and downstream *dlp* operon flanks (approximately 750 base pairs) were inserted using *KpnI* and *BamHI*, with an overlapping region designed with an internal *SmaI* site. This intermediate plasmid was cut with *SmaI* and the *kan-sacB* cassette was cut with *XhoI* and *SmaI*, treated with klenow fragment to fill in the overhangs to form two blunt ends and ligated together to form the complete plasmid. The cassette is under a *H. pylori flaA* promoter.

5.2.2 *H. PYLORI TRANSFORMATION*

The *H. pylori* strain (*Table 2.1*) of interest was prepared as outlined in *Section 2.4.1*. A patch of growth from this strain preparation was taken and smeared over the surface of fresh Blood agar (*Table 2.3*). The plate was then incubated overnight in a microaerophilic environment at 37°C (CampyGen 2.5L). Once grown, a loop of cells was transferred to a new plate and smeared into a new patch. This was performed for each individual transformation, including a negative control to which no DNA was added. These patches were incubated in a microaerophilic environment at 37°C for 6 h (CampyGen 2.5L), after which varying concentrations of plasmid DNA (including 0.2, 0.5, 1 and 2 µg) were added to each patch of cells and left overnight in a microaerophilic environment at 37°C (CampyGen 2.5L). All patches were lifted and smeared over their own new blood agar plate (*Table 2.3*) for selection with 20 µg/mL Kanamycin (Km20) and left to incubate in a microaerophilic environment at 37°C for 5 days (CampyGen 2.5L). The colonies that appeared after this incubation, they were re-patched onto a fresh blood agar plate (*Table 2.3*) with Km20 selection for further growth. These were then screened by PCR for the correct insertion of the cassette.

This same procedure was followed for the complementation of the deletion strain with the WT *dlp operon* with a few minor changes. These changes include; instead of plasmid DNA being used, the WT B128.7.13 *dlp operon* was amplified using strain specific primers (*Table 5.1*) in PCR (*Section 2.5.4*) and the purified product sequenced (*Section 2.5.5*). This was then used as the transformation template DNA source and the selection of successful clones was made on blood agar plates (*Table 2.3*) with 20% sucrose. These were then screened by PCR for the correct insertion of the cassette.

5.2.3 WHOLE GENOME SEQUENCING AND DATA ANALYSIS OF B128.7.13 STRAINS

gDNA was purified from B128.7.13 WT, Δdlp C1, Δdlp C2 and $\Delta::dlp$ C2 as outlined in Section 2.5.3. Whole genome sequencing was performed by Genome Quebec on the MiSeq Platform and the subsequent demultiplexed 150 bp paired end reads were placed into FASTQ files and analysed at UTS. Each FASTQ file underwent quality control analysis using the FastQC tool. The template genome, B128.7.13 was constructed using the scaffolds available on National Centre for Biotechnology Information (NCBI) (Accession - LAQK000000000) (221), aligned to the B8 genome (222) by Mauve (223). The complete draft B128.7.13 genome generated was then annotated using the program Prokka (224). The quality checked FASTQ files for WT B128.7.13, B128.7.13 Δdlp C1 and B128.7.13 Δdlp C2 along with B128.7.13 $\Delta::dlp$ C2 were aligned to this template using Bowtie2.0 (225) with default settings and the output alignment visualised using Artemis (226) and BRIG (227). Statistics of sequencing reads were generated by using the SAMtools software package and the Flagstat command (228)

5.2.4 HAEMOLYSIS OF BLOOD AGAR

B128.7.13 WT, Δdlp C2, $\Delta::dlp$ C2 strains from glycerol stocks were patched onto fresh Blood agar (Table 2.3) with selection and incubated for 48 h in a microaerophilic environment at 37°C (CampyGen 2.5L). A single patch from all strains was inoculated into 5mL of fresh Brucella Broth (Table 2.3), with selection and incubated for a further 24 h in a microaerophilic environment at 37°C with shaking (120 rpm) (CampyGen 2.5L). Cultures were diluted down to OD_{600nm} 0.05 and left to grow to an OD_{600nm} of approximately 0.4. The cultures were then normalised to 0.4 and 25 µl was patched onto fresh blood agar (Table 2.3) in technical replicate (n=4) and incubated for 24 h in a microaerophilic environment at 37°C (CampyGen 2.5L) and photographed at the end of the incubation.

5.2.5 MEMBRANE INTEGRITY STAINING

B128.7.13 WT, Δdlp C1, Δdlp C2 and $\Delta::dlp$ C2 strains from glycerol stocks were patched onto fresh Blood agar (*Table 2.3*) with selection and incubated for 48 h in a microaerophilic environment at 37°C (CampyGen 2.5L). A single patch from all strains was inoculated into 5 mL of fresh Brucella Broth (*Table 2.3*), with selection and incubated for 24 h in a microaerophilic environment at 37°C with shaking (120 rpm) (CampyGen 2.5L). All cultures were diluted down to an OD_{600nm} of 0.05 and left to grow to an approx. of OD_{600nm} 0.4. Cells (1 mL) were taken and centrifuged at 1000 x g at RT for 5 min. Then resuspended in 100 µl of 1 x PBS (*Table 2.5*), 5 mM FM4-64 and 1 mM Sytox green dye and left for 30 min at RT. 2 µl of each culture was spotted onto 1% 1 X PBS (*Table 2.5*) agarose pads. Microscopy was then carried out using a Zeiss (Axioplan 2 Model) fluorescence microscope, with a 1.4 NA 100x phase contrast objective lens, under oil immersion. The Zeiss FITC and Rhodamine fluorescence filter-sets with 100 ms and 50 ms exposures were used for visualisation of the Sytox and FM4-64 fluorescence, respectively. The AxioCam MRm camera was used for image collection. Images were processed and analysed using AxioVision software, version 4.8 (Zeiss).

Further microscopy analysis was used for the visualisation of the Sytox and FM4-64 fluorescence using FITC and TRITC filter sets at 500 ms and 300 ms respectively on the DeltaVision Elite inverted research microscope with a 1.4 NA 100 x UPlanFLN objective, InsightSSI illumination, SoftWorX acquisition software, fitted with a WeatherStation environmental chamber (Applied Precision, GE Healthcare, Issaquah, WA, USA) and a scientific CMOS 15-bit camera (pco.edge, PCO AG, Kelheim, Germany). Images were analysed by ImageJ software version 1.49 (NIH).

5.2.6 GROWTH CURVE IN NEUTRAL CONDITIONS

B128.7.13 WT, Δdlp C1, Δdlp C2, $\Delta::dlp$ C1 and $\Delta::dlp$ C2 strains from glycerol stocks were patched onto fresh Blood agar (*Table 2.3*) with selection and incubated for 48 h in a microaerophilic environment at 37°C (CampyGen 2.5L). A single patch from all strains was inoculated into 5 mL of fresh Brucella Broth (*Table 2.3*), with selection and incubated for 24 h in a microaerophilic environment at 37°C with shaking (120 rpm). The cultures were then diluted to an OD_{600nm} of 0.005 and left to grow in the same conditions until they reached mid log phase. Following this, the cultures were further diluted to an OD_{600nm} of 0.05 and split into individual flasks for OD_{600nm} measurements at the selected time points. All flasks were then left to grow for their allocated time 12, 15, 18, 24 and 48 h, after which OD_{600nm} measurements were recorded.

5.2.7 EXPOSURE TO ACID

B128.7.13 WT, Δdlp C1, Δdlp C2, $\Delta::dlp$ C1 and $\Delta::dlp$ C2 strains from glycerol stocks were patched onto fresh Blood agar (*Table 2.3*) with selection and incubated for 48 h in a microaerophilic environment at 37°C (CampyGen 2.5L). A single patch from all strains was inoculated into 5 mL of fresh Brucella Broth (*Table 2.3*), with selection and incubated for 24 h in a microaerophilic environment at 37°C with shaking (120 rpm). The cultures were then diluted to an OD_{600nm} of 0.005 and left to grow in the same conditions until they reached mid log phase (greater than an OD_{600nm} of 0.5). Following this, the cultures were centrifuged (1000 x g, RT, 5 min) and resuspended to an OD_{600nm} of 0.5 in Brucella broth (*Table 2.3*) pH adjusted using HCl to either pH 2.1 or pH 4.9. When acid was present in the media, 10 mM or 100 mM urea was also added. All cultures were left to grow for a further 1 h with shaking (120 rpm) in a microaerophilic environment at 37°C (CampyGen 2.5L) after which OD_{600nm} were recorded. A two tailed, paired unequal variance t-test was performed using Excel on all replicates.

5.2.8 ONE POINT MEASUREMENT OF GROWTH IN ACID

B128.7.13 WT, Δdlp C1, Δdlp C2, $\Delta::dlp$ C1 and $\Delta::dlp$ C2 strains from glycerol stocks were patched onto fresh Blood agar (*Table 2.3*) with selection and incubated for 48 h in a microaerophilic environment at 37°C (CampyGen 2.5L). A single patch from all strains was inoculated into 5 mL of fresh Brucella Broth (*Table 2.3*), with selection and incubated for 24 h in a microaerophilic environment at 37°C with shaking (120 rpm). The cultures were then diluted to an OD_{600nm} of 0.005 and left to grow in the same conditions until they reached mid log phase (greater than an OD_{600nm} of 0.5). Following this, the cultures were centrifuged (1000 x g, RT, 5 min) and resuspended to an OD_{600nm} of 0.5 in Brucella broth (*Table 2.3*) pH adjusted using HCl to either pH 3 or pH 5. When acid was present in the media, 10 mM or 100 mM urea was also added. All cultures were left to grow for a further 72 h with shaking (120 rpm) in a microaerophilic environment at 37°C (CampyGen 2.5L) after which OD_{600nm} were recorded. A two tailed, paired unequal variance t-test was performed using Excel on all replicates.

5.2.9 GROWTH CURVE IN ACID

B128.7.13 WT, Δdlp 1, Δdlp 2, $\Delta::dlp$ 1 and $\Delta::dlp$ 2 strains from glycerol stocks were patched onto fresh Blood agar (*Table 2.3*) with selection and incubated for 48 h in a microaerophilic environment at 37°C (CampyGen 2.5L). A single patch from all strains was inoculated into 5 mL of fresh Brucella Broth (*Table 2.3*), with selection and incubated for 24 h in a microaerophilic environment at 37°C with shaking (120 rpm). The cultures were then diluted to an OD_{600nm} of 0.005 and left to grow in the same conditions until they reached mid log phase. Following this, the cultures were further diluted to an OD_{600nm} of 0.05 in Brucella Broth (*Table 2.3*) pH adjusted with HCl to pH 4.9 along with 10 mM urea. This diluted culture was split into individual flasks for each time selected for OD_{600nm} measurement. All flasks were then left to grow for their allocated time, 12, 15, 18, 24 and 48 h, after which OD_{600nm} measurements were recorded. At each time point, along with the OD_{600nm} measurements, the culture was centrifuged at 1000 x g at RT for 5 min, the supernatant removed, and the pH measured.

5.3 RESULTS

5.3.1 CREATION OF B128.7.13ΔDLP STRAINS

In designing the deletion studies, a number of different constructs were created in order to study the function of a number of elements within the *dlp* operon in *H. pylori* (Figure 5.2). In the *H. pylori* strain utilised for protein production and expression studies, 26695, the phenotypic effects of knocking out the entire operon, along with the individual genes themselves, Dlp1 and Dlp2 were to be studied (Figure 5.2A, B, C). A key functional feature of the dynamin superfamily is the ability to bind to and modulate membrane dynamics. In an effort to understand the importance of this domain in *H. pylori* DLP's, a strain was to be constructed in which the predicted transmembrane domain alone, had been knocked out (Figure 5.2D). Furthermore, future *in vivo* mouse studies would allow us to observe the effect of knocking out the dynamin-like proteins, on the ability of *H. pylori* to colonise the mouse stomach and cause infection. Therefore, constructs for knockout of the entire *dlp* operon were created for two mouse adapted strains; B128.7.13 (Figure 5.2E) and SS1 (Figure 5.2F) (sourced from Richard Ferrero) both of which share approximately 70% sequence identity to the 26695 *dlp* operon.

Using the pBlueScriptII –SK backbone, 750bp recombination areas of homology, either side of the region of interest for deletion, was inserted. A cassette (220) containing a kanamycin resistance gene (*kan*), *sacB* for complementation selection, and the *flaA* promotor from *H. pylori* was inserted in between the two flanks, creating the plasmid for recombination (Figure 5.1). Ultimately, this cassette would be placed into the *H. pylori* genome in the place of the *dlp* genes through natural recombination of the upstream and downstream flanks of the region of insert. Each plasmid was used with the appropriate strain to attempt a natural recombination transformation. The strains were then monitored for growth over the time frame of 2 days on blood agar plates and then spotted onto fresh plates for the transformation, in a spot with approximately a 2 cm diameter. Varying plasmid DNA concentrations for each transformation were tested from 0.2 µg - 2 µg, and a number of variations to the method were tested for the transformation, including linearisation of the DNA, ligation of the DNA and both plate and liquid transformations.



Figure 5.2 Overview of the Planned *H. pylori* Dynamain Deletion Strains

A number of different constructs were designed to knockout varying parts (--) of the operon to study the function of dynamins in *H. pylori*. These constructs include; (A) a construct to investigate the phenotypic effects of a complete knockout of the *dlp* operon in the type strain 26695 (B), two constructs to study the function of the individual genes, *dlp1* and *dlp2* in 26695 (B) and (C) and a construct to explore the importance of the transmembrane domain in 26695 (D). For *in vivo* mouse studies, two constructs were designed to look at phenotypic effects in two mouse-adapted strains, B128.7.13 and SS1 (E) and (F).

Despite exhaustive attempts, the only strain that successfully underwent natural recombination to incorporate the plasmid for deletion of the full *dlp* operon was the gerbil adapted strain, B128.7.13. As seen in *Figure 5.3*, PCR was performed to confirm the position of the cassette in the *H. pylori* genome, using both internal and external primers. Firstly, to ensure the cassette was inserted into the correct position in the genome, the flanking primers of the operon were used, observing for a slight shift in size from 3.9 kb to 4.5 kb with the insertion of the cassette (*Figure 5.3A*). The orientation and location of the cassette was then confirmed in the genome using a downstream primer external to the flanking region looking for the presence of a product that would be absent in the WT, where the cassette would be absent (*Figure 5.3B*). From this, it was confirmed that the *dlp* operon had been successfully knocked out in B128.7.13. Herein, two individual biological clones of the Δdlp which were named Δdlp C1 and Δdlp C2 were utilised for study.

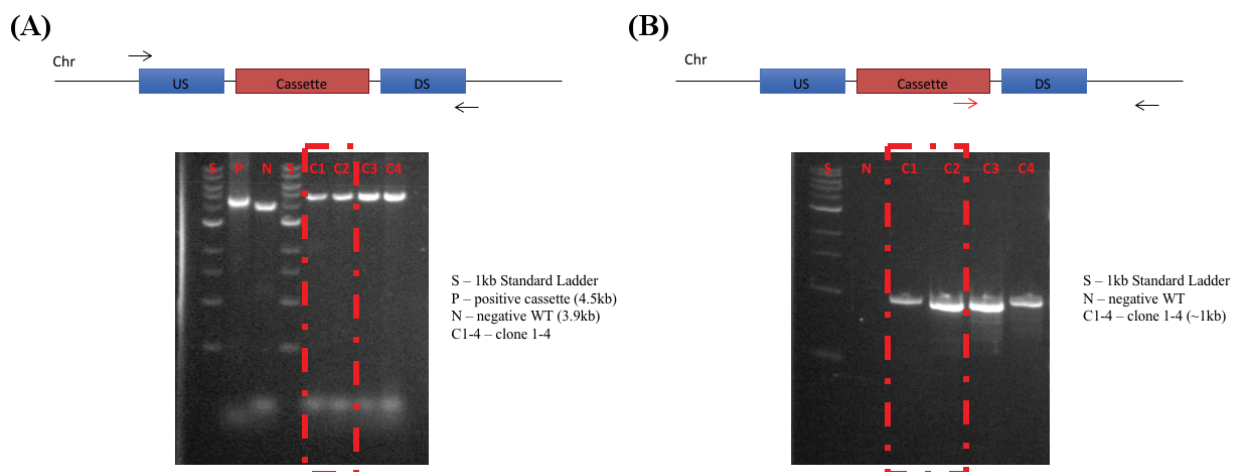


Figure 5.3 Confirmation of Cassette Insertion into B128.7.13

The insertion of the kan-sacB cassette was confirmed using PCR with primers designed (above corresponding gels) in the B128.7.13 genome to ensure the correct location. (a) 1% agarose gel showing the PCR products of correct insertion (P), WT (N) and transformation clones (C1-C4). (b) 1% agarose gel showing the PCR products of WT (N) and transformation clones (C1-C4) confirming the orientation of the cassette in the genome using the primers marked above. 1 Kb standards used (Bio-Rad). Clone 1 and clone 2 were selected for WGS and further phenotype work (highlighted in red box).

5.3.2 CREATION OF A COMPLEMENTED B128.7.13 Δ DLP STRAIN

With the creation of a deletion strain in B128.7.13, a complemented strain was required to confirm any downstream phenotypes that became apparent with the absence of the *dlp* operon. Due to no plasmids being available in the lab able to complement the absent *dlp* operon, a chromosomal complemented strain was created for each clone. Each biological clone, Δ *dlp* C1 and Δ *dlp* C2, was complemented using the WT *dlp* operon, once again using natural recombination. Instead of using a plasmid for this recombination as previously used in the deletion transformation (*Figure 5.2*), a single sequenced PCR product of the WT B128.7.13 *dlp* operon was utilised (data not shown). The same natural transformation was performed and clones that lost the cassette were selected for using 20% sucrose; which is toxic to Δ *dlp* C1 and Δ *dlp* C2 due to the incorporation and expression of the *sacB* gene in the cassette.

As demonstrated in *Figure 5.4*, PCR checks were used to confirm the loss of the cassette. The WT size of the operon was measured using the flanking primers and the loss of the cassette tested by using a cassette specific primer. The expected size of the WT operon of 3.9 kb was seen for both $\Delta::$ *dlp* C1 (*Figure 5.4 Lane 3*) and $\Delta::$ *dlp* C2 (*Figure 5.4 Lane 2*). The loss of the cassette was clear with no product produced when this primer was used for PCR (*Figure 5.4 Lane 1*). These specific checks suggest that the cassette was lost and that both clones of the B128.7.13 Δ *dlp* were successfully complemented with the WT *dlp* operon (B128.7.13 $\Delta::$ *dlp*).

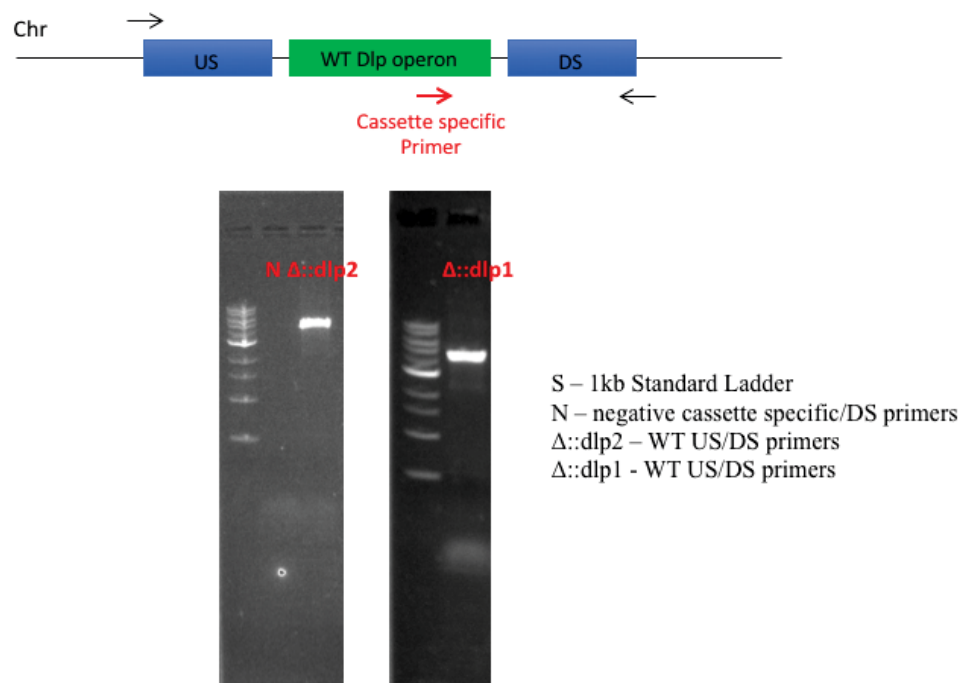


Figure 5.4 Confirmation of the Complementation of B128.7.13Δdlp

Loss of the kan-sacB cassette was confirmed using PCR with primers both for the WT operon and the cassette. A 1% agarose gel clearly showed that for clone 1 and clone 2, the deletion had been complemented with the WT dlp operon, with a product seen at the correct size of (3.9kb). The loss of the cassette was evident with the absence of a product seen in the reaction that used a cassette specific primer (N). 1 Kb standards were used (Bio-Rad).

5.3.3 CONFIRMATION OF B128.7.13 Δ DLP AND B128.7.13 Δ ::DLP STRAIN SEQUENCES

H. pylori is well noted to have a dynamic, unstable genome (229). Therefore, following on from the PCR checks, whole genome sequencing (WGS) using MiSeq Next generation sequencing technology was performed by Genome Quebec (Canada) on all strains that were used in the phenotype testing. These included B128.7.13 WT, Δ *dlp* C1 and Δ *dlp* C2 and Δ ::*dlp* C2. The B128.7.13 WT strain was sequenced as this strain remains in scaffold form in the NCBI database and it was necessary to align the sequences against the ‘in house’ strain. The WGS was performed to ensure no other major changes had occurred to the genome throughout the recombination as much difficulty was faced in inserting the cassette into the genome and was successful in only one strain. This would support that any phenotype that was evident in the *dlp* knock out strains, and was able to be complemented, was due to the *dlp* operon alone.

The B128.7.13 genome is not published as a closed genome, but rather as scaffolds and therefore, to be able to assemble the WGS the reference/back bone sequence from the scaffolds also had to be assembled; Mauve was used for this purpose (223). The genome was then annotated using the program Prokka (224) which relies on external feature prediction tools to identify the coordinates of genomic features within contigs. Prokka carries out its search in a hierarchical manner, starting with a smaller trustworthy database, moving to medium-sized but domain-specific databases, and finally to curated models of protein families. By default, an *e*-value threshold of 10^{-6} was used with the following series of included databases: Blast and BlastP (230), UniProt (231), RefSeq followed by a series of Hidden Markov model profile databases including Pfam (232) and TIGRFAMs (233) and finally if no match was found it was labelled as a hypothetical protein (*see S5 for full annotation*). The paired end reads from the MiSeq of our WT B128.7.13, B128.7.13 Δ *dlp* C1 and B128.7.13 Δ *dlp* C2 along with B128.7.13 Δ ::*dlp* C2 were aligned using Bowtie2.0 (225) with default settings and the output alignment visualised using Artemis (226) and BRIG (227).

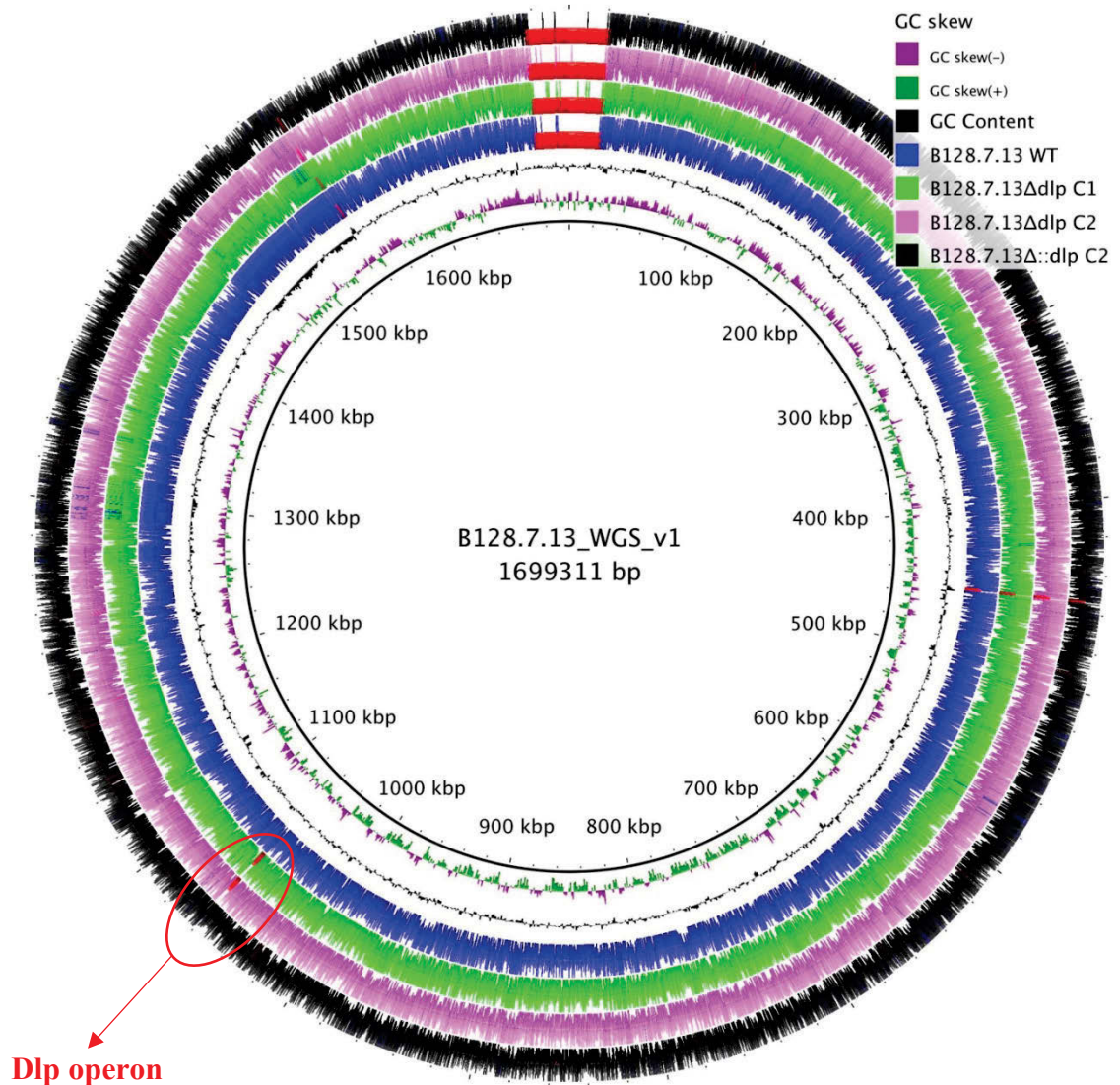


Figure 5.5 Whole Genome Alignments of B128.7.13, Δ dlp and $\Delta::$ dlp Strains

All MiSeq reads were mapped onto a reference genome using Bowtie2 and visualised using BRIG. B128.7.13 WT (blue ring), Δ dlp C1 (green ring), Δ dlp C2 (pink ring) and $\Delta::$ dlp C2 (black ring) were mapped onto the reference strain and visualised as a circularised genome despite them not being closed, indicated at the top of the circle (red). The inner most ring maps the GC skew of the genomes and the second most inner ring maps the GC content (~37%). Of interest to us was the deletion of the dlp operon (labelled) which is evident in the Δ dlp C1 (green) and Δ dlp C2 (pink) sequences only.

High coverage was achieved using the short run sequencing technology (*Figure 5.5*), with the closure of the genome not being possible as the backbone had not been closed previously (*Figure 5.5* top red region). There are a number of slight differences that can be seen in all the genomes compared to the reference backbone used in this analysis, but due to their conservation in our strains, they are considered not to be due to the transformation procedure. Furthermore, FlagStat was used to analyse this coverage further and matches of the reads to the reference sequence of the assembled B128.7.13 scaffolds. From this analysis, all strains achieved a high coverage above 90% across the genome with a high percentage of these reads being paired (Table 5.2).

Table 5.2 Summary of FlagStat Analysis of the Alignment of B128.7.13 MiSeq Reads

	Mapped Reads	Properly Paired
B128.7.13 WT	95.72%	88.32%
<i>Δdlp C1</i>	91.83%	79.79%
<i>Δdlp C2</i>	91.46%	83.83%
<i>Δ::dlp C2</i>	95.04%	87.18%

The *dlp* operon (*Figure 5.5* black arrow) has not previously been annotated in this genome and therefore a multiple sequence alignment was performed with several other *H. pylori* dynamins we have identified in this study. This was performed to ensure these ‘hypothetical’ genes within the B128.7.13 draft genome were in fact the *dlp* genes of interest (bottom panel of *Figure 5.6*). Key conserved residues of the GTPase domain, in particular the G1 P loop and the predicted transmembrane domain, are highlighted. The presence of these features and the high sequence homology along the length of the protein to other *H. pylori* dynamins, gave confidence to the identification of these genes as DLPs.

Quite remarkably these genes are structured differently to other *H. pylori dlp* genes, with three opening reading frames seemingly encoding *dlp2* (noted in *Figure 5.6*). The reason behind this is not fully understood. Interestingly, the sequencing suggests that these separate

open reading frames encoding *dlp2* are also present in our WT strain (data not shown). This is highly unique and has not been described in any other *H. pylori* genome or the wider dynamin superfamily.

The upstream *dlp* gene *dlp1*, shows the conserved split that has been evident in most other *H. pylori* dynamins, with the previously noted exception of *dlp1* in 26695 (Section 3.3.2). The split point varies once again in this gene, but the conservation of the transmembrane position towards the end of the longer *dlp1b* gene is maintained. As seen in Figure 5.6, deletion of the *dlp* operon in B128.7.13 was successful. No- reads were detected in this region indicated by the drop in the consensus lines for the deletion clone genomes (red and green) to zero. It is evident that the deletion is not complete however, with over hanging regions of the *dlp* operon still present. However, the predicted GTPase domains of both proteins and the TM domain of *dlp1* have all been removed making the operon mostly inactive. A full complementation was successful, with the restoration of the operon to WT coverage (black line).

To summarise, the transformation of *H. pylori* was successful for one strain, resulting in a single construct. This opens up the possibility for study of *in vivo* phenotypes that arise from the deletion of the *dlp* genes in a *H. pylori* strain e.g. in future mouse investigations. Whole genome sequencing was performed to confirm successful knockout of the *dlp* genes, as well as complementation of the WT *dlp* operon.

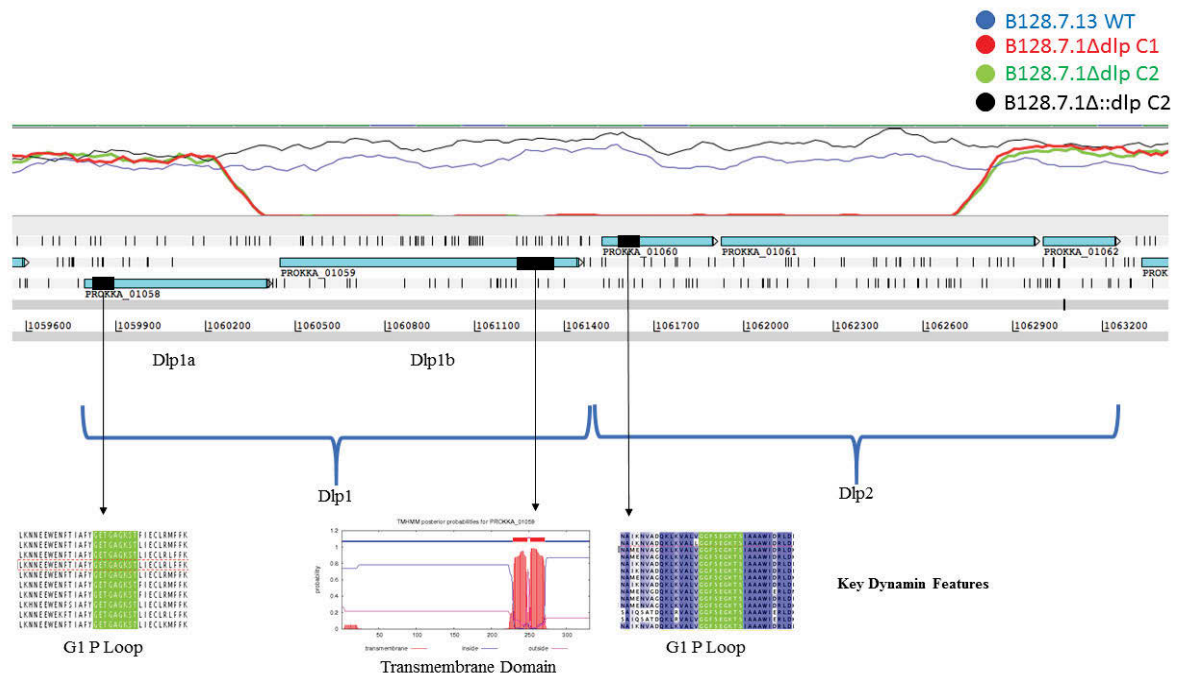


Figure 5.6 Dynamain-like Protein Operon in B128.7.13 WT Δ dlp and Δ ::dlp Strains

The assembled B128.7.13 WT genome was annotated using PROKKA. Key dynamin features (lower panel) have been annotated (black), with the GTPase domain G1 P loop of both Dlp1 and Dlp2, which align (green) with homologs of DLPs in a range of *H. pylori* strains and the transmembrane domain (TM) predicted by TMHMM. The assembled Miseq reads for Δ dlp C1 (red line), Δ dlp C2 (green line) and Δ ::dlp (black line) strains were aligned to B128.7.13 WT (blue line) and visualised in Artemis. Complete coverage of the dlp operon was achieved in the WT (blue) and Δ ::dlp (black). In the Δ dlp C1 (red), Δ dlp C2 (green) there was a clean deletion, with no reads in this region of the genomes of both strains indicated by the drop in the line. The deletion is not complete, with overhanging regions of the operon still present but any protein products produced would not be functional with all key functional domains not present.

5.3.4 Δ DLP CAUSES AN INCREASE IN BLOOD AGAR HAEMOLYSIS

Routinely *H. pylori* is grown on blood agar plates, composed of a blood agar base, along with 8% defibrillated horse blood (DHB) and a selective antibiotic supplement (DENT). Upon growth, *H. pylori* can exhibit signs of haemolysis, whereby the bacteria break down the red blood cells in the media (234). Interestingly, the B128.7.13 Δ dlp strain showed earlier signs of this haemolysis, within 24 h of repeated standardised growth, when compared to the WT and complemented strains which took 48-72 h (*Figure 5.7*). A halo of clearance was seen on the B128.7.13 Δ dlp agar plate (*Figure 5.7* circled in black) after 24 h of growth indicating the release of an intracellular haemolysin into the media by *H. pylori*. When this is compared to the inoculation sites of the WT and complemented strains, minimal to no haemolysis is seen (*Figure 5.7*). The haemolysis demonstrated by this phenotype is typical of beta haemolysis (234).

From this experiment alone, it is unclear whether the release of the haemolysin is a direct effect of the deletion of *dlp* or whether it is an indirect effect; for example, if the cells had become more permeable or “leaky” due to the deletion.

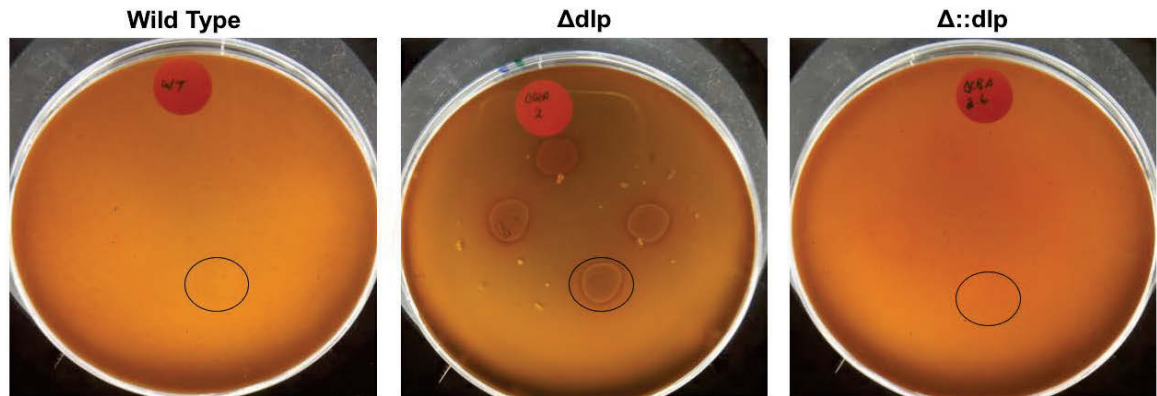


Figure 5.7 Haemolysis of Blood Agar in the Presence of Δ dlp Strain

Normalised cultures of WT, Δ dlp and $\Delta::$ dlp were spotted (circle) in replicate (four times) onto nonselective blood agar plates and incubated for 24 h in a microaerophilic environment and images were taken. A halo of clearance was evident around the four inoculated spots of Δ dlp and absent from the WT plate. Early signs of clearance were beginning to form on the $\Delta::$ dlp plate. This halo of clearance is caused by the breakdown of the red blood cells in the media, known as haemolysis.

5.3.5 HAEMOLYSIS IS DUE TO A MEMBRANE DEFECT IN Δ DLP *H. PYLORI*

After noting the increased haemolysis caused by the B128.7.13 Δ dlp strain, the cells were examined under the microscope to observe any evidence of membrane defects, given that bacterial dynamins have been implicated in the maintenance of membrane integrity previously (100). To visualise membrane integrity, an indirect approach was taken at first, whereby the membrane was stained with FM4-64, a lipophilic dye that stains membranes and appears red (excitation/emission maxima ~515/640 nm). To study membrane permeability, the cells were also stained with Sytox Green (excitation/emission maxima ~504/523 nm), a nucleic acid dye, that is impermeant to live cells, making it a useful indicator of dead or damaged cells within a population. B128.7.13 WT, Δ dlp C1 and Δ dlp C2 strains were grown to mid log (OD_{600nm} 0.4) centrifuged gently and resuspended both in FM4-64 and Sytox Green and incubated at RT for 30 min and visualised on a Zeiss microscope using the 100X objective.

There was an evident difference between B128.7.13 WT and the Δ dlp C1 and Δ dlp C2 strains; with minimal Sytox staining seen in the WT (*Figure 5.8 Panel 1*), compared with both Δ dlp C1 and Δ dlp C2 strains (*Figure 5.8 Panel 2 and 3*, respectively). WT cells had an intact membrane with the exclusion of the Sytox Green dye with only background staining (*Figure 5.8 Panel 1C*). When compared to Δ dlp C1, (*Figure 5.8 Panel 2C*) which exhibited green fluorescence and when overlaid with the membrane staining seen in Panel 2B, this corresponded directly with the *H. pylori* cells (*Figure 5.8 Panel 2D*). There were also several events where the Sytox dye has stained the DNA outside the cell, indicating leakage of the DNA from the cell possibly due to a compromised membrane. The same phenotype was confirmed in Δ dlp C2 (*Figure 5.8 Panel 3*), where the overlay showed clear DNA staining in cells indicating a compromised membrane (*Figure 5.8 Panel 3D*)

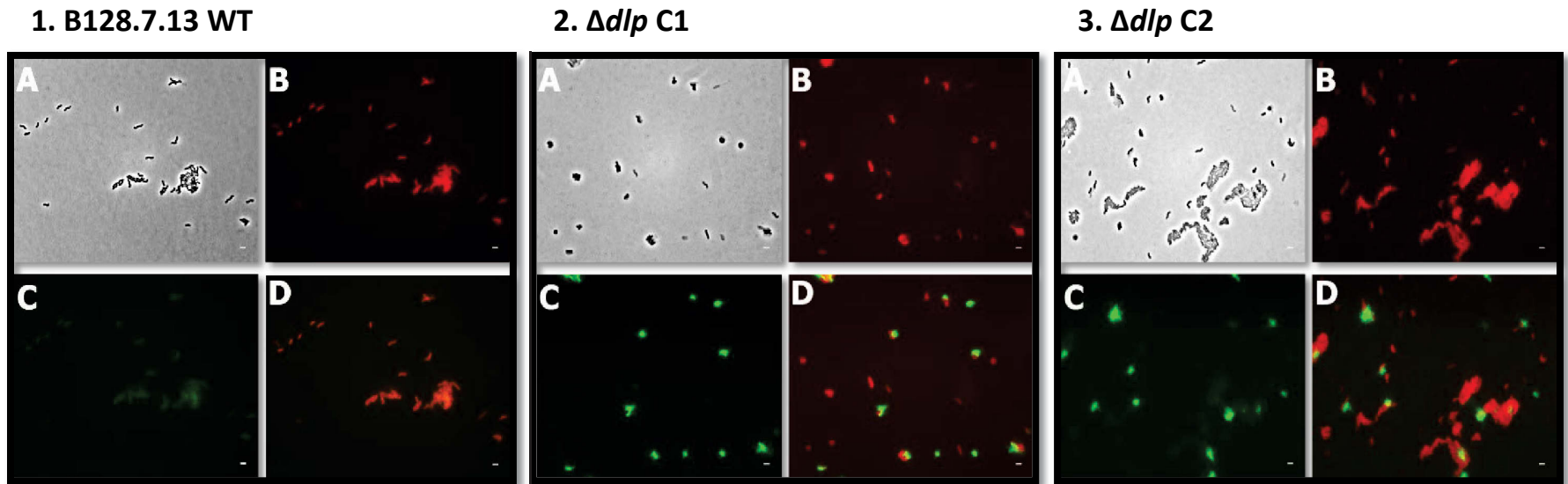


Figure 5.8 Δ dlp Shows Evidence of a Fragile Membrane

The WT, Δ dlp C1 and Δ dlp C2 strains were grown to mid log (OD_{600nm} 0.4), centrifuged, resuspended in media with FM4-64 and Sytox green fluorescent dyes, mounted onto 1% PBS agarose pads and visualised using the Zeiss microscope. FM4-64 stained the membrane (red, (B)) and sytox green (green, (C)) is an impermeable dye, excluded from the cell unless the membrane is compromised in some way and then it is able to pass through and bind to the DNA (green). Minimal permeability is evident in the WT strain (panel 1 (D)) where no sytox staining is seen in the *H. pylori* cells. In the absence of dlp (Panel 2D and Panel 3D) there is a clear increase in membrane permeability. Both biological clones of the Δ dlp show evidence of a fragile membrane, with several cells in the field of view showing a stained membrane, but the DNA is also stained inside and out of the cell. Scale bar is $2\mu m$.

With visual differences clearly evident between the WT, Δdlp C1 and Δdlp C2 strains, these events were manually counted. This was necessary due to clumping of the *H. pylori* cells, particularly the WT (*Figure 5.9* Bottom Panel A) and $\Delta::dlp$ C2 (*Figure 5.9* Bottom Panel D) strains which made automated counting difficult. Approximately 300 cells were counted for each strain across 2-3 replicate cultures, using a 100x objective on the Delta Vision Elite inverted research microscope using the filter sets of FITC (500 ms) and TRITC (300 ms) for Sytox Green (nucleic acids) and FM4-64 (membrane) staining respectively.

Upon counting these events it was found that of the 300 cells observed for B128.7.13 WT, 13% showed evidence of a compromised membrane with green fluorescence within or directly adjacent to the cell (*Figure 5.9* Top Panel). As seen in the representative image of the WT cells minimal green fluorescence was witnessed, and the cells showed increased levels of clumping after the centrifugation step, though membrane integrity seemed to remain intact (*Figure 5.9* Bottom Panel A). Evidence of membrane damage was pronounced in both Δdlp C1 and Δdlp C2 strains, with 67% and 51% of the cells demonstrating membrane damage respectively (*Figure 5.9* Top Panel). Further support of this is seen in the representative images of both Δdlp C1 and Δdlp C2 strains in *Figure 5.9* in panels B and C respectively. Green fluorescence was seen in most cells, with spotted staining of the DNA in and around the cells. Quite interestingly, decreased clumping was witnessed in both these strains. The membrane defect was completely restored in the $\Delta::dlp$ C2 strain where only 9% of cells observed showed evidence of a compromised membrane (*Figure 5.9* Top Panel). As observed previously, these cells exhibited little evidence of Sytox staining along with the same level of clumping as WT, seen in the representative image in *Figure 5.9* Panel D (*See S6 for separate channel images of all strains*).

From this analysis it is clear that in the absence of dynamins the integrity of the *H. pylori* membrane is compromised, and that this integrity can be restored upon complementation with the WT *dlp* operon.

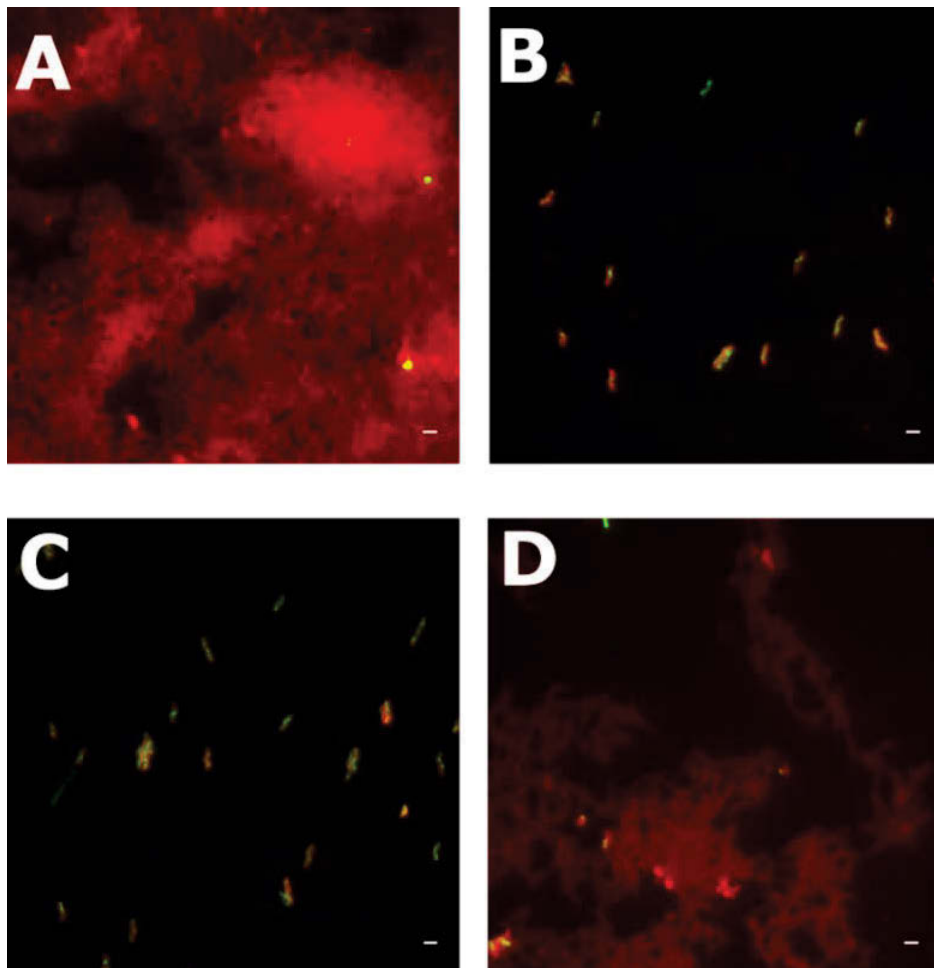
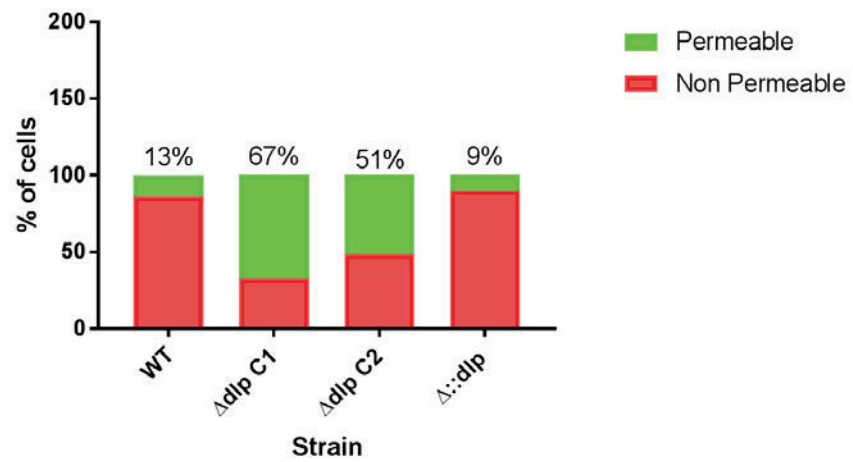


Figure 5.9 Quantification of Δ dlp Membrane Permeability

Cells were quantified for their membrane permeability over 2-3 culture replicates. Approximately 300 cells were manually counted for WT (A), Δ dlp C1 (B), Δ dlp C2 (C) and

$\Delta::dlp$ (D) and representative images are seen below analysed by ImageJ. Cells were scored to have membrane permeability (green) or no membrane permeability (red) based on the presence of green fluorescence in or in the direct surrounding environment. The total permeability (green) was transformed into a percentage of the number of cells counted (above corresponding column) and graphed using GraphPad. WT showed 13% of cells counted to have an increase in permeability compared to 67% and 51% for Δdlp C1 and Δdlp C2 respectively. This increase in permeability was able to be rescued in the $\Delta::dlp$ strain where 9% of cells counted showed permeability. Scale bar is 2 μ m.

5.3.6 H. PYLORI ΔDLP IS SENSITIVE TO ACID EXPOSURE

H. pylori is a pathogen of the stomach, where it is exposed to a low pH environment. A key defence against this harsh environment is the presence of a strong, fortified membrane (Chapter I). Therefore, a compromised cell membrane would be hypothesised to render the B128.7.13 Δdlp strain quite vulnerable to the acidic environment of the stomach. As such, the strains, B128.7.13 WT, Δdlp C1 and Δdlp C2, $\Delta::dlp$ C1 and $\Delta::dlp$ C2, were exposed to acid over short exposure times in order to test their viability and survival (methods outlined in Section 5.2.7). All strains were grown to an OD_{600nm} of 0.5 (as indicated by black line in Figure 5.10), centrifuged gently, resuspended in media of varying pH (7.4, 4.9 and 2.1) and left to grow for another h at which another OD_{600nm} measurement was taken. In its natural environment, *H. pylori* scavenges urea from its surroundings to combat acid exposure through the essential urease pathway, therefore 10 mM urea was added to all acidic media to mimic conditions found in the stomach.

Both Δdlp C1 and Δdlp C2 strains were found to have a slight but significant decrease in OD_{600nm} at neutral pH when compared to the WT (p-value < 0.01) and complemented strains (Figure 5.10). This could be due to the Δdlp strains not responding well to the centrifugation and resuspension steps required in this experiment, with a decrease in their OD_{600nm} of 0.5 to approximately 0.35; most likely attributed to their fragile membranes (seen in Section 5.3.5). At a slightly acidic pH of 4.9, there was a drastic decrease in the ability of strains Δdlp C1 and Δdlp C2 to cope with the acid exposure, with another significant drop from 0.5 to

approximately 0.2 in OD_{600nm} (p-value < 0.05). This was despite the presence of 10 mM urea. This decrease was not seen in the WT and complemented strains, which were able to maintain growth at approximately 0.5 OD_{600nm}; the exception being B128.7.13Δ::*dlp* C2 which dropped significantly to approximately 0.4 (p-value < 0.05). Once again, this could possibly be attributable to the resuspension step. At pH 2.1, the inability of Δ*dlp* C1 and Δ*dlp* C2 to combat the acid exposure was more evident with nearly all cells lysing, resulting in a dramatic and once again significant drop in OD_{600nm} to an average of 0.05 (p-value < 0.01). Interestingly, despite the presence of urea, the WT and the complemented strains also dropped in OD_{600nm} to approximately 0.2 (not significant). This drop is to be expected given the acidic shock experienced by the cells and therefore some level of lysis is expected.

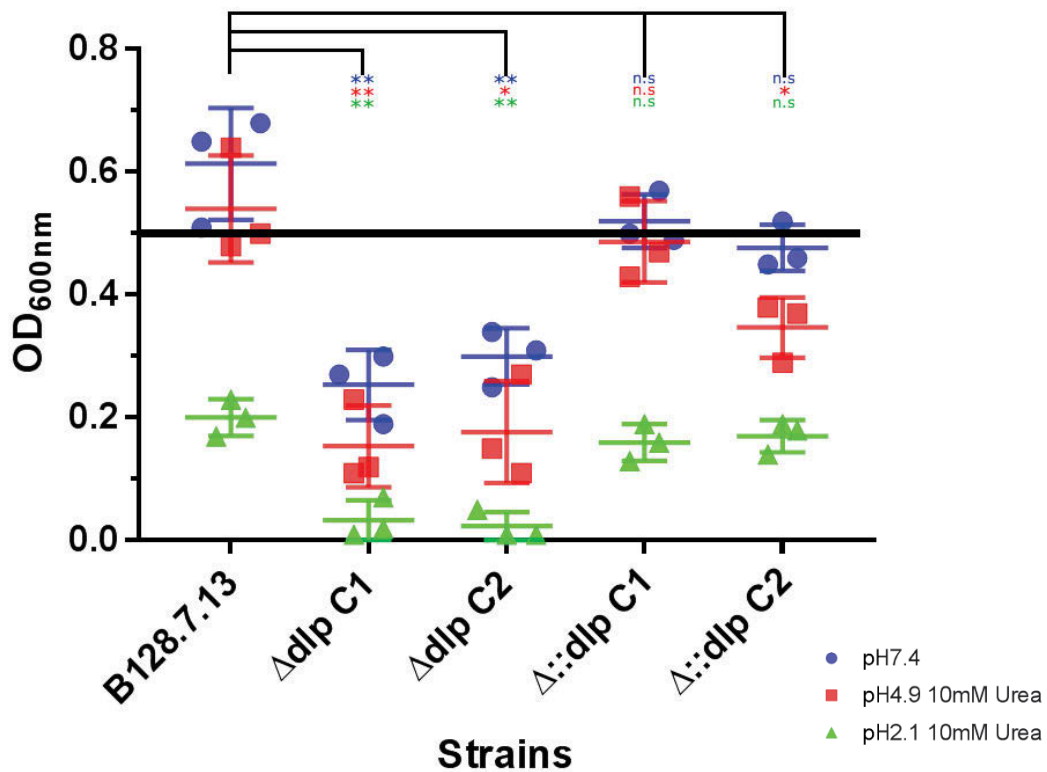


Figure 5.10 Δ dlp is Unable to Overcome Acid Exposure

The B128.7.13, Δ dlp C1, Δ dlp C2, $\Delta::$ dlp C1 and $\Delta::$ dlp C2 strains were grown in a microaerophilic environment at 37°C with shaking to mid log (OD_{600nm} 0.5), centrifuged, resuspended back to an OD_{600nm} of 0.5 in media at varying pH, 7.4 (blue), 4.9 (red) and 2.1 (green). The acidic media (4.9 and 2.1) was supplemented with 10 mM urea. All cultures were allowed to grow for another h under the same conditions in the new media and then final OD_{600nm} measurements were taken (individual dots). The average and the +/- SEM are indicated for each pH tested. B128.7.13 WT maintained and grew in both pH 7.4 and 4.9 to an OD_{600nm} of approximately 0.6 where as both Δ dlp clones significantly dropped to an OD_{600nm} of approximately 0.1-0.3 for pH 4.9 and 7.4 respectively. This was rescued with the complemented strains, which both showing maintained growth at pH 7.4 and 4.9. At the pH of 2.1, all strains showed a drop-in cell density with WT and the complemented strains having OD_{600nm} of 0.2. But the most drastic drop was seen in the Δ dlp clones where the most cell lysis was evident with an OD_{600nm} barely above zero. P-values are indicated by n.s = not significant * < 0.05 and ** < 0.01, coloured coded according to the pH being tested. (N=3)

Surprisingly, an oversupply of urea (100 mM) was able to rescue this phenotype to some degree, with both Δdlp C1 and Δdlp C2 now able to combat this exposure with 10 times the amount of urea present in the media (*Figure 5.11*). At a neutral pH, no strains could successfully overcome resuspension into new media and this is due to the presence of urea. The presence of urea at a neutral pH is toxic to *H. pylori* and has been previously noted in the literature (235). The oversupply of urea to all strains allowed them to combat the acidic stress more strongly than seen in *Figure 5.10*. However, Δdlp C1 and Δdlp C2 were able to maintain cell density with only a slight decrease to an approximate OD_{600nm} of 0.4 for both acidic conditions tested (*Figure 5.11*). WT and $\Delta::dlp$ C1 and $\Delta::dlp$ C2 were also able to maintain their cell density and even grow slightly in the hour exposure with an approximate OD_{600nm} of 0.5 (*Figure 5.11*).

The ability to combat acid exposure is essential for the survival of *H. pylori*, and therefore with Δdlp C1 and Δdlp C2 both not being able to do this while still having a seemingly active urease pathway. It seems like DLPs are playing a key role in protecting the *H. pylori* cell from this stress potentially by ensuring membrane integrity.

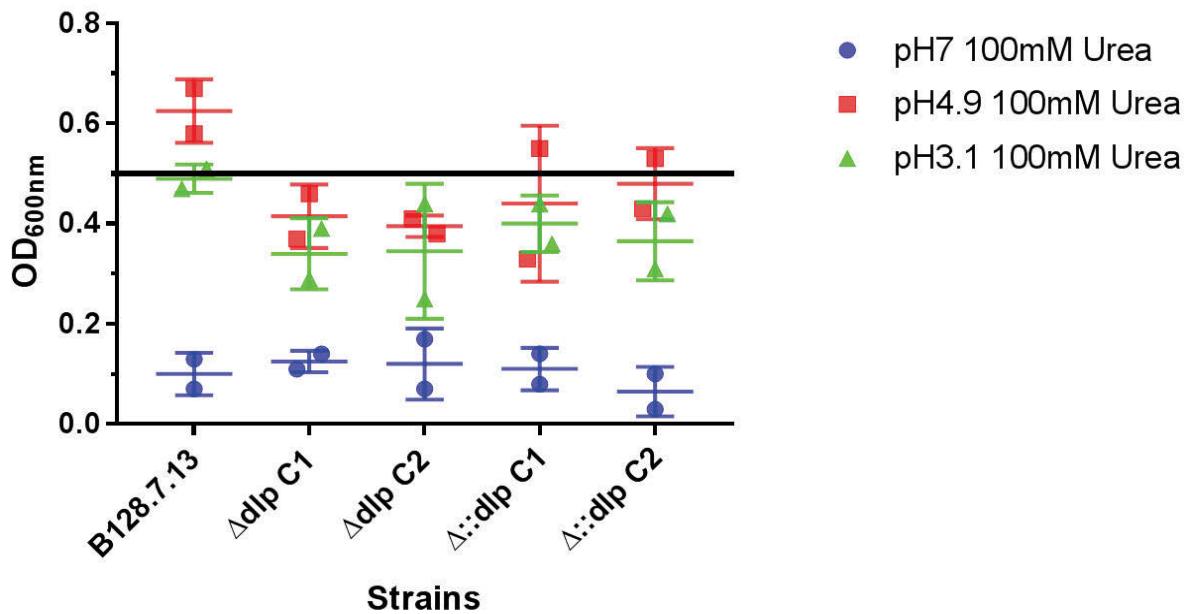


Figure 5.11 Additional Urea Can Rescue Δ dlp in Acid

The B128.7.13 WT, Δ dlp C1, Δ dlp C2, $\Delta::$ dlp C1 and $\Delta::$ dlp C2 strains were grown in a microaerophilic environment at 37°C with shaking to mid log (OD_{600nm} 0.5), centrifuged, resuspended back to OD_{600nm} 0.5 in media at varying pH, 7.4 (blue), 4.9 (red) and 2.1 (green). All new media was supplemented with 100 mM urea. All cultures were then allowed to grow under the same conditions for another h in the new media and then final OD_{600nm} measurements were taken (individual dots). The average and the +/- SEM are indicated for each pH tested. All strains showed a drop-in cell density at pH 7 and this can be attributed to the presence of 100 mM urea in the media. At both pH 4.9 and 3.1, the extra supplementation of urea rescued both the Δ dlp strains with an OD_{600nm} of approximately 0.4 recorded, very comparable to both WT and complemented strain OD_{600nm} after the 1 h acid exposure. (N=2)

5.3.7 *H. PYLORI* Δ DLP HAS A MINIMAL GROWTH DEFECT IN NEUTRAL CONDITIONS

With microscopy confirming reduced membrane integrity (Figure 5.9), the growth for each strain was subsequently examined, with it hypothesised that membrane defects would hinder the growth rate of *H. pylori*. To address this a growth curve was performed over the course of 48 h in duplicate, where OD_{600nm} measurements were taken at regular intervals. It was found that B128.7.13 Δ dlp showed a slight growth defect early in the lag phase of growth, but once an OD_{600nm} of approximately 0.1 was achieved, growth continued in line with the WT and both Δ ::dlp C1 and Δ ::dlp C2 (Figure 5.12). As this growth in normal conditions showed no major differences, the effect of Δ dlp on *H. pylori* growth was examined in the presence of acid, an environment which is more representative of that in the stomach.

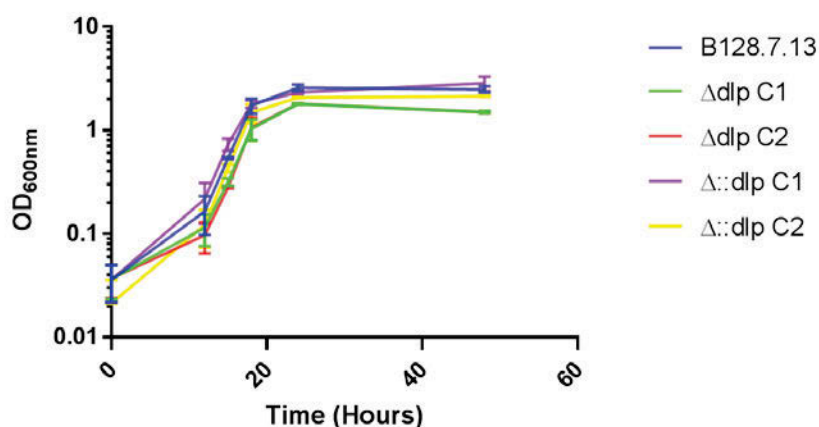


Figure 5.12 Δ dlp Shows No Growth Defect in Neutral pH Conditions

The WT (blue), Δ dlp C1 (green), Δ dlp C2 (red), Δ ::dlp C1 (purple) and Δ ::dlp C2 (yellow) strains were grown from an OD_{600nm} between 0.02-0.05 for 48 h in a microaerophilic environment at 37°C with shaking. A number of time points were selected for OD_{600nm} measurements (12, 15, 18, 24 and 48 h). The average and the +/- SEM are indicated for each strain. There is a slight growth defect of both Δ dlp C1 and Δ dlp C2 seen early in the growth phases but once mid log and stationary phase is reached, the growth defect is no longer seen. WT and Δ ::dlp C1, Δ ::dlp C2 all showed comparable growth throughout the growth curve (N=3).

5.3.8 *H. PYLORI Δ DLP IS UNABLE TO GROW IN AN ACIDIC ENVIRONMENT*

The ability of *H. pylori* to neutralise the acid in its surrounding environment and to continue to grow is a key feature of its fundamental biology. To test whether membrane defects such as that witnessed after dynamin deletion, was enough to affect the growth of *H. pylori* in acid, a one-point growth test was performed (methods outlined in *Section 5.2.8*). Here, WT, Δ *dlp* C1, Δ *dlp* C2, $\Delta::dlp$ C1 and $\Delta::dlp$ C2 strains were passaged four times over the course of 4 days, diluted to a starting OD of 0.05 in three different pHs, 7.4, 4.9 and 2.1 and left to grow for 70 h. A single OD_{600nm} measurement was taken at the endpoint, of which the results are seen in *Figure 5.13*. Again, 10 mM urea was supplemented into the acidic media to allow the strains to utilise the urease pathway.

A significant growth defect was once again witnessed in the Δ *dlp* strain with the Δ *dlp* C1 and Δ *dlp* C2 strains showing a lag of approximately one generation time, reaching an OD_{600nm} of only 2 compared to the WT strain which reached approximately 4 (p-value < 0.05) (*Figure 5.13*). The slight difference between this and the growth curve in *Figure 5.12* could be due to the extended period for which these cultures were left (an extra 30 h) and therefore greater time over which cell lysis could occur. At pH 4.9 and supplemented with 10mM urea, WT and complemented strains were able to achieve normal growth despite the acidic challenge reaching a respectable OD_{600nm} of approximately 2.5 (*Figure 5.13*). In contrast, Δ *dlp* C1 and Δ *dlp* C2 could not overcome this growth challenge as effectively and achieved an OD_{600nm} of approximately 1 after 70 h (p-value < 0.01). At pH 2.1, the WT and complemented strains only achieved an OD_{600nm} of approximately 1 after 70 h and Δ *dlp* C1 and Δ *dlp* C2 strains showed minimal signs of growth with an OD_{600nm} not much above zero (p-value < 0.001). This was despite supplementation with 10 mM urea. This result strongly indicates that Δ *dlp* C1 and Δ *dlp* C2 are unable to overcome and thrive in this harsh environment due to cellular changes caused by a lack of the *dlp* operon.

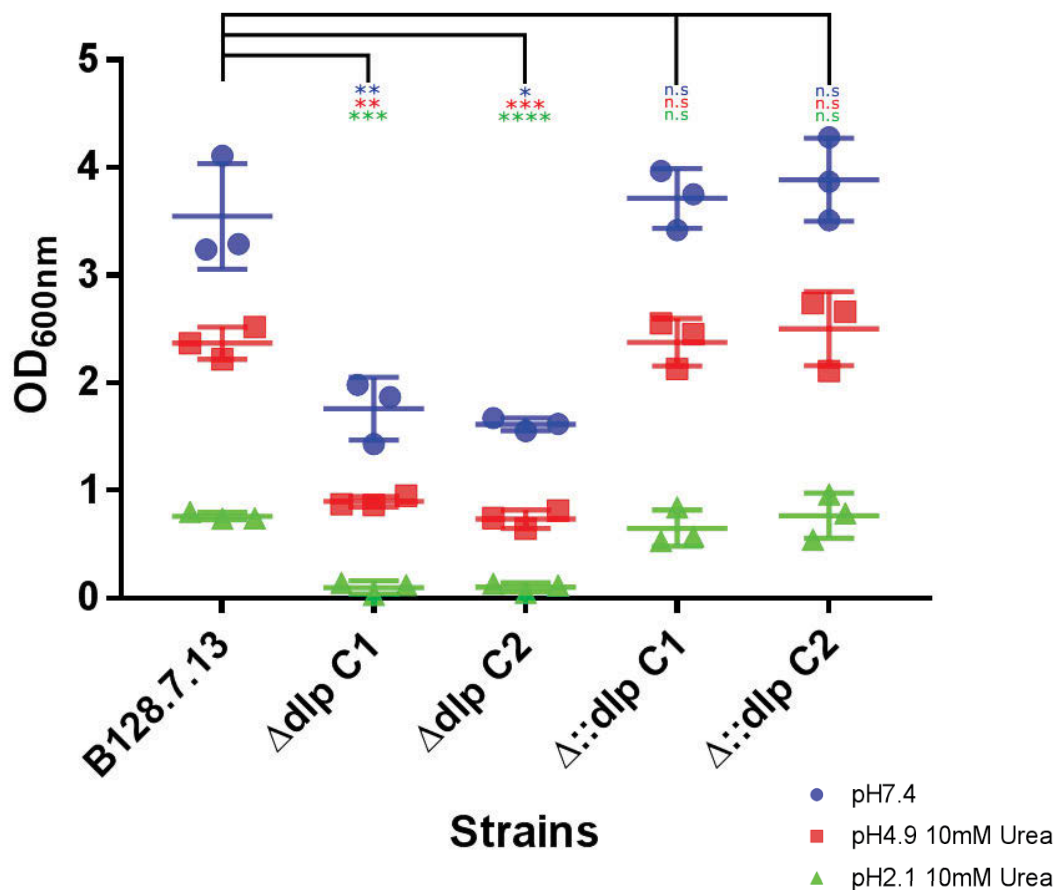


Figure 5.13 Δ dlp is Unable to Grow in Acidic pH Conditions

The B128.7.13 WT, Δ dlp C1, Δ dlp C2, $\Delta::$ dlp C1 and $\Delta::$ dlp C2 strains were grown from an OD_{600nm} between 0.02-0.05 for 70 h in a microaerophilic environment at 37°C with shaking. Each strain was grown in the three pHs being tested, 7.4 (blue), 4.9 (red) and 2.1 (green). All acidic media was supplemented with 10 mM urea. A single end time point (70 h) was selected for an OD_{600nm} measurement. The average and the $\pm$ SEM are indicated for each pH tested. B128.7.13 WT was able to achieve an OD_{600nm} of approximately 4, 2.5 and 1 in pH 7.4, 4.9 and 2.1 respectively. Both Δ dlp C1 and Δ dlp C2 showed a significantly reduced growth at pH 7.4 which was further seen in the acidic conditions with minimal to no growth detectable in pH 2.1. This reduced growth was able to be rescued in the complemented strains ($\Delta::$ dlp C1 and $\Delta::$ dlp C2) with growth seen back to WT levels. P-values are indicated by n.s = not significant, * < 0.05, ** < 0.01, *** < 0.001 and **** < 0.0001, coloured coded according to the pH being tested. (N=3)

When oversupplied with 100 mM urea, this growth defect was rescued (*Figure 5.14*). Both Δdlp C1 and Δdlp C2 showed growth even at the lowest pH tested. As seen previously (*Figure 5.12*) growth was inhibited at the neutral pH in the presence of urea in all strains. At a pH of 4.9, both Δdlp C1 and Δdlp C2 showed somewhat of a rescue with OD_{600nm} measurements of approximately 2, back to similar levels of growth compared to the WT and complemented strains which reached approximately 3 (*Figure 5.14*). This strongly indicates that the urease pathway is not affected in the Δdlp strains.

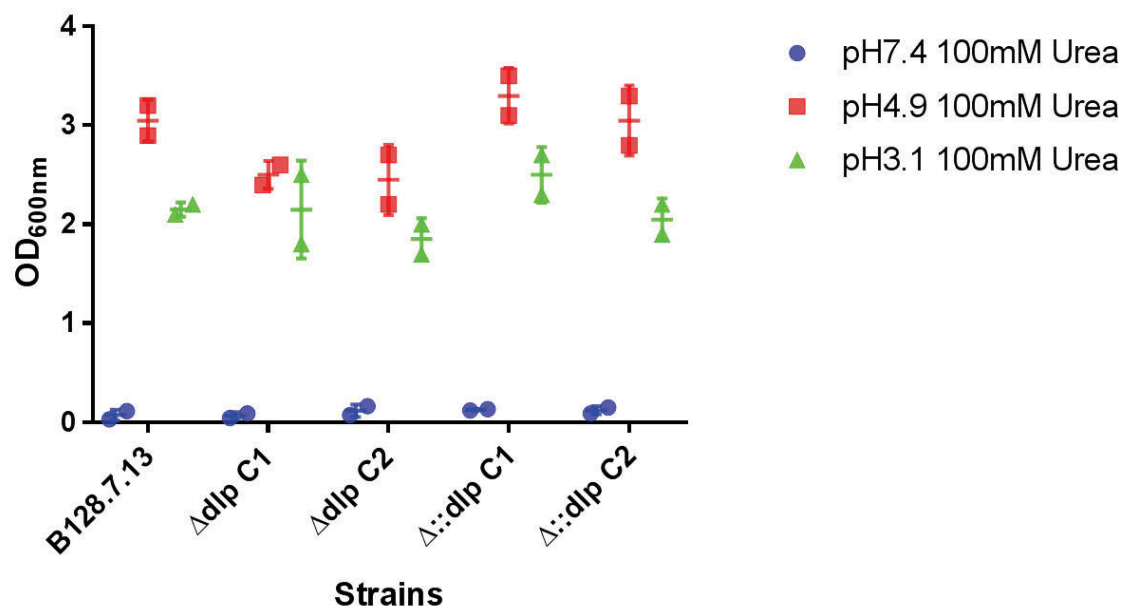


Figure 5.14 Additional Urea Can Rescue Δ dlp Growth in Acid

The B128.7.13 WT, Δ dlp C1, Δ dlp C2, $\Delta::$ dlp C1 and $\Delta::$ dlp C2 strains were grown from an OD_{600nm} between 0.02-0.05 for 70 h in a microaerophilic environment at 37°C with shaking. Each strain was grown in the three pHs tested, 7.4 (blue), 4.9 (red) and 3.1 (green). All media was supplemented with 100 mM urea. A single end time point (70 h) was selected for an OD_{600nm} measurement (individual dots). The average and the spread are indicated for each pH tested. All strains showed a drop-in cell density at pH 7, attributed to the presence of 100 mM urea in the media, seen previously in Figure 5.12. At both pH 4.9 and 3.1, the extra supplementation of urea was able to rescue both the Δ dlp strains with an OD_{600nm} of approximately 2 recorded, very comparable to both WT and complemented strains OD_{600nm}. (N=2)

From this, the one-point growth test was expanded in order to study growth in acid over a time course. All strains were passaged for 4 days and then subcultured into growth media at a pH of 5.2, supplemented with 10 mM urea and diluted to a starting OD_{600nm} of 0.05. Each culture was split into individual time points and OD_{600nm} measurements taken at these selected time points (*Figure 5.15A*). Along with the OD_{600nm} measurements, pH readings were also taken of the supernatant to monitor the neutralisation of the acid in the media over the growth curve for all strains (*Figure 5.15B*).

Consistent with previous findings, Δdlp C1 and Δdlp C2 strains showed slower growth (*Figure 5.15A*). The WT and $\Delta::dlp$ C1 and $\Delta::dlp$ C2 strains exhibited a normal lag, exponential and stationary phase profile when compared with Δdlp C1 and Δdlp C2 strains, which showed a steady growth after 15 h to reach an OD_(600nm) of approximately 1.5 at the 48 hour timepoint. Interestingly, the WT and $\Delta::dlp$ C1 strains both completely neutralised the acid in the media after 12 h with a pH reading of the supernatant being pH 7 (*Figure 5.15B*). The Δdlp C1, Δdlp C2 and $\Delta::dlp$ C2 strains at 12 h had partially neutralised the acid with the supernatant being a pH of 6 and by 15 h they had completely neutralised the acid in the media also (*Figure 5.15B*). The media only control showed no variation in the pH from 5 (*Figure 5.15B*). Despite neutralising the acid in the media, Δdlp C1 and Δdlp C2 were not able to overcome the initial shock of the acid.

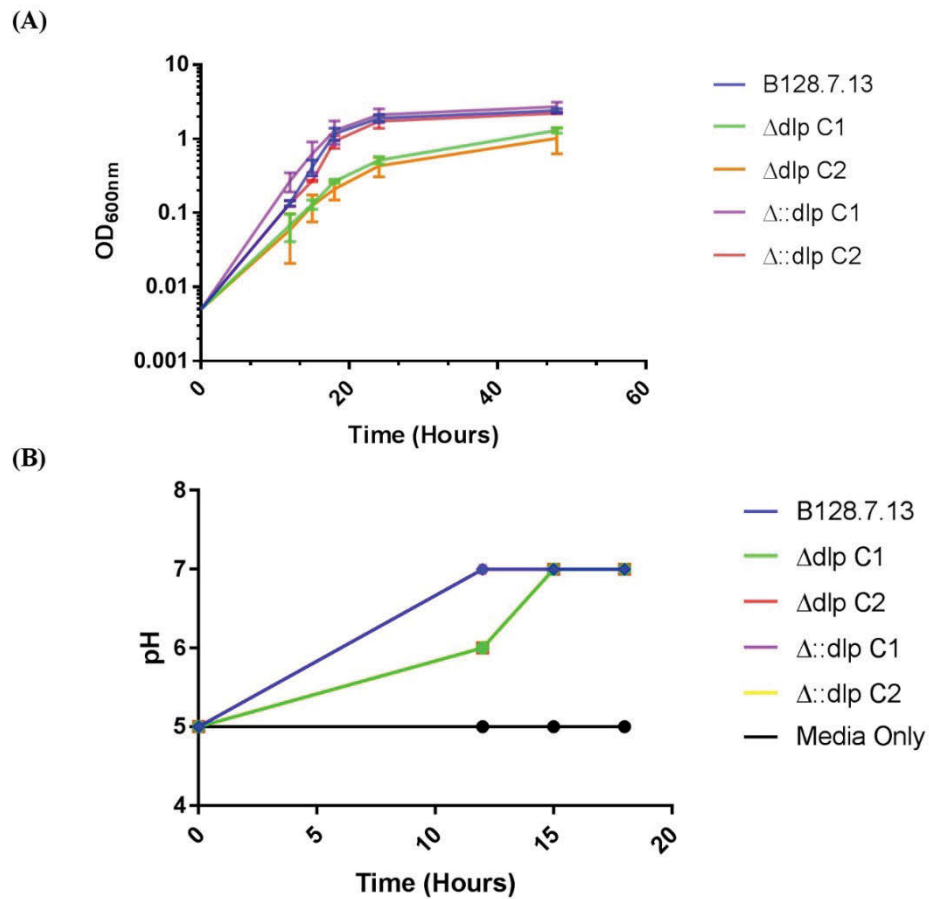


Figure 5.15 Acidic Growth Curve and Corresponding pH Neutralisation

(A) The B128.7.13 WT (blue), Δdlp C1 (green), Δdlp C2 (red), Δ::dlp C1 (purple) and Δ::dlp C2 (yellow) strains were grown from an OD_{600nm} between 0.02-0.05 for 48 h in a microaerophilic environment at 37°C with shaking. A number of time points were selected for OD_{600nm} measurements (12, 15, 18, 24 and 48 h). The average and the +/- SEM are indicated for each strain. It can be seen that there was a moderate growth defect of both Δdlp C1 and Δdlp C2 seen early in the growth phases that continued into the mid log and stationary phases. WT Δ::dlp C1 and Δ::dlp C2 all showed comparable growth throughout the growth curve. (B) Throughout the early time points, pH measurements were taken for each culture. The starting pH was measured to be 5 and by the first time point of 12 h, WT (blue) and Δ::dlp C1 (purple) had neutralised the acid to pH 7. Δdlp C1 (green), Δdlp C2 (red) and Δ::dlp C2 (yellow) were slightly slower and took 15 h to neutralise the acid. (Line colours may be hidden due to overlay of the strains) (N=2)

Overall, it appears that Δdlp results in a compromised membrane inhibiting *H. pylori* from overcoming acidic conditions following both short and long-term exposure. However, the ability to rescue Δdlp with excess amounts of urea emphasises that the urease pathway is not affected by Δdlp . Rather, it may be the defect in the membrane seen using microscopy, independent of urease that is subsequently making the strains more susceptible to the acid. This demonstrates how important the membrane is to *H. pylori* to overcome this natural environmental stress.

5.4 DISCUSSION

An essential aspect of enhancing our knowledge of bacteria is the development of functional studies of proteins with unknown function. It has been stated that approximately 40% of all proteins in *H. pylori* that have little to no known function (236), and to begin to investigate their function could lead to new breakthroughs in understanding and producing therapeutic treatments for this important human pathogen. One such group of proteins in *H. pylori* are the DLPs. Generally in bacteria, despite being widely spread throughout the domain, DLPs are understudied and their functions are largely unknown. Studying the role of these hypothetical dynamin proteins may aid in developing new therapeutic strategies to target *H. pylori*, given the fundamental role dynamins play in eukaryotes and increasing evidence for a similar role in bacteria.

To investigate the function of the dynamins in *H. pylori*, a knockout strain was required. *H. pylori* is a naturally competent bacterium that is able to take up DNA from its environment without any pre-treatment and incorporate it into its genome through recombination events (237). The work presented in this chapter commenced with a simple mutagenesis procedure designed specifically for gene replacement in *H. pylori*, where a double cross over recombination event of the plasmid DNA containing a cassette with the flanking regions of interest for the deletion of the *dlp* operon (*Figure 5.1*). The design of the cassette allows for the procedure to incorporate both a positive selection step, in the presence of Kanamycin, and a negative selection step, utilizing the counter selectable *sacB* gene of *B. subtilis* under the *flaA* promoter from *H. pylori* to replace any gene in a markerless fashion (220). Unfortunately, despite many efforts using both natural transformation and electroporation methods, not all constructs required for the investigation of the *dlp* operon in *H. pylori* were successfully developed in this study. Several attempts were made to construct the strains under various conditions including; varying the DNA concentration, time of incubation with the DNA and amount of kanamycin used in the selection step. It has been previously seen that by linearising and ligating the DNA of interest, transformation efficiency can be greatly improved in naturally competent bacteria like *B. subtilis* (238). Implementing this strategy in the transformation of *H. pylori* didn't drastically improve the transformation efficiency,

with similar number of clones seen in the transformation for the complementation of the *dlp* operon into the B128.7.13 Δ *dlp* strains and it was not successful in the creation of the other *H. pylori* Δ *dlp* strains. There are potentially several reasons why the transformation was not successful in the strains 26695 and SS1 but a possible reason would be that the deletion of dynamin in these strains is lethal. This is supported by the extremely low transformation efficiency in B128.7.13 with only 4 clones and the radically improved transformation efficiency when complementing the operon back to the chromosome. The successful creation of B128.7.13 Δ *dlp* and B128.7.13 $\Delta::dlp$ was however sufficient to develop a fundamental understanding of the function of these proteins in the carcinogenic strain B128.7.13.

The gerbil-adapted *H. pylori* strain B128.7.13 has been previously sequenced and placed in the NCBI database in scaffold form, meaning it has not been assembled as a complete and closed genome (in NCBI ASM17295v1). For the purposes of this study, this parent genome needed to be assembled to ensure all phenotypes that are noted can only be attributed to the deletion of the *dlp* operon. With *H. pylori* being known to significantly and rapidly alter its genome, this was an important confirmation (229). Using Mauve to assemble the scaffolds of the backbone, and the annotation performed by Prokka, it was able to be shown that the deletion of the dynamin operon was successful as was the chromosomal complementation of these strains (*Figure 5.6*). Notably, the structure of the *dlp* operon is predicted to be different to what has been previously seen for *H. pylori* dynamins. Despite, the dynamin operon structure in various bacteria having been previously shown to vary slightly, it is noted in the literature that bacterial dynamins seemingly come in pairs in the genome, usually back-to-back (3-5, 98). Specifically, in most *H. pylori* genomes there is a split in the upstream dynamin gene despite both proteins being assumed to be required for function, which has been noted previously in the ETEC dynamin operon of LeoCBA (5). The split in the upstream dynamin gene, B128.7.13 is no different. The main finding that was extremely interesting was the split in the downstream gene, Dlp2, which had not been noted previously. Despite the split in the gene, all the necessary domains, such as the GTPase domain, needed for a functional dynamin protein were present (*Figure 5.6*). This was the first case noted where Dlp2 has been split in the genome and the obvious question that came to mind was, why split the gene when a full protein product is needed for function? The answer to this question is

still unknown and warrants further investigation using long read technologies such as PacBio. WGS was performed for the sole reason to confirm the deletion of the dynamin operon and to ensure that no major inversions, deletions, or other changes had occurred to the genome. These were all confirmed and although it is beyond the scope of this study, this data could be used to close the B128.7.13 genome and examine other changes that may have occurred such as SNPs.

Haemolysis is a common occurrence in bacterial infection including *H. pylori*, with bacteria targeting the erythrocytes to access important nutrients, such as iron, for survival (239). Gram positive and Gram-negative bacteria all possess the ability to exploit the host in this way, but specifically Gram-negative bacteria have several factors that assist in causing infection to the host and in the case of *H. pylori* this includes but is not limited to the vacuolating cytotoxin (VacA) and cytotoxin-associated gene A (CagA). Haemolysis is the breakdown or lysis of red blood cells with the contents of these cells being released into the extracellular milieu. Haemolysins, sometimes referred to as cytolysins, are bacterial toxins that form a pore in the target host cell (239). *H. pylori* 26695 is known to encode three putative cytolysins in its genome namely, HP1086, HP0499 and HP1490, with only two being properly characterised and studied (240). HP1086 and HP0499 are both TlyA-like proteins that show distinctive homology to a potential virulence factor, *tlyA*, from both *M. tuberculosis* and *Serpulina hyodysenteriae*, that equally display potent haemolytic activity on host cells (241, 242). So, it would seem that haemolytic activity in *H. pylori* could be the result of the action of multiple proteins.

Most haemolysins are pore forming toxins which can lyse a number of cells including, red blood cells, leukocytes and even gastric epithelial cells which could be important for establishing *H. pylori* infection (239, 243). Haemolysins are usually secreted as water-soluble molecules which diffuse to the target membrane where they can bind to specific receptors and undergo polymerisation that facilitates their insertion into the host/ target membrane (239). Interestingly, in *H. pylori* it has been shown that the beta- haemolysis that the bacteria exhibits increases in iron limiting conditions (244). In this study, increased haemolysis was evident in the Δdlp clones after 24 h growth on blood agar when compared to normalised WT

and $\Delta::dlp$ clone growth (Figure 5.7). This could either be attributed to a direct or indirect interaction between dynamin and the putative haemolysins in *H. pylori*. Previously it has been shown that expression of the TlyA-like protein in *H. pylori* causes agglutination of liposome and erythrocyte membranes, followed by the fusion of the membranes and their subsequent disruption and increased permeabilisation (245). This is intriguing because the Δdlp clones show reduced levels of clumping when compared to the WT and complemented strains (Figure 5.9). This may indicate that the DLPs could be involved in membrane fusion needed to cause infection of the host by *H. pylori*. It is possible that DLPs aid in the release of the TlyA-like protein from *H. pylori*, similar to the role LeoA has in the release of LT toxin from ETEC (5, 7). Alternatively, it could be an indirect effect, where the gene deletion causes the membrane to be compromised, similar to what has been seen in *B. subtilis* in the absence of DynA (100). Therefore, these possibilities were further investigated using fluorescent microscopy whereby *H. pylori* cells were stained with FM4-64, a lipophilic dye that stains the membrane, and a Sytox nuclear stain, that is membrane impermeable and will only stain cells (DNA) with compromised membranes. WT cells showed a low level of DNA staining with only approximately 13% of cells displaying any level of permeability, when compared to the Δdlp clones, where a much higher level of permeability was detected at approximately 50-60% (Figure 5.8 and Figure 5.9). This phenotype was able to be complemented with a recovery of membrane integrity back to WT levels of approximately 9% (Figure 5.9). Interestingly, even though the membrane seemed mostly intact, there were points of weakness in the membrane that could be seen; indicating regions where the membrane was compromised, and the stain was able to move through with the DNA seemingly coming out of the cell from this point (Figure 5.8). This phenotype was further exacerbated when the cells were subjected to a concentrating step where by most cells imaged showed some level of membrane permeability (Figure 5.9). Taking these results together it can be postulated that in the absence of dynamin, the *H. pylori* membrane becomes more fragile, especially in times of duress, a function quite reminiscent of DynA. However, as the proteins are homologs of LeoCBA, further investigation is warranted into whether the release of the TlyA-like cytotoxin is impeded in the absence of dynamin (5).

With a compromised or weakened membrane, it could be hypothesised that the Δdlp clones would exhibit some sort of growth defect but interestingly, this was not observed under neutral conditions (*Figure 5.12*). However, this is not the native or “normal” conditions that *H. pylori* encounters within the stomach in order to survive and cause infection. The stomach is a hostile environment that was once considered sterile, a theory that has since been disproven, with a thriving population of diverse bacteria found in all humans. Very few bacterial species can survive in the low pH environment found within the stomach; *H. pylori* however, thrives in it. *H. pylori* has developed multiple mechanisms to deal with the environmental stress of acid, including urease enzymes that actively buffer the direct environment surrounding the bacterium and the ability to sense and move away from low pH areas of the stomach (122). The urease pathway consisting of the *ureAB* operon, encoding the enzymatic subunits of urease, and the *ureIEFGH* operon, encoding the accessory proteins, is the most well studied pathway in *H. pylori* and has been shown to be essential for growth and colonisation of the stomach lumen (246-249). The urease enzyme neutralises the hydrochloric acid (HCl) in direct proximity to the bacterium by converting host derived urea into ammonia and carbon dioxide. The ammonia then reacts with the HCl, neutralising it, and provides a “bubble” of neutral pH directly around the bacterium to survive short exposures. To test the ability of B128.7.13 Δdlp strain to continue to carry out this important acid adaptive function, the Δdlp clones were exposed to varying amounts of acid for 1 h and it was found that they were unable to overcome this stress despite being supplemented with 10 mM urea. This was evident with the reduction in the starting OD_{600nm} from 0.5 to 0-0.2 in the Δdlp clones whereas the WT and complemented strains could maintain or even expand the cell population during the exposure (*Figure 5.10*). The inability of the Δdlp strains to combat this short acid exposure could indicate a disruption to the urease pathway, causing an inability to neutralise the HCl in the media fast enough to continue growth. This could be a direct or indirect disruption to this essential pathway, with the inner and outer membrane being the location at which this reaction occurs. An indirect disruption to this pathway could be that the fragile membrane as seen using microscopy in the Δdlp clones could make them more susceptible to the acidic environment compared to the WT and $\Delta::dlp$ clones. Further investigation into the fragile membrane using different stressing techniques apart from acid

and higher resolution microscopy would allow a more comprehensive and detailed look at the membrane to see if it is able to overcome these types of pressures.

The enzymatic activity of urease is controlled by the inner membrane, pH gated channel, UreI which regulates the access of urea to the bacteria (137, 250). The pH responsive transcriptional induction of this pathway is mediated by the ArsRS two component system via the binding of the phosphorylated response regulator ArsR to the promoters of operons, *ureAB* and *ureIEFGH* (251). However, this transcriptional regulatory system has been found to control a number of other genes in *H. pylori*, with many thought to have urease-independent mechanisms for acid adaption including *amiE* and *amiF* which are both involved in ammonia producing pathways separate to urease. Interestingly this regulation has been found for genes that target membrane processes like numerous OMPs as well as secreted flagella proteins (135). Interestingly the dynamin-like protein operon in 26695, HP0731-HP0733 was one of the operons found to also be positively regulated by this two-component system at low pH (135). This could indicate that the dynamin operon works in concert with the urease pathway in some way, and that is why acid adaption after short exposure has been hindered in the absence of the operon. It could also indicate a novel urease-independent acid adaption mechanism utilised by *H. pylori*, with this transcriptional regulatory system seemingly important in the expression of membrane targeting proteins like OMPs and the flagella which are of high significance in acid adaption.

Ammonia producing pathways to survive short exposure to acid are not the only form of acid adaption in *H. pylori*, with other key acid adaptive responses include motility and chemotaxis, where the bacteria can actively move away from low pH using its flagella combined with a well characterised chemotaxis pathway (252). It is thought that these systems allow for rapid and efficient entrance into the mucus layer and an escape from the stomach lumen. Once inside the mucus layer the diffusion rate of chemicals is much lower and allows the HCl to be neutralised more simply and the establishment of neutral conditions can be maintained for longer (253). To test the ability of the B128.7.13 Δ *dlp* strains to carry out this maintained acid adaptive function, the Δ *dlp* clones were exposed to varying amounts of acid for 72 h and were not able to overcome this stress despite being supplemented with

10 mM urea. This was evident with a low final OD_{600nm} of anywhere between 0 - 1 for Δdlp clones whereas the WT and complemented strains could grow quite reproducibly to a final OD of 1 - 2.5 after 72 h (*Figure 5.12*). When this was further investigated with a full growth curve, the Δdlp clones showed a reduced overall growth rate when compared to WT and the complemented strains; despite being able to neutralise the acid after 15 h (*Figure 5.15*). This indicates that the urease pathway is being switched on and is functioning, causing the neutralisation of the acid in the media, though there was a delay to this. A fragile membrane could cause this delay. The overall inability of the Δdlp strains to combat this growth stress could indicate a direct or indirect disruption to these alternative pathways that allow longer term survival. Taken together the inability of the Δdlp clones to overcome both a 1 h acid exposure as well as not being able to grow in acidic conditions indicates that there is something in common in both stresses, most likely the membrane. Intriguingly, dynamin could be maintaining the integrity of the membrane under these acid conditions to ensure these pathways can maintain homeostasis.

The most notable finding was that the oversupply of urea at a concentration of 100 mM could rescue the acid exposure as well as the extended acid growth defects in the Δdlp clones. This was shown with the Δdlp clones being able to maintain their starting OD_{600nm} of 0.5 in the exposure study (*Figure 5.11*) and their ability to grow to comparable OD_{600nm} over the course of 72 h (*Figure 5.14*). Across both studies, it was seen that at neutral pH, urea is toxic which has been seen previously (235). The oversupply rescue indicates that the urease pathway itself isn't directly impacted by the absence of dynamin but rather may not be working at a sufficient rate to overcome the extra acid stress caused by the fragile membrane, highlighting the importance of the membrane for bacterial survival within this harsh environment. The oversupply rescue of the Δdlp clones shows that dynamin could be playing a role in an alternative pathway of acid adaptation that acts through maintaining the integrity of the membrane.

There are still a number of unanswered questions that surround the function of dynamins in *H. pylori*. The unusual operon structure with multiple splits identified by bioinformatics (Figure 5.6) not only in the *dlp1* gene but also the *dlp2* gene requires further study as to whether this holds true biologically and whether it is important in the regulation and expression of these genes as seen previously in this thesis in strain 26695. The increased haemolysis, reduction in clumping and increased cell permeability in the absence of dynamin warrants further investigation also. This further investigation may lead to seeing dynamin potentially involved in the release or movement of the important cytolysin, TlyA-like protein, across the membrane to cause infection to the host; a function reminiscent of their closest homologs LeoCBA. The clear reduction in the ability of the bacteria to overcome acid stress both in short and long-term exposures indicates a more general weakening and seemingly fragile membrane in the absence of dynamin, with the membrane also being important factor in the function of the TlyA-like protein.

This chapter details preliminary developments and data in the understanding of the function of dynamins in *H. pylori*. Further creation of additional mutant strains is needed, as they are essential for the comprehensive investigation of the function(s) of dynamins in this important human pathogen. This study focussed on investigating the function of dynamin in *H. pylori* under one key stressor, acid, and so to further understand dynamins potential function in general membrane integrity for *H. pylori*, several membrane stressing conditions should be considered including antibiotic exposure, temperature and oxidative stress. Considering the need for membrane integrity for resisting the lytic effects of antibiotics, it would be hypothesised that without dynamin, *H. pylori* will now be sensitive to a larger range of antibiotics. To compliment the more traditional growth curve based analysis presented in this chapter, a high throughout, molecular level analysis method would be ideal, focussing on the proteomic changes occurring in the cell. A label free quantitative mass spectrometry methodology, SWATH-MS, to determine the differences in the abundance of proteins in all the mentioned conditions both in the presence and absence of dynamin is the subject of the following chapter (see Chapter VI). This methodology will allow a global, untargeted approach to the investigation of *H. pylori* dynamins potential role in general membrane integrity under varying stressful conditions.

The eukaryotic members of the dynamin superfamily have a diverse range of functions within a cell and it is no surprise that the bacterial members of this family are starting to show the same diversity in their cellular function despite maintaining the dynamin fundamental role in membrane dynamics.

CHAPTER VI

EXPLORING THE GLOBAL PROTEOMIC CHANGES TO B128.7.13 IN RESPONSE TO ACID

PREFACE

In this thesis, the presence of Dlp in *H. pylori* has been confirmed, with their expression *in vivo* shown (*Chapter III and IV*). They exhibit the hallmarks of previously characterised bacterial dynamins, such as having the ability to change the shapes of liposomes and localising to a membrane fraction of the cell. Also, evidence has been gathered to show that they potentially could be interacting to carry out their function. Stemming from this, their function has been explored through the creation of a Dlp knock out strain in the *H. pylori* gerbil adapted strain, B128.7.13 (*Chapter V*). This knockout strain showed evidence of a loss of membrane integrity, with an increase in haemolysis of the strain on blood agar (*Section 5.3.4*) and permeability to a DNA stain, SYTOX visualised by microscopy (*Section 5.3.5*). Finally, the Dlp knock out strain was tested under the important *H. pylori* environmental stress, acid and was not able to overcome both a short or long-term exposure (*Section 5.3.6 and Section 5.3.8*). However, with an oversupply of urea this was rescued (*Section 5.3.6 and Section 5.3.8*), indicating that this inability of *H. pylori* to survive in acid in the absence of dynamin was urease independent. To further investigate the effect that the absence of Dlp has on *H. pylori*, the global proteomic changes were explored and will be the focus of *Chapter VI*.

6.1 INTRODUCTION

The accurate quantification of proteome changes poses a significant challenge in biology, especially when looking at the global proteome level. Previously, the only methods available to carry out quantification were the data-dependent acquisition (DDA) mass spectrometry (MS) techniques of either label free approaches which can be inaccurate, or labelled quantification, such as the metabolic incorporation of stable isotope containing amino acids (SILAC) or the chemical labelling methods of iTRAQ and Tandem Mass Tags (TMT), which can be quite expensive, with both not being overly comprehensive in an acceptable timeframe. Sequential window acquisition of all theoretical spectra mass spectra (SWATH-MS) is a variant of the class of data-independent acquisition (DIA) methods that record fragment ion spectra of all ionized species in a sample (254). SWATH-MS is used for unbiased and untargeted quantification of proteins using MS to investigate global proteome changes across a range of conditions. This DIA method is able to quantitatively describe the global differential abundances of proteins without the inherent biases of labelling strategies (255-257). SWATH-MS can be applied to a broad range of biological conditions with any number of variables to be compared, with biological and technical replication. A key feature of SWATH-MS is that it is a label free strategy and therefore allows a more “native” quantification of the proteome under a range of conditions that aren’t limited by a labelling strategy. SWATH-MS takes the best part about discovery proteomics, performing an untargeted analysis where knowledge of expected protein abundance changes is not required, and merges it with the best part of targeted proteomics, specifically selected reaction monitoring MS which takes multiple measurements of a unique peptide for accurate quantification. This is a state of the art technology in the field of proteomics which is well-established as a method capable of simultaneously describing changes to multiple, complete pathways (258). This type of analysis then enables a more detailed study of the proteins discovered under conditions of interest, as potential future targets for developments in areas such as drug discovery.

As a DIA MS method, SWATH-MS fragments all precursor ions present at a particular moment in time and measures all of the fragment ions generated from multiple precursor ions in a single spectrum. This prevents the use of the usual proteomics MS search algorithms, such as Mascot (259), to identify the peptides present in a sample. Thus, the underpinning data for SWATH-MS is a comprehensive ion library generated for the specific organism of interest to form a library for comparison; if a peptide representing the protein isn't in the library, identification and quantification of unknown peptides in SWATH-MS is not possible. This ion library can be built from a comprehensive whole proteome sample in DDA acquisition mode, with the most ideal being a sample from all of the conditions that are being tested. The strength of this method is that the library can be continually expanded as more conditions are added throughout a study. SWATH-MS has been successfully performed on a number of bacteria (260-262) including the *H. pylori* strain J99 which examined proteins involved in the shape changes that occurred over long periods of growth. This study was able to successfully build a comprehensive library that covered approximately 62% of the proteome and quantitatively identified significant changes to a number of proteins and pathways that had not been previously implicated in prolonged growth for *H. pylori* (263). This seminal study using SWATH in *H. pylori* highlights that global proteomic analysis can help bring to light previously forgotten or not observed connections between pathways and conditions.

H. pylori is a dynamically changing organism which is known to drastically change its genome under different conditions, this in turn causes changes to its proteome to ensure survival even in the harshest of conditions, like survival in the stomach (200, 229). This inherent adaptability is a fundamental feature of bacterial pathogens, allowing them to respond to the many stressors that they encounter during their lifecycles. Stress adaptations can come in a wide variety of forms; from genomic changes seen in phase variation regulation, to transcriptional regulation of gene expression to even proteomic changes via post translational modifications. These survival responses can lead to multiple beneficial changes, which is evident in *H. pylori* with the acquisition of new defence and host-cell interaction mechanisms and enhanced protein multi-functionality (also known as protein

moonlighting for secondary functions) allowing it to carry out more than one function within the cell (195, 199, 200, 264, 265) .

For *H. pylori*, acid stress is arguably the most important environmental stressor that it, and most enteric pathogens for that matter, face within the gastrointestinal tract in order to colonise this niche. To overcome acid stress, *H. pylori* has evolved various acid-tolerance mechanisms to deal with acid such as the urease pathway, as well as chemotaxis and cell motility to neutralise and escape from areas of low pH respectively (122). Both of these mechanisms rely heavily on the membrane of *H. pylori* being intact and rather impermeable to the acidic environment to ensure homeostasis within the bacterium. Previous work in this thesis has shown that in the absence of the newly identified dynamin genes, *dlp1* and *dlp2* in *H. pylori*, the membrane becomes compromised and the bacteria are not able to tolerate acid conditions (*Chapter V*). Therefore, understanding the global proteome changes in the absence of dynamin genes may elucidate a potential function for these genes and any putatively relationship to acid tolerance.

Acid tolerance is absolutely crucial to the survival of *H. pylori* and there has been no comprehensive proteomic study to date on the proteins that are changing in response to low pH, specifically in the gerbil adapted strain of *H. pylori* B128.7.13. Furthermore, studying the global proteome response and changes in protein abundance by SWATH-MS, in the absence of dynamins, will lead to a greater understanding of what role these proteins are playing in this important clinical pathogen.

The aims of this chapter were to;

1. Construct a comprehensive ion library for the gerbil adapted strain of *H. pylori* B128.7.13
2. Investigate and quantify the proteomic changes in both neutral and acidic conditions using SWATH-MS in the presence and in the absence of dynamin-like proteins

6.2 MATERIALS AND METHODS

6.2.1 PREPARATION OF SAMPLES FOR SWATH

All samples for SWATH and ion library production were prepared as outline in *Section 5.2.6* with a few minor amendments as follows; only normal media (pH7) and acidic testing media (pH 5) supplemented with 10mM urea were used.

The cells were harvested by centrifugation at 4000 x g for 10 min at 4°C and washed twice in 1 x TBS (*Table 2.5*). The pellet was resuspended in UTC7 (*Table 2.5*) (1 mL) and sonicated on ice (3 x 30 sec intervals, 40% maximum output [VC 50T, Sonic]). The sonicated sample was centrifuged at 10 000 x g for 30 min and the pellet discarded. Reducing and Alkylating Solution (*Table 2.5*) (1 mL) was added to the sample and incubated at RT for 90 min. After reduction, acetone (10 mL) was added and the sample placed at -20°C for 30 min to allow protein precipitation. The sample was centrifuged at 4000 x g for 15 min and the supernatant discarded. The pellet was resuspended in UTC7 (500 µL). A Bradford assay (Sigma-Aldrich, USA) was used to determine the protein sample concentrations. 100 µg of each sample was combined with 6 volumes of 100 mM NH₄CO₃ to dilute the urea out of the sample. Trypsin Solution (*Table 2.5*) (1 µL) was added to the sample and incubated at 37°C overnight. After digestion, the sample was centrifuged for 5 min at 16,000 x g to remove any particles. A Silicycle HLB SPE syringe column (GE Biosciences) was suspended over a 15 mL falcon tube, with the appropriate solution applied to the top of the column and allowed to flow via gravity. The column was prepared as follows; to activate the column resin, 100% acetonitrile (1 mL) was added. Following this, the column was prepped for sample application by adding MS loading solvent (*Table 2.5*) (1 mL). The sample was then applied and allowed to run through completely. To wash away any contaminants, salts and unbound peptides from the column, MS loading solvent (*Table 2.5*) (1 mL) was added. Following this, elution solvent (*Table 2.5*) (400 µl) was applied. The flow through (~400 µL), containing eluted peptides which were quantified using a peptide assay (Thermo Scientific. Pierce Quantitative Colorimetric Peptide Assay).

6.2.2 DATA DEPENDENT ACQUISITION (LC/MS/MS) FOR ION LIBRARY CONSTRUCTION

For the creation of the Ion Library, all samples were mixed together in equal ratio and 20 µg of peptide was loaded at 10 µL/min in 90% Acetonitrile, 0.1% Trifluoroacetic acid (TFA; solvent B) into a TSK-Amide 80 HILIC column (3 µm particle size, 320 µmID x 17 cmL). Loading occurred over 10 min before the flow rate was dropped to 6µL/min and a linear gradient elution was performed with 0.1% TFA in water (solvent A) using the following program: 0-30% solvent A over 27 min, 30-70% solvent A over 4 min, 70% solvent A for 1 min, 70-0% solvent A over 4 min. Peptides were collected into 10 fractions manually in volumes depending on the UV absorbance, monitored at 220nm.

Data Dependent Acquisition (DDA) was performed using a TripleTOF 6600 mass spectrometer (Sciex, CA) at the Charles Perkins Centre (CPC), University of Sydney (USYD) in order to obtain ion spectra for peptides in the WCL prepared above. This was achieved based on methods previously described by Tacchi *et al* (266). Using an Eksigent 415 autosampler connected to a 415 nanoLC system (Eksigent, USA), 5 µL of the sample containing ~1 µg of peptide was loaded at 300 nL/min with MS buffer A (2% Acetonitrile + 0.2% Formic Acid) by direct injection onto an in-house manufactured fritless column (75 µmID x 150 mm) packed with ReproSil-Pur C18AQ resin (1.9 µm particle, 120Å pore size, Dr Maisch, Germany). Peptides were eluted from the column and into the source of a 6600 TripleTOF hybrid quadrupole-time-of-flight mass spectrometer using the following program: 2-35% MS buffer B (80% Acetonitrile + 0.2% Formic Acid) over 90 min, 35-95% MS buffer B over 9 min, 95% MS buffer B for 10 min, 80-2% for 2 min. The eluting peptides were ionised at 2300 V. An Intelligent Data Acquisition (IDA) experiment was performed, with a mass range of 350-1500 Da continuously scanned for peptides of charge state 2⁺-5⁺ with an intensity of more than 400 counts/s. Up to 50 candidate peptide ions per cycle were selected and fragmented and the product ion fragment masses measured over a mass range of 100-2000 Da. The mass of the precursor peptide was then excluded for 15 sec. A BSA digest control was run before and after the WCL to ensure optimal instrument function.

6.2.3 DATA INDEPENDENT ACQUISITION (SWATH) AND DATA ANALYSIS

SWATH acquisition of the whole cell proteome for all samples was performed in order to identify proteins in each sample. The same mass spectrometer, chromatography column and gradient conditions were as described above (*Section 6.2.2*). SWATH acquisition was achieved by first performing a full MS scan over the range of 350-1500 Da with an accumulation time of 50 ms. This was followed by 22 scan windows of 40 Da in width with a 1 Da overlap between windows and a 100 ms accumulation time per window. This resulted in a complete cycle time of ~2.3 sec. The cycle was repeated through the 120 min acquisition time. BSA controls were run before and after each sample.

The MS/MS data file produced by DDA analysis above was imported into ProteinPilot (v5.0). The Peptide library was generated from the 10 HILIC fractions described above along with samples of unfractionated WCL that were combined to create a single common peptide library used for the SWATH analysis. The files were searched against the B128.7.13 predicted proteome, generated using PROKKA (outlined in *Chapter V*) where the identified peptides and their corresponding mass spectra were assigned as an ion library database. The identified peptides with a 1% FDR served as the ion library to apply to the SWATH acquisition runs. The parameters set in ProteinPilot were as stated in Loke *et al*, and are as follows, fixed modifications: none, variable modifications: propionamide, oxidised methionine, deamidated asparagine, enzyme: semitrypsin, number of allowed missed cleavages: 3, peptide mass tolerance: 20 ppm, MS/MS mass tolerance: 0.1 Da, charge state: 2⁺, 3⁺ and 4⁺ and 1% False Discovery Rate (FDR) (263).

For SWATH-MS data, information such as retention time, amino acid sequence, protein information for the correct peptide peaks was extracted, compiled from the ProteinPilot ion library and processed using the SWATH Acquisition Microapp (v1.0) in the PeakView software (v1.2). Following the workflow for SWATH outlined by Loke *et al*, the number of proteins first imported into the SWATH analysis software was restricted to the reported protein number based on false discovery rate of 1% and peak mass accuracy of 50 ppm by user input. Peaks were extracted based on software criteria (non-user input), such as mass

accuracy (within 50 ppm), charge state, retention time, peak width and area, etc. Parameters for identification of a peak included, an extraction mass window of 50 ppm and within 4 min of their expected retention time. Each peak area for each protein was then extracted using 3 peptides (excluding shared peptides if any identified) and 3 transitions or fragment masses for all the samples. The peptide list was also further filtered by peptide confidence level (set as 95% based on ProteinPilot scoring definition) (263). All quantitative analyses including normalisation and statistical calculations were processed using the MarkerView (v1.2.1) software. Samples were grouped according to the experimental design; WT Normal, WT Acid, Mutant Normal and Mutant Acid and triplicate data was imported. Normalisation was performed based on total sum of the areas under the peaks in Markverview 1.2.1. Student's t-test were performed to compare each group, using p-value scoring of less than 0.05 as a cutoff for significance and an area fold change (> 2) as a significant expression change.

6.3 RESULTS

6.3.1 ION LIBRARY CONSTRUCTION

To investigate the global proteomic changes that occur in *H. pylori* when exposed to acid, both in the presence and absence of dynamin using SWATH MS, it was necessary to construct a reference ion library. Both B128.7.13 WT and B128.7.13 Δ *dlp* C2 strains were grown to mid-log phase and then exposed to an acidic environment (pH 5) for an hour with shaking. Whole cell lysate was collected from both strains at both pH 7 and pH 5. This was repeated in triplicate. All samples that were used in the SWATH acquisition (IDA) were also used in a data dependent acquisition (DDA) for library construction. Each sample was extensively fractionated using HILIC to ensure maximum coverage of the proteome during the data dependent acquisition. Due to the lack of previous studies of this *H. pylori* strain, B128.7.13, the genome first needed to be annotated in order to generate the predicted proteome required for MS spectral identification; this was done using PROKKA (as seen in *Chapter V, Figure 5.3*). This annotation produced a predicted proteome that consisted of 1655 open reading frames that were subsequently used as the database to search the DDA spectra against.

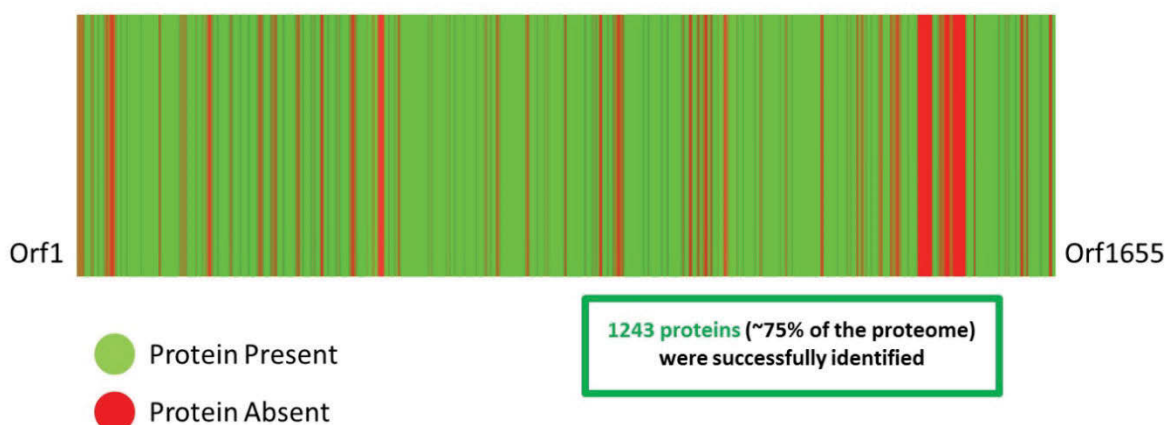


Figure 6.1 Coverage of the Predicted *H. pylori* B128.7.13 Proteome

*Visual representation of the proteome coverage achieved for the predicted proteome *H. pylori* B128.7.13. Protein identification was established with a minimum of 1 unique peptide and 95% confidence. If a protein met these criteria, it was labelled as present (green), if a protein didn't meet these criteria it was labelled as absent (red). Moving from left to right the proteome is depicted. Coverage of 75% of the predicted proteome was achieved. The ion library had the additional criteria of 2 unique peptides for protein identification and has 72% coverage, or 1197 proteins identified as present out of 1655 predicted proteins for the B128.7.13 proteome.*

The ion library consists of ~72% of the proteome with 1197 identified proteins with a minimum unique peptide match of 2 and a > 95% confidence (*see S7a*). If the data is expanded to include proteins that have been identified by one unique peptide, an additional 50 proteins can be noted, however this is not considered statistically significant for proteins of standard size and amino acid distribution. As seen in *Figure 6.1*, the coverage of the proteins detected spans the entire predicted proteome with just a few notable hot spots where a group of proteins were not detected. These hotspots are most notably towards the end of the proteome, consisting of mainly hypothetical proteins. Of the approximately 400 proteins not detected in this study, ~56% of them are predicted hypothetical ORFs and/or putative proteins. The lack of detection of peptides representing these proteins is most likely due to the low abundance of the protein in the samples or that the protein is not being produced under these culture conditions. Also with this proteome being a predicted proteome based on an annotation of an open genome, the number of ORFs may be overestimated.

In this dataset, the *dlp* operon was not detected by any peptides in the ion library, most likely due to low expression levels resulting in protein abundance below the limit of detection under these specific conditions. It is difficult to determine the reason for this as Dlp1 and Dlp2 have only been detected by Western blot (*Chapter IV*) and have not been detected by MS in *H. pylori* to date and may not produce peptides of sufficient ionisation potential for MS. However, the previous characterisation of this strain (*Figure 5.3*) confirmed the deletion of the *dlp* operon in the knockout strain, therefore any changes in the proteome can be confidently attributed due to the absence of these genes.

6.3.2 PRINCIPLE COMPONENTS ANALYSIS OF SWATH-MS DATA

Principle components analysis was used to statistically analyse and visualise the SWATH-MS data, which can be described by multiple variables. Principal component analysis (PCA) is a mathematical procedure that transforms a number of correlated variables into a number of uncorrelated variables called principal components (PC). The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. Discriminant analysis (PCA-DA) can be employed, a supervised version of PCA where prior knowledge of sample groups is used to determine the variables that maximise the variation between groups and those which minimise the variation within a group. PCA-DA (supervised) with parteo scaling was performed to gain a general insight and overview of the dataset. The resulting analysis (*Figure 6.2*) demonstrates that B128.7.13 WT and B128.7.13 Δdlp were able to be discriminated using the SWATH-MS quantitative dataset, validating the analysis, as there is already a known difference between these strains. Further to this, it was shown that the four samples are able to be separated discretely using the first two PCs (D1 and D2 scores), confirming that the analysis is capable of predicting proteome changes in response to acid stress.

The first and second principle components did not greatly differ from each other, with 34.6% of the variation accounted for in D1 (strain) and then a further 33.6% accounted for in D2 (pH) (*Figure 6.2*). This indicates that the data generated in each replicate corresponds well with each other and can be discernible from each of the other samples in the analysis. This is a good result for this dataset and the analysis of these biological conditions, with it known that drastic variation in proteome sampling and MS can prevent accurate quantitative analysis by any other method.

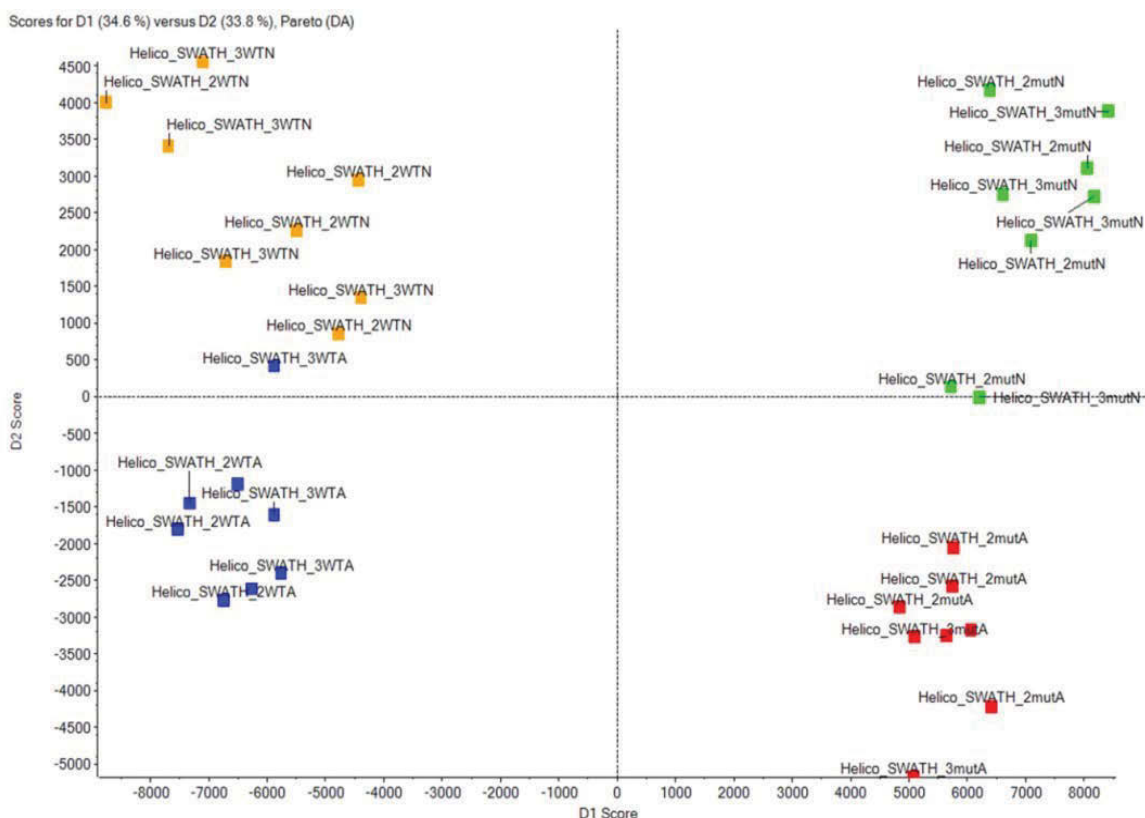


Figure 6.2 Variance in the SWATH-MS Sample Groups

PCA analysis carried out using Pareto-supervised test highlights difference between sample groups by plotting D1 and D2 scores; B128.7.13 WT - acid treated (blue), B128.7.13 WT – normal (yellow), B128.7.13Δdlp – acid treated (red) and B128.7.13Δdlp –normal (green). 34.6% of variance is accounted for in D1 (strain) and 33.8% of variance is accounted for in D2 (treatment). Adapted from the output from Markerview.

6.3.3 SUMMARY OF SIGNIFICANT PROTEIN CHANGES DURING SWATH ACQUISITION

Using the SWATH-MS analysis software Markerview 1.3.1 (AbSciex), 979 proteins from the B128.7.13 WT and B128.7.13 Δ dlp proteome were able to be quantified using the stringent statistical identification and quantification conditions described in the previous section (*see S7b, S7d, S7f and S7h*). This represents approximately 60% of the predicted proteome and 82% of the measurable proteome detected in our ion library. An orthogonal analysis was able to be performed on the data due to the experimental design chosen, therefore a range of t-tests were performed using Markerview to compare; B128.7.13 WT pH 7, B128.7.13 Δ dlp pH 7, B128.7.13 WT pH 5 and B128.7.13 Δ dlp pH 5 (represented in *Figure 6.3*). A subset of the SWATH hits were considered statistically significant (p-value < 0.05), using this statistical method and a summary of these results is indicated in *Figure 6.3*.

When comparing B128.7.13 WT under neutral and acidic conditions, 28 differentially abundant proteins were found to have a significant p-value, with fold changes between 0.36 - 3.28 (not transformed). Under the same conditions B128.7.13 Δ dlp was found to have 45 proteins that had significant changes in abundance, with fold changes varying from 0.44 - 2.16. In neutral conditions, there was found to be 120 proteins with significant differences between B128.7.13 WT and B128.7.13 Δ dlp with fold changes varying from 0.3 - 4.11. However, the most significant proteome changes were seen between the two strains under acidic conditions, with 170 proteins identified with fold changes from 0.32 - 3.4. These proteome comparisons will be the focus of the following sections.

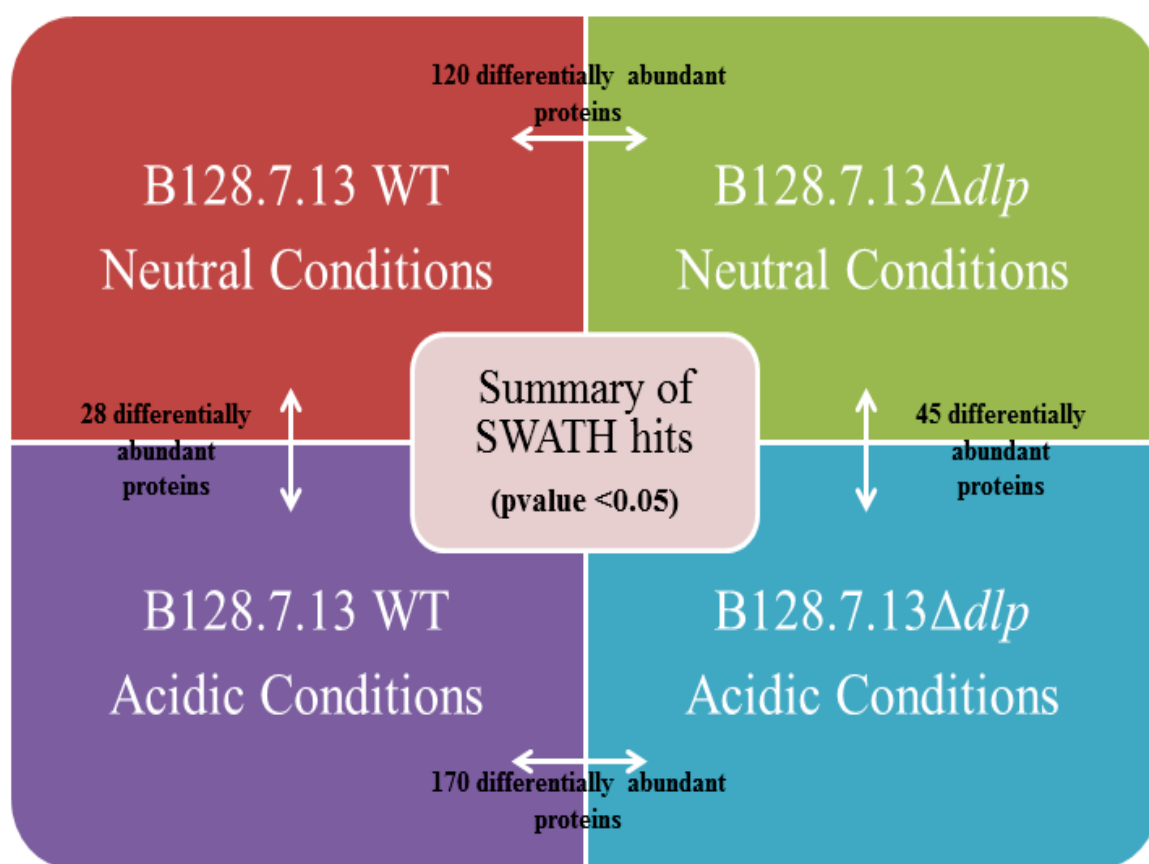


Figure 6.3 Summary of Significant Hits Quantified using SWATH-MS

SWATH-MS data analysis was carried out using Markerview 1.3.1 and a visual representation summarising the significant changes in abundance identified by SWATH-MS between each of the sample groups is seen above. A significant change to a proteins abundance between different sample groups was identified by a p-value of less than or equal to 0.05.

6.3.4 PROTEOME CHANGES TO B128.7.13 WT AFTER AN ACIDIC SHOCK (PH5)

B128.7.13 WT was grown for 1 h in both neutral (pH 7) and acidic conditions (pH 5) and the subsequent whole cell proteomes compared using SWATH-MS. This analysis demonstrated that the proteome showed minimal changes, after the exposure to acid as you could expect due to the short time for which it was exposed (*Figure 6.4A, see S7b for full list*). However, there were a number of proteins still shown to change in abundance due to the acid stress; the most significant of these are listed in *Table 6.1 (see S7c for full list)*.

Table 6.1 Significant Proteome Changes to B128.7.13 WT After an Acidic Shock

	Accession ¹	Protein Name ¹	p-value ²	Fold Change ³	Log Fold Change ⁴
Most Significant Change	PROKKA_01248	hypothetical protein	0.00054	1.6539	0.2185*
	PROKKA_00221	Ribonuclease J 1	0.00275	1.2425	0.0943*
	PROKKA_01300	DNA gyrase subunit B	0.00339	1.1699	0.0681*
	PROKKA_00669	Cysteine desulfurase	0.00484	0.8177	-0.0874*
	PROKKA_01444	CobQ/CobB/MinD/ParA nucleotide binding domain protein	0.00606	0.7152	-0.1456*
Lower abundance In acid	PROKKA_01174	Endonuclease MutS2	0.0439	0.3678	-0.4344
	PROKKA_00867	Putative heat shock protein HspR	0.04553	0.6148	-0.2113*
	PROKKA_00773	CMP-N,N'-diacetyllegionaminic acid synthase	0.02137	0.4737	-0.3245
Higher abundance In acid	PROKKA_01248	hypothetical protein	0.00054	1.6539	0.2185*
	PROKKA_01161	hypothetical protein	0.03873	1.8538	0.2681*
	PROKKA_00934	D-glycero- α -D-manno-heptose-1,7-bisphosphate 7-phosphatase	0.02257	3.2820	0.5161

¹ Accession and protein name - given by annotation of B128.7.13 WT by PROKKA

² p-value - given for each protein taking into account peptide coverage and relative frequency of the fold change across all replicates (significant cut off < 0.05)

³ Fold Change - the comparison based on relative intensity of the spectra across replicates between sample groups

⁴ Log Fold Change – Log_{10} (fold change)

*Proteins that lie just outside the fold change criteria but have a significant p-value

All 28 proteins found to have significant fold changes are shown in *Figure 6.4B*, with the proteins found to have a lower abundance at pH 5 in the left column and those with a higher abundance in the right column.

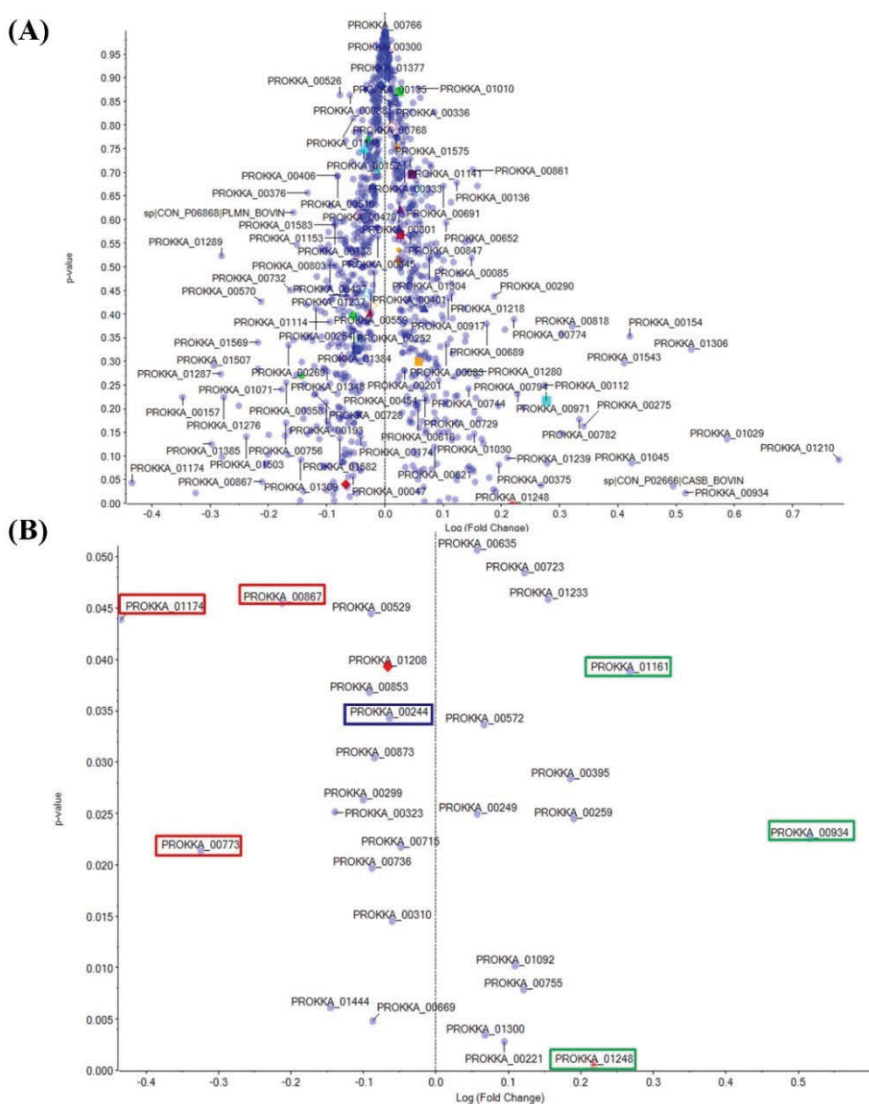


Figure 6.4 SWATH-MS Analysis of Proteome Changes to B128.7.13 WT After an Acidic Shock

Markerview analysis of SWATH-MS comparing B128.7.13 WT proteomic changes between pH 7 and pH 5. (a) Volcano plot of Log (fold change) in relation to the p-value of the complete quantified proteome (979 proteins). Each dot indicates a protein, coloured dots highlight tRNA proteins. (b) Volcano plot of Log (fold change) in relation to the p-value of the significantly ($p\text{-value} < 0.05$) changing quantified proteome, consisting of 28 proteins. 15 proteins were found to have a lower abundance in acidic conditions (left) with the three proteins with most change in their abundance highlighted (boxed red) and 14 proteins were found with a higher abundance in acidic conditions (right) with the three proteins with most change in their abundance highlighted (boxed green). A known acid adaptation protein, UreA (boxed blue) was noted to have a significant change that laid outside the 2x fold change criteria. Adapted from the output from Markerview.

When B128.7.13 WT was exposed to acid, the greatest fold change to an individual protein's abundance was of the Endonuclease MutS2. Its abundance was reduced by a fold change of 0.37 ($p\text{-value} = 0.0439$). MutS2 has been shown to be involved in the suppression of homologous recombination and is thought to have a key role in genetic diversity amongst *H. pylori* (267). The reduction in its abundance is most likely due to not being needed when the bacteria are under stress and is discussed further in *Section 6.4.1*. In addition, the CMP acid synthase protein (fold change 0.47; $p\text{-value}$: 0.02137) which is involved in catalysing an important step in a general glycosylation pathway and a putative heat shock protein, HspR (fold change: 0.61; $p\text{-value}$: 0.04553), also showed significant abundance changes when the WT was exposed to acid (*Figure 6.4B*, boxed red).

When B128.7.13 WT was exposed to acid shock, the protein with the greatest and most significant fold change was D-glycero-alpha-D-manno-heptose-1,7-bisphosphate 7-phosphatase with a fold change of 3.28 ($p\text{-value} = 0.02257$). This enzyme is involved in the biosynthesis of GDP-D-glycero-alpha-D-manno-heptose, a core oligosaccharide of the LPS of the bacteria (268). An increase in this enzyme during acid shock could be due to the need for maintenance of the membrane and this is further discussed in *Section 6.4.2*. Along with this enzyme, two hypothetical proteins, PROKKA_01248 ($p\text{-value} = 0.00054$) and PROKKA_01611 ($p\text{-value} = 0.039$) were also shown to have a significantly changing

abundance in response to acid with fold changes of approximately 1.7 (*Figure 6.4B*, boxed green). PROKKA_01248 shows high homology to Rho transcriptional terminator factor and PROKKA_01611 shows high homology to hydE which is associated with acid tolerance (269). In particular, hydE has previously been linked to the support of growth of *H. pylori* through its hydrogenase activity, important in colonisation of the stomach niche (269).

It was expected to observe components of known acid adaptation pathways, like the urease proteins, to appear in this data set. Despite not being scored as “most significant”, the urease subunit alpha (UreA) was found to have a slight change (0.8) with a p-value of 0.045 (*Figure 6.4B* boxed blue). The other members of the urease pathway, UreB and the accessory proteins show no significant changes with fold changes centring on 1, indicating no change. This is likely due to either the exposure not being for long enough (1 h) and/or the pH not being low enough to induce a rapid urease response (discussed further in *Section 6.4.2*). Further time points would allow for a more comprehensive look at the acid response in the WT strain in regard to this. Despite this, further comparisons are still possible with B128.7.13 Δ dlp, with the acidic exposure not strong enough to kill the cells but strong enough to induce the early changes that are potentially urease independent and that need to happen for acid tolerance in *H. pylori*.

6.3.5 PROTEOME CHANGES TO B128.7.13 Δ DLP AFTER AN ACIDIC SHOCK (PH5)

The aim of this study was primarily to increase our understanding of the Dlp's and so, under the same conditions as the WT (previous section), B128.7.13 Δ dlp was grown for 1 h in both neutral (pH 7) and acidic conditions (pH 5) and the subsequent whole cell proteomes compared using SWATH-MS. Changes that were seen to the whole proteome can be seen in *Figure 6.5A* (see *S7d* for full list) and the most significant of these are outlined in *Table 6.2* (see *S7e* for full list).

Table 6.2 Significant Proteome Changes to B128.7.13 Δ dlp After an Acidic Shock

	Accession ¹	Protein Name ¹	p-value ²	Fold Change ³	Log Fold Change ⁴
Most Significant Change	PROKKA_00571	Flagellin B	1.40E-05	2.1636	0.3352
	PROKKA_00924	Flagellar hook protein FlgE	0.00031	2.1101	0.3243
	PROKKA_01165	Quinone-reactive Ni/Fe-hydrogenase small chain precursor	0.00034	1.3618	0.1341*
	PROKKA_00960	Spermidine synthase	0.00084	0.6411	-0.1931*
	PROKKA_01278	hypothetical protein	0.00214	0.4429	-0.3537
Lower abundance In acid	PROKKA_01278	hypothetical protein	0.00214	0.4429	-0.3537
	PROKKA_01239	Acyl carrier protein	0.04071	0.4827	-0.3163
Higher abundance In acid	PROKKA_00924	Flagellar hook protein FlgE	0.00031	2.1101	0.3243
	PROKKA_00571	Flagellin B	1.40E-05	2.1636	0.3352
	PROKKA_00205	CtpB precursor	0.0341	1.6805	0.2253*

¹ Accession and protein name - given by annotation of B128.7.13 WT by PROKKA

² p-value - given for each protein taking into account peptide coverage and relative frequency of the fold change across all replicates (significant cut off < 0.05)

³ Fold Change - the comparison based on relative intensity of the spectra across replicates between sample groups

⁴ Log Fold Change – Log₁₀ (fold change)

*Proteins that lie just outside the fold change criteria but have a significant p-value

All 45 proteins established to have significant fold changes can be seen in *Figure 6.5B*, with proteins found to have a lower abundance in pH 5 shown on the left and those found to have higher abundance in pH 5 on the right.

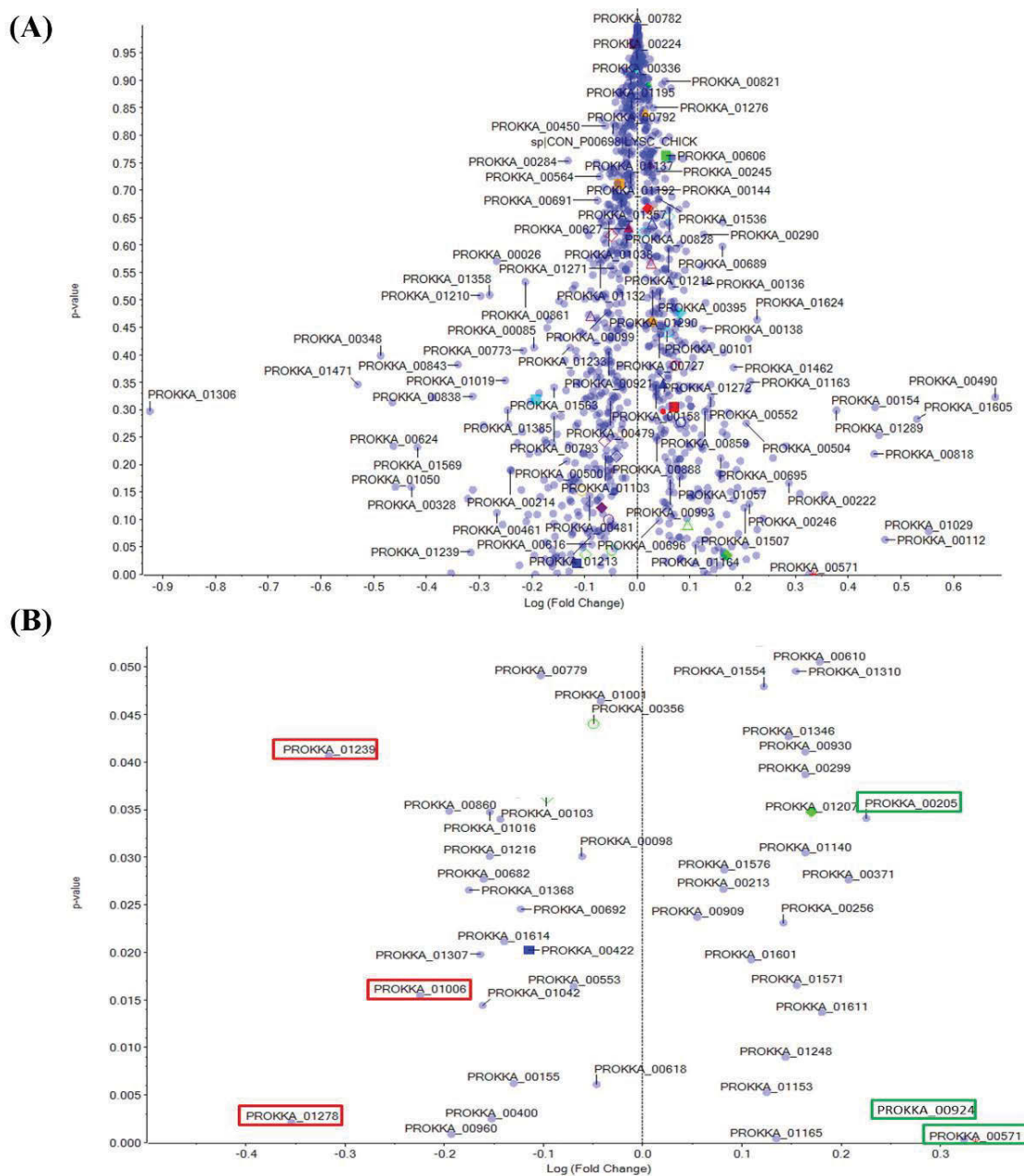


Figure 6.5 SWATH-MS Analysis of Proteome Changes to B128.7.13Δ*dlp* After an Acidic Shock

Markerview analysis of SWATH-MS comparing B128.7.13Δdlp proteomic changes between pH 7 and pH 5. (a) Volcano plot of Log (fold change) in relation to the p-value of the complete quantified proteome (979 proteins). Each dot indicates a protein, coloured dots highlight tRNA proteins. (b) Volcano plot of Log (fold change) in relation to the p-value of the significantly (p-value < 0.05) changing quantified proteome, consisting of 45 proteins. 23 proteins were found to have a lower abundance in acidic conditions (left) with the three proteins with most change in their abundance highlighted (boxed red) and 22 proteins were found with a higher abundance in acidic conditions (right) with the three proteins with most change in their abundance highlighted (boxed green). Adapted from the output from Markerview.

When B128.7.13Δdlp was exposed to acid the protein whose abundance was most significantly lower with a fold change of 0.45 (p-value= 0.00214) was PROKKA_01278. This hypothetical protein shows high homology to CagU, an inner membrane protein associated with the important *H. pylori* Type IV secretion system (270). Along with this protein hit, both an Acyl carrier protein involved in fatty acid synthesis and an outer membrane lipoprotein carrier protein precursor, showed fold changes of 0.48 (p-value= 0.04071) and 0.59 (p-value= 0.01552) respectively lowering their abundance in the proteome of B128.7.13Δdlp when the bacteria was exposed to acid (*Figure 6.5B*, boxed red). All these changes lay on the borderline of the 2-fold change ($\log_{10}0.3$).

When B128.7.13Δdlp was exposed to an acidic environment of pH 5, the proteins with the greatest and most significant fold changes in the whole data set were associated with the flagella; Flagellin B (FlaB) (p-value = 1.4×10^{-5}) and the flagellar hook protein (FlgE) (p-value = 0.00031) both with a fold change of approximately 2.1 ($\log_{10}0.33$). Motility is a key acid adaption that *H. pylori* implements to colonise the stomach and the importance of the flagella machinery is discussed in *Section 6.4.4*. The increase in these proteins is possibly the first evidence that B128.7.13Δdlp is switching on a stress response in reaction to acid that is different to WT. Along with these was a carboxy-terminal processing protease CtpB precursor, that also significantly (p-value = 0.0341) showed an increase in its abundance in response to acid with a fold change of 1.68 (*Figure 6.5B*, boxed green). An increase in

proteases, i.e. HtrA and CtpB is well established in *H. pylori* to be linked to the inflammation seen during the early stages of disease and often expression is stress regulated, i.e. acid (271)

Interestingly spermidine synthase was noted with a fold change that sits just outside our criteria ($\text{Log}_{10} \pm 0.3$), with a log fold change of -0.1931 but it was calculated to have a p-value of 0.00084 indicating that the change was quite significant despite being relatively small. Recently it has been shown that spermidine synthase does not have a role in the production of spermidine in *H. pylori* but rather interacts with the flagellar motor protein FlhA and is necessary for the normal motility that is essential for *H. pylori* colonization and infection in gastric mucosa (272).

6.3.6 PROTEOME CHANGES UNDER NORMAL CONDITIONS (PH7) TO B128.7.13ΔDLP

With the establishment of changes in each of the strains under normal lab conditions of pH 7 and the condition of interest, pH 5, the comparison between the strains under the same condition was now the focus. This specific comparison allowed the differentially abundant proteins between the strains under the same condition to be seen, and as seen in *Table 6.3* and *Figure 6.6B* there are many more changes that are of significance when compared to the changes seen in the individual strains.

Table 6.3 Significant Proteome Changes to B128.7.13Δdlp Under Neutral Conditions

	Accession ¹	Protein Name ¹	p-value ²	Fold Change ³	Log Fold Change ⁴
Most Significant Change	PROKKA_01167	hypothetical protein	7.63E-09	2.4568	0.3904
	PROKKA_01165	Quinone-reactive Ni/Fe-hydrogenase small chain precursor	1.50E-07	0.5002	-0.3008
	PROKKA_00096	Formamidase	2.38E-06	3.1083	0.4925
	PROKKA_00256	Cytochrome c551 peroxidase precursor	2.44E-06	0.4772	-0.3213
	PROKKA_01164	Quinone-reactive Ni/Fe-hydrogenase large chain	7.59E-06	0.5385	-0.2688*
Lower abundance In Δdlp	PROKKA_01174	Endonuclease MutS2	0.0263	0.3056	-0.5149
	PROKKA_01140	Ferritin	0.01294	0.3380	-0.4711
	PROKKA_01159	hypothetical protein	0.01583	0.3674	-0.4348
	PROKKA_00600	Cbb3-type cytochrome c oxidase subunit CcoP	0.04378	0.4216	-0.3751
	PROKKA_01503	fatty acid biosynthesis transcriptional regulator	0.04276	0.4310	-0.3655
	PROKKA_00084	Cytochrome c-553 precursor	0.01753	0.4543	-0.3426
	PROKKA_00256	Cytochrome c551 peroxidase precursor	2.44E-06	0.4772	-0.3213
	PROKKA_01165	Quinone-reactive Ni/Fe-hydrogenase small chain precursor	1.50E-07	0.5002	-0.3008
Higher abundance In Δdlp	PROKKA_00401	4-hydroxythreonine-4-phosphate dehydrogenase	0.0005	2.0035	0.3018
	PROKKA_01108	Fe(3+) dicitrate transport protein FecA precursor	4.14E-05	2.0494	0.3116
	PROKKA_00745	Aliphatic amidase	0.00123	2.2398	0.3502
	PROKKA_00791	Flagellar M-ring protein	6.76E-05	2.2782	0.3576
	PROKKA_00264	Branched-chain-amino-acid transaminase I	0.01693	2.4303	0.3857
	PROKKA_01167	hypothetical protein	7.63E-09	2.4568	0.3904
	PROKKA_00980	hypothetical protein	0.04757	2.5508	0.4067
	PROKKA_01588	UvrABC system protein B	0.04947	3.0158	0.4794
	PROKKA_00096	Formamidase	2.38E-06	3.1083	0.4925
	PROKKA_00556	hypothetical protein	0.03933	4.1186	0.6148

¹ Accession and protein name - given by annotation of B128.7.13 WT by PROKKA

² p-value - given for each protein taking into account peptide coverage and relative frequency of the fold change across all replicates (significant cut off < 0.05)

³ Fold Change - the comparison based on relative intensity of the spectra across replicates between sample groups

⁴ Log Fold Change – Log₁₀ (fold change)

*Proteins that lie just outside the fold change criteria but have a significant p-value

In *Figure 6.6A*, it can be seen in the volcano plot of the Log (fold change) vs. p-value that the majority of the proteome shows minimal change, with many proteins with fold changes in the range of -0.3 - 0.3 and a p-value greater than 0.05 i.e. no change in abundance (*see S7f for full list*). This is expected in a global approach to investigate the proteomic changes between two strains under the same condition because on this large scale, majority of the proteins will not significantly change in abundance. This is the strength of SWATH-MS analysis, enabling accurate quantification of important proteins amongst a large volume of unchanging protein MS data. When the significance threshold is narrowed to less than 0.05, as seen in *Figure 6.6B*, there are 120 proteins that fit this criterion of significance (*see S7g for full list*). The next criterion was to look at the proteins with a fold change greater than or equal to 2 (a greater than two-fold lower change is a value of less than or equal to 0.5) or equal to and less than 0.5 in the case of reduced abundance. These criteria allow data of high biological significance to be drawn from this large data set which would otherwise be difficult to analyse. These data are summarised in *Table 6.3* below and presented in full in *Figure 6.6C-D*.

In the baseline condition of pH 7, where the cut off of fold changes equal to or greater than 2 (a greater than two-fold lower change is a value of less than or equal to 0.5), there were 8 proteins that had significantly lower abundance in B128.7.13 Δ *dlp* when compared to the WT, shown in *Figure 6.6C*. The lowest of these was an endonuclease MutS2 protein with a fold change of 0.3056 (*Figure 6.6C boxed red*). MutS2 was previously noted in *Section 6.3.4* as a significantly changing protein for WT under acidic conditions and this is discussed further in *Section 6.4.1*. Also in this data it was noted that 3 cytochrome C associated proteins were in lower abundance by a factor of approximately 2 compared to the WT proteome; Cbb3-type cytochrome c oxidase subunit CcoP, Cytochrome c-553 precursor and Cytochrome c551 peroxidase precursor. Each of these had a p-value of significance, 0.04, 0.02 and 2.44×10^{-6} , respectively, with Cytochrome c551 peroxidase precursor having one of the most significant p-values in this comparison. The cytochrome C family of proteins are known to have roles in energy synthesis and quite interestingly are located in the inner membrane of the bacterial cell and their potential function in normal growth in the absence of dynamin is discussed further in *Section 6.4.1* (273).

Under neutral growth conditions there were 10 proteins in B128.7.13 Δ *dlp* that showed a significantly increased abundance in the proteome when compared to the WT (*Figure 6.6D*, right of line). The highest fold change in the B128.7.13 Δ *dlp* proteome under neutral conditions was PROKKA_00556. This hypothetical protein shows homology to an AANH-like protein in BLAST. Quite interestingly, in this data set, both AmiE and AmiF, two characterised proteins involved in acid adaptation through maintenance of cytoplasmic homeostasis, were seen with increases to their abundance in B128.7.13 Δ *dlp* with fold changes of 2.2 and 3.1 respectively. This interesting finding is discussed further in *Section 6.4.1*.

From these results, it can be seen that the deletion of dynamin from the *H. pylori* genome has an effect on the *H. pylori* proteome in a pH 7 neutral environment, with a number of proteins changing in their abundance. Of particular interest is the increased abundance of cytoplasmic pH maintenance proteins, AmiE and AmiF, induced by the absence of dynamin.

Markerview analysis of SWATH-MS comparing proteomic changes at pH 7 between B128.7.13 WT and B128.7.13Δdlp. (a) Volcano plot of Log (fold change) in relation to the p-value of the complete quantified proteome (979 proteins). Each dot indicates a protein, coloured dots highlight tRNA proteins. (b) Volcano plot of Log (fold change) in relation to the p-value of the significantly (p-value < 0.05) changing quantified proteome, consisting of 120 proteins that are differently expressed. (c) 8 proteins were found to have an abundance that was equal to twofold or less in B128.7.13Δdlp. (d) 10 proteins were found with an abundance that was equal to twofold or higher in B128.7.13Δdlp (right of the dashed green line). Adapted from the output from Markerview

6.3.7 PROTEOME CHANGES AFTER AN ACIDIC SHOCK (PH5) TO B128.7.13ΔDLP

The final and arguably the most important proteome analysis was to examine the differences between the strains, B128.7.13 WT and B128.7.13Δdlp after 1 h exposure to acidic conditions. These conditions mimic what the bacteria encounters in the stomach, revealing any changes that elucidate important pathway(s) that dynamin may interact with during an acid challenge. This comparison allowed the differentially abundant proteins between the wild-type and deletion strains under the same biologically relevant condition to be seen. The results presented in *Table 6.4* and *Figure 6.7B* demonstrates that there are many significant effects of dynamin deletion on the proteome of *H. pylori*. Some of the largest proteome changes were observed between these two groups, particularly when compared to the changes seen in these strains under the same conditions, which has substantial implications for the role dynamins play in *H. pylori* biology.

The majority of the proteome showed minimal changes between the two strains in acid conditions, with most proteins showing fold changes between -0.3 - 0.3 and a p-value greater than 0.05 (*Figure 6.7A* and *see S7h for full list*). When the data is narrowed to the significant threshold of 0.05, 170 proteins were identified (*Figure 6.7B*). Restricting this list to proteins with a fold change greater than or equal to 2 and less than or equal to 0.5, 16 proteins were

found to be significantly changing in *Figure 6.7C-D* and are summarised below in *Table 6.4* (see *S7i* for full list),.

Table 6.4 Significant Proteome Changes to B128.7.13 Δ dlp Under Acidic Conditions

	Accession ¹	Protein Name ¹	p-value ²	Fold Change ³	Log Fold Change ⁴
Most Significant Change	PROKKA_01108	FecA precursor	1.50E-08	2.5497	0.4065
	PROKKA_00571	Flagellin B	1.56E-08	3.1198	0.4941
	PROKKA_01140	Ferritin	2.34E-08	0.3267	-0.4858
	PROKKA_00745	Aliphatic amidase	3.92E-08	1.9209	0.2835*
	PROKKA_00562	Cystathionine gamma-synthase	5.35E-08	2.4398	0.3874
Lower abundance In Δdlp	PROKKA_01140	Ferritin	2.34E-08	0.3267	-0.4858
	PROKKA_00451	Alginate biosynthesis protein AlgA	0.02133	0.4729	-0.3252
	PROKKA_01066	tRNA-dihydrouridine synthase C	0.00255	0.4745	-0.3238
	PROKKA_01598	Anti-sigma-28 factor, FlgM	0.00662	0.4792	-0.3195
Higher abundance In Δdlp	PROKKA_00299	Helicobacter outer membrane protein	0.00102	2.0756	0.3171
	PROKKA_01342	Flagellar hook protein FlgE	0.00702	2.0829	0.3187
	PROKKA_01346	Helicobacter outer membrane protein	0.00054	2.0971	0.3216
	PROKKA_01167	hypothetical protein	2.76E-05	2.1895	0.3404
	PROKKA_00588	hypothetical protein	0.01743	2.3537	0.3717
	PROKKA_00482	3-hydroxyacyl-[acyl-carrier-protein] dehydratase FabZ	0.01296	2.4164	0.3832
	PROKKA_00791	Flagellar M-ring protein	0.00305	2.4300	0.3856
	PROKKA_00562	Cystathionine gamma-synthase	5.35E-08	2.4398	0.3874
	PROKKA_01108	FecA precursor	1.50E-08	2.5497	0.4065
	PROKKA_00924	Flagellar hook protein FlgE	1.80E-05	2.8644	0.4570
	PROKKA_00571	Flagellin B	1.56E-08	3.1198	0.4941
	PROKKA_00096	Formamidase	3.06E-07	3.4098	0.5327

¹ Accession and protein name - given by annotation of B128.7.13 WT by PROKKA

² p-value - given for each protein taking into account peptide coverage and relative frequency of the fold change across all replicates (significant cut off < 0.05)

³ Fold Change - the comparison based on relative intensity of the spectra across replicates between sample groups

⁴ Log Fold Change – Log_{10} (fold change)

*Proteins that lie just outside the fold change criteria but have a significant p-value

Of the top 5 most significant protein hits, 4 were observed to have a higher abundance in the B128.7.13 Δ dlp after the acidic challenge (*Table 6.4*). These include the proteins aliphatic amidase (AmiE), FecA precursor, Flagellin B and Cystathionine gamma-synthase, with the last three also showing a fold change greater than 2. Each of these proteins was quantitatively identified with high confidence, with significance values (p-values) in the order of $\times 10^{-8}$.

After the acid shock, with the cut off of a fold change equal to or greater than 2, there were 4 proteins, shown in *Figure 6.7C*, which had significantly lower abundance in B128.7.13 Δ dlp when compared to the WT. The protein that showed the lowest abundance as well as one of the most significant changes (p-value = 2.34×10^{-8}) was ferritin with a fold change of 0.32 (*Figure 6.7C*). In *H. pylori*, iron can be stored in ferritin (Pfr) and it has been shown that ferritin allows the bacteria, under severe iron starvation, to continue to grow and therefore protect the bacteria from acid-amplified iron toxicity (274). This function in iron metabolism is essential for survival in the gastric mucosa. Interestingly, Ferritin was also a significant hit found between the strains under pH 7 conditions (*Table 6.3*) and this relationship will be discussed further in *Section 6.4.1*.

When B128.7.13 Δ dlp was exposed to an acidic shock there were 12 proteins that showed a significant increase to their abundance within our criteria when compared to the WT and these can be seen in *Figure 6.7D*. Remarkably with the combination of biological triplicates, 7 of these 12 proteins are directly related to the *H. pylori* membrane in some way; 4 of these proteins are related to the flagella, Flagellin B, Flagellar hook protein FlgE (PROKKA_01342), homolog of Flagellar hook protein FlgE (PROKKA_00924) and the Flagellar M-ring protein and 3 proteins shown to have homology to outer membrane proteins, PROKKA_00299, PROKKA_01346 and PROKKA_00588. The protein with the greatest increase to its abundance was found to be Formamidase (AmiF) with a fold change of 3.4. This was closely followed by Flagellin B with a change of 3.1. Looking further into the data,

the protein partner of AmiF, AmiE also showed a significant increase to its abundance that sat just outside our criterion of a fold change greater than 2, with a change of 1.91. The presence of all these proteins in higher abundance in the B128.7.13 Δ *dlp* compared to WT after acid exposure indicates that *H. pylori* may be undergoing a stress response in that it is attempting to maintain intracellular homeostasis while being exposed to acid.

When these results are considered as a whole, it appears that B128.7.13 Δ *dlp* with a compromised membrane, as described in *Chapter V*, undergoes a significant stress response when challenged with acidic shock. The bacteria seems to be attempting to neutralise and escape the higher levels of intracellular acid with the higher abundance of flagella proteins, OMPs and AmiE and AmiF enzymes, when compared to the B128.7.13 WT.

Markerview analysis of SWATH-MS comparing proteomic changes at pH 5 between B128.7.13 WT and B128.7.13Δdlp. (a) Volcano plot of Log (fold change) in relation to the p-value of the complete quantified proteome (979 proteins). Each dot indicates a protein, coloured dots highlight tRNA proteins. (b) Volcano plot of Log (fold change) in relation to the p-value of the significantly ($p\text{-value} < 0.05$) changing quantified proteome, consisting of 170 proteins that are differently expressed. (c) 4 proteins were found to have an abundance that was equal to two-fold or less in B128.7.13Δdlp (left of the dashed red line). (d) 12 proteins were found with an abundance that was equal to twofold or higher in B128.7.13Δdlp. Adapted from the output from Markerview.

6.4 DISCUSSION

Over the past two decades Mass Spectrometry (MS) has emerged as a powerful, high throughput tool in the technology-driven field of biochemistry and biological research and even more so in the field of microbiology. For MS-based proteomic analysis there are two complemented strategies, targeted and untargeted. Untargeted proteomics aims to identify proteins in a proteome sample, without prior knowledge of which proteins may be present or changing, by digesting proteins into peptides for the highest sensitive in MS (known as “shotgun” proteomics). This “discovery” mode of untargeted analysis works best on proteins with high abundance and peptide ionisation potential, and is considered to be the most powerful method for high-throughput protein analysis, as well as maximum proteome coverage. On the other hand, targeted approaches, such as selected/multiple reaction monitoring (SRM/MRM), targets a specific protein(s) by identification of number of peptides and their corresponding fragment ions which allows for a highly specific, label-free quantification, with unparalleled reproducibility and sensitivity, at the cost of proteome coverage (275). These methods have a significant method development phase for each targeted experiment which requires prior knowledge of the proteins of interest, as such these are inherently biased strategies as unknown proteome changes will not be measured. Sequential window acquisition of all theoretical mass spectra (SWATH-MS) is essentially a large-scale MRM with a data-independent acquisition strategy and no reliance on known targets. Although prior knowledge of the sample being analysed is needed in order to build a library of all detectable peptides in the proteome, the method is unbiased in that all changes in the detectable proteome will be analysed i.e. it retains the strengths of untargeted MS methods. SWATH-MS allows a whole proteome to be reliably quantified in a label-free manner without extensive prior sample characterisation and for this reason was chosen for this study (258).

SWATH-MS relies on the building of comprehensive spectral libraries by data dependent approaches for peptide feature extraction and thus identification and subsequent quantification. Therefore, the quality and coverage of the proteome of interest in these libraries is crucial to the method (254). In this study, a comprehensive coverage of the

B128.7.13 predicted proteome was achieved, with 72% of the predicted proteome (1197 proteins identified/1655 predicted open reading frames) being identified by at least 2 unique peptides and a 95% confidence for identification, which can be increased by additional 50 proteins if proteins identified by 1 unique peptide are included (*Figure 6.1*). To date this is one of the highest coverages of a *H. pylori* proteome for quantification studies, with 72% previously achieved in another *H. pylori* strain 26695 in iTRAQ studies (276). Furthermore, from this data set 60% of the proteome was able to be quantified using SWATH, and again this is one of the highest quantifications to date of a *H. pylori* proteome, with the two previous studies achieving 60% and 46% respectively (263, 276). In the PROKKA annotation of B128.7.13 there may have been an overestimation of the predicted ORFs, with not all predicted ORFs encoding for a protein and for this reason the coverage and quantification found in this study may be under estimated. In the B128.7.13 ion library, key virulence factors and known acid adaption proteins were captured, including the components of the flagella machinery, proteins involved in the hook, the basal body and the filament. Along with the flagella coverage, the urease subunits and accessory proteins as well as the chemotaxis proteins used in acid sensing among other things, including CheA, CheV and CheW, were identified. The presence of these proteins allowed the known acid adaption changes to be monitored quite closely (*see S7a for ion library*).

Conversely, there are approximately 300 predicted proteins missing from the ion library. However, over 50% of them were classified as hypothetical by PROKKA, meaning that their presence in the proteome has not been confirmed. For this study, the proteins of interest, Dlp1 and Dlp2 were not able to be identified by DDA MS and therefore they are not in the ion library. This is likely due to low abundance under the conditions tested or poor peptide ionisation potential, as their expression has been detected by Western blot in the whole cell lysate previously (*Chapter IV*). It is well established that Western blotting is more sensitive in the detection of proteins in a whole cell lysate when compared to the “shotgun” proteomics approaches used to create the ion library due to signal amplification which is not possible in MS (275). Unfortunately, the lack of relevant peptides in the ion library does mean that SWATH-MS is not able to monitor any changes in abundance of the dynamin proteins themselves, specifically in the WT during acid exposure. However, the absence of these

proteins in the ion library does not inhibit the ability of SWATH-MS to measure proteome disturbances resulting from deletion of these genes upon exposure to acid stress, with WGS confirming the absence of the operon in the B128.7.13 Δ *dlp* strain previously (*Chapter V*). The failure to detect these proteins in DDA MS under all of these conditions suggests that their abundance is remaining below the limit of detection and is not changing dramatically between conditions. Further exploration of the different conditions where the *dlp* operon may be upregulated enough for identification by DDA MS, may enable the addition of the dynamin proteins to the ion library in the future. This is another strength of the SWATH-MS methodology, as the ion library is able to be continuously expanded as more data becomes available and previously acquired SWATH data, such as this study, can be reanalysed using the updated library to ascertain if the proteins are present and able to be quantified.

As shown in the results (*Section 6.3*), 60% of the proteome was able to be quantified in this study; with peptide evidence for 979 of 1655 ORFs identified and quantified using SWATH-MS (*see S7 for all quantification data*). This is one of the highest levels of quantification to date that we are aware of for *H. pylori* using whole cell proteomics. It is also amongst one of the highest rates of proteome quantification for any bacteria with the highest that we are aware of being 70% in *Pseudomonas aeruginosa* (262). The experimental design allowed a number of different comparisons to be made to highlight and observe key differences in the proteomes between strains and between conditions (*Figure 6.3*). Using a supervised PCA it was very clear that the 3 biological and 4 technical replicates in both strain and acid tests were reproducible and discrete (*Figure 6.2*), indicating a high level of confidence in the results obtained in this global quantification of the proteome. All relationships were analysed using a t-test statistical analysis, and then restricted to all proteins that were identified as significantly changing in their abundance given a p-value of equal to or less than 0.05 and a fold change greater than or equal to 2 ($\log_{10} 0.3$) or less than or equal to 0.5 ($-\log_{10} 0.3$) in the case of reduction of abundance. This stringent criterion allowed us to restrict our data set to focus on the most significantly changing proteins under these experimental conditions. However, despite the implementation of these stringent criteria on the data set, all proteins identified by SWATH-MS could be used to look at the global changes to the proteome, with changes considered insignificant statistically by Markerview potentially providing more

information on a pathway that may have significant changes occurring to it under these conditions. For the purposes of this discussion, and due to the extremely large data set generated, not all significantly changing proteins will be discussed. To address the aims of this thesis, the analysis has been focused on the proteins of most biological significance in terms of *H. pylori*'s acid tolerance mechanisms and dynamin proteins.

6.4.1 PROTEOMIC DIFFERENCES BETWEEN B128.7.13 WT AND B128.7.13 Δ DLP

A comparison between B128.7.13 WT and B128.7.13 Δ dlp under neutral conditions helped elucidate the changes that occurred in normal growth in the presence and absence of the dynamin operon. This formed a baseline for investigating the potential function, if any, dynamins have in normal growth. As seen in *Table 6.3*, there are a number of changes to the proteome in the absence of dynamin under these conditions. Most notably is the significantly lower abundance of three proteins of the cytochrome superfamily, Cytochrome c-553 precursor, Cytochrome c551 peroxidase precursor and Cbb3-type cytochrome c oxidase subunit CcoP. In Gram-negative bacteria, these are integral membrane proteins, that have a range of roles linked with iron but most notably in *H. pylori* they are thought to be expressed in response to lower oxygen tensions, allowing it to survive and proliferate under microaerophilic conditions (277). They are also known to be involved in the storage electrons in the periplasmic space (273). Therefore, their two-fold lower abundance in B128.7.13 Δ dlp under neutral conditions is quite perplexing with no obvious link to previously established dynamin functions. This change could be a by-product of the fragile membrane of the mutant and they are simply leaking out of the cell into the supernatant and therefore would not have been captured in the SWATH analysis which was performed on the whole cell. However, this would also suggest the mutant is also osmotically compromised and obviously the export/exchange of proteins is much harder than ions or simple sugars etc. and usually requires a transporter of some sort, whether passive or complex self-transporter. Therefore, the absence of dynamin may be also having a direct or regulatory effect on these pathways, but targeted analysis would be required in order to determine these types of dynamics or potential interactions.

The protein that had the greatest abundance reduction in the absence of dynamin under neutral conditions compared to WT was Endonuclease MutS2 with a log fold change of -0.5 (Table 6.3). Endonuclease MutS2 is involved in the suppression of homologous recombination and therefore has a key role in the control of bacterial genetic diversity in *H. pylori* (278). Quite interestingly, it was noted in Chapter V, that the deletion of the *dlp* operon in both 26695 and SS1 was attempted with no success, but the transformation protocol was proven effective with the complementation of the B128.7.13 mutant strain being highly successful. The SWATH data suggests that when dynamin was deleted this protein was dramatically affected, and so it could be possible that MutS2 may have been a contributing factor to the difficulty in producing a *dlp* knockout transformation in *H. pylori* by binding to the recombinant DNA and stopping incorporation into the chromosome (267). Notably MutS2 was observed with a significantly lower abundance when the WT was exposed to acid (Table 6.1, Figure 6.4). Therefore, it may be speculated that the mutagenesis of *H. pylori* under acidic conditions may aid the incorporation of DNA into the chromosome, although this hypothesis would need controlled experimentation to draw this conclusion.

Proteins found in significantly higher abundance in the absence of dynamin under neutral conditions were comprised mainly of hypothetical proteins (Table 6.3). This data highlights that these proteins are being expressed in *H. pylori* and therefore can no longer be considered hypothetical. BLAST matches identified them as a DUF262 domain containing protein, SAM dependent methyltransferase and an AANH-like protein, suggestive of their functions in *H. pylori*. Considering they are significantly affected by the absence of dynamin, they warrant further investigation to determine any link to the general function of dynamins in *H. pylori*. The protein found to have the most significantly reduced abundance in B128.7.13 Δ *dlp* under acid stress was Ferritin (p-value = 2.3×10^{-8}). However, upon closer inspection it can be seen that this was also seen to be lower in the comparison of the pH conditions in B128.7.13 Δ *dlp*, implying this protein isn't exclusively changing in response to acidic conditions but rather that the lower abundance is correlated with the absence of dynamin (Table 6.2). Previously, it has been found that HP-Ferritin may have a dual function by providing an iron-storage role with ferroxidase activity, as well as protecting chromosomal DNA from the increase in free radicals that target DNA under acid stress (279#, 280). These functions are controlled by the

binding of the Fur repressor to the Ferritin promotor (274). Lower abundance of Ferritin may indicate that the cells are attempting to secure the availability of free iron in the cytoplasm, especially if for some reason iron is scarce in the absence of dynamin (281). Furthermore, this SWATH data suggests that in B128.7.13 Δ *dlp* the Fur repressor potentially hasn't reduced its binding to the HP-Ferritin promoter under the mild acid stress, which it would normally do (274). Under neutral conditions the repressor should bind to the HP-Ferritin promoter stopping protein production, but for some reason in B128.7.13 WT this doesn't seem to be occurring, with higher levels of Ferritin in the proteome compared to B128.7.13 Δ *dlp*. On the other hand, normal levels and activity of ferritin are observed in the WT and there is a strong repression of the production of Ferritin in the absence of dynamin, regardless of growth conditions. This second hypothesis is supported by the ion detection levels detected for ferritin during SWATH-MS between the sample groups (*Figure 6.8*).

These conclusions warrant further investigation into the mechanism by which B128.7.13 WT carries out iron storage both under normal and stress conditions, as there is a well-established link between metal responsiveness and acid tolerance in *H. pylori*, with a potential further function for HP-Ferritin that is dependent on the presence of dynamins (282).

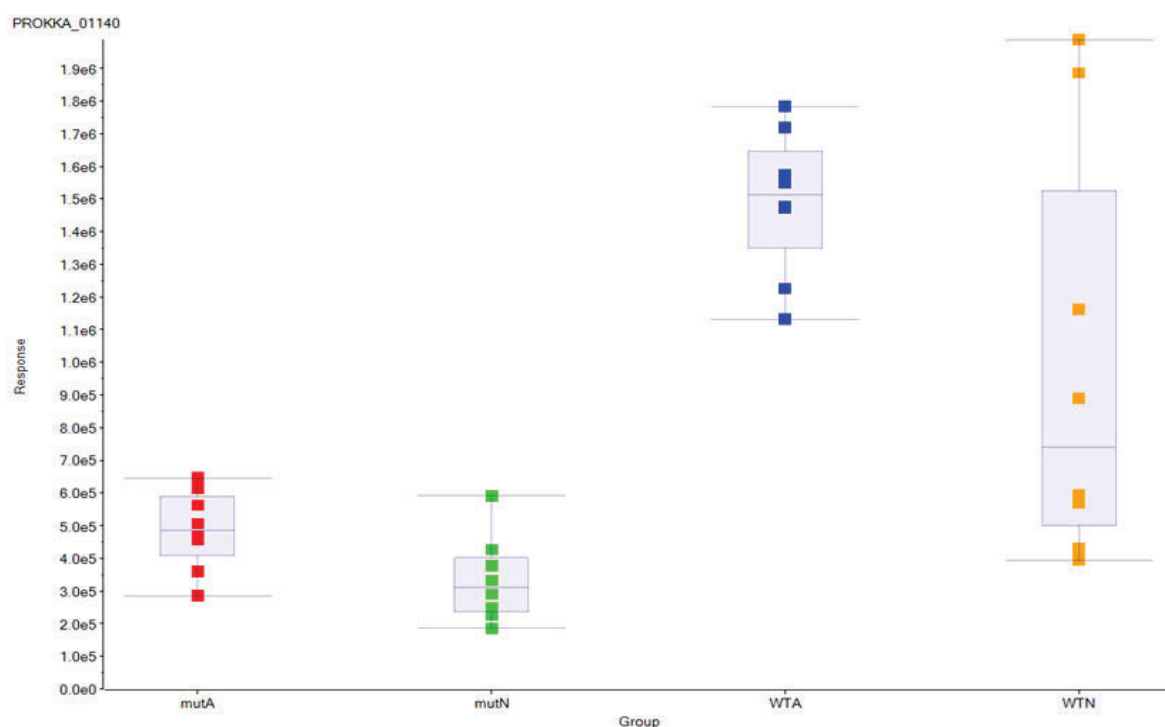


Figure 6.8 The Ion Intensity for PROKKA_01140 – Ferritin

Box and whisker plots generated by Markerview show the response (ion intensity) for each of the samples, mutA - B128.7.13 Δ dlp acid treated (red), mutN - B128.7.13 Δ dlp no treatment (green), WTA - B128.7.13 WT acid treated (blue) and WTN - B128.7.13 WT no treatment (orange) for 4 replicates. There is a clear reduction in the mutant groups of the response detected for Ferritin when compared to the WT. The response is indicative of the ion intensity, relative to the amount of a protein in a sample. There is a clear increase in the variability in ferritin in the WT strain compared to the mutant

Also in this data set setting the benchmark for changes between B128.7.13 WT and B128.7.13 Δ *dlp* are the ammonia producing enzymes, AmiE and AmiF. The SWATH data showed that these proteins are in higher abundance with fold changes of 3.1 and 2 in B128.7.13 Δ *dlp* under acidic conditions compared to WT (Table 6.4). This is quite interesting as they are also significantly changing in response to the treatment of just B128.7.13 Δ *dlp* with a similar fold change (Table 6.2). Similar to Ferritin, these proteins are more likely to be strain dependant changes rather than condition specific. *H. pylori* is able to survive long-term exposure in the gastric lumen where the pH is quite variable, with the bacterium's response to this being largely thought to be multifactorial but is still poorly understood. As seen in Figure 6.9, as hydrogen ions pass through the membrane when *H. pylori* is in an acidic environment, the two-component transcriptional regulator system ArsS and ArsR, mentioned previously (Chapter V) is activated. From this there are numerous downstream effects such as the activation of the urease pathway for acid neutralisation in the periplasm and the extracellular milieu (283). The activation of *amiE* and *amiF* by this system, to produce ammonia in the cytoplasm of the cell to maintain internal homeostasis in the presence of a stress like acid, is of utmost importance to this study. Both AmiE and AmiF showed a significant increase in their abundance in the absence of dynamin, suggesting that with a compromised membrane, B128.7.13 Δ *dlp* has a disruption to the cytoplasmic pH and the bacteria are responding by producing more AmiE and AmiF to try and maintain homeostasis. Quite interestingly, as previously mentioned in Chapter V, this two-component system has been shown to activate the *dlp* operon (in another strain) in response to acid (pH 5) (284). Therefore, additional research into the link between this transcriptional regulating system and dynamin is needed to determine whether there is an impact on the expression of other genes under the same condition, like *amiE* and *amiF*.

Overall, the changes observed between B128.7.13 WT and B128.7.13 Δ *dlp* bring to light that there are a number of changes to the proteome in response to the absence of dynamin which defines the baseline to begin analysing the role of dynamins in acidic stress conditions.

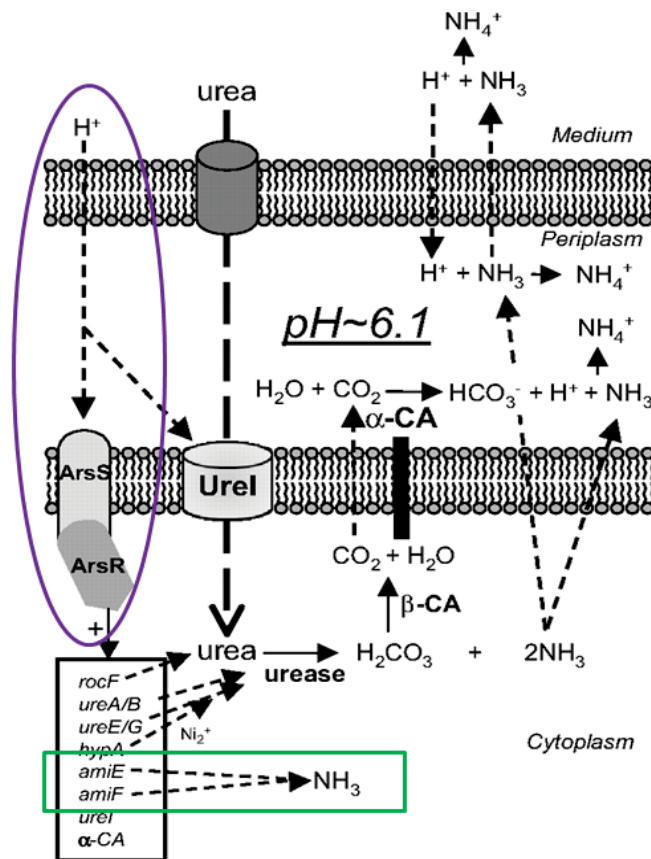


Figure 6.9 Acid Tolerance Transcriptional Regulation by the ArsS/R system

The acid acclimation genes that are up-regulated by ArsR (HP0166) are shown in the box. At pH 6.1, the ArsRS two-component system (circled in purple) is also activated and results in the up-regulation of the genes, *ureE* and *ureG*, with both contributing to the increase in urease activity by converting the apoenzyme into active urease. The potential ammonia producing enzymes, *amiE* and *amiF* (green box), and the urea producing *rocF* (arginase) are also up-regulated by this two-component system, as is carbonic anhydrase (α -CA) and *ureI*. Hydrolysis of urea results in the production of H_2CO_3 and 2NH_3 . The former is converted into CO_2 and H_2O by the cytoplasmic α -CA and both the CO_2 and NH_3 diffuse rapidly through the inner membrane into the periplasm. There, the CO_2 is converted by α -CA into HCO_3^- and H^+ . The bicarbonate buffers the periplasm, maintaining periplasmic pH at 6.1. Cytoplasmic NH_3 can neutralize protons entering the cytoplasm. In this SWATH analysis, both *amiE* and *amiF* were found to be in higher abundance in the bacteria under acid shock in the absence of dynamin. Adapted from (283).

6.4.2 PROTEOMIC CHANGES IN B128.7.13 WT UPON ACID SHOCK

Acid tolerance is an important feature for bacteria both Gram-positive and Gram-negative. Many environments that these bacteria encounter have changing pHs, whether it be in a human host or an environmental source such as food for human consumption. For example it has been shown that *Salmonella spp* as well as *E. coli* O157:H7 can survive in acidic environments they encounter, such as in food (285-287). The way that Gram-positive bacteria overcome acid stress is quite interesting and many parallels can be drawn to how Gram-negatives, more specifically *H. pylori*, overcome this stress. The general mechanisms of acid tolerance include; proton pumps to bring about an increase in internal pH, repair systems that involve chaperones, proteases, and heat shock proteins that results in the protection of proteins as well the repair of DNA damaged due to the low internal pH. Furthermore, cell density affects cell-to-cell communication, cell envelope alterations can enhance protection of the cells by changing its architecture, composition, stability, and activity and finally the production of alkali by a system such as the urease system increasing the pH of the environment in direct proximity to the cell (288). Therefore, it can be seen that there are a number of survival strategies that the cell can utilise when exposed to a stress such as acid (289). Specifically for *H. pylori*, it has adapted to this stress and is able to successfully colonise the stomach, through the development of a flagella for motility, chemotaxis for acid sensing and an extremely well controlled urease pathway (122). When analysing the SWATH data for both B128.7.13 WT and B128.7.13 Δ dlp, all of these pathways were considered, as this is the first investigation into the B128.7.13 WT acid stress response.

Looking at the response of strains to an acidic environment brought to light the first evidence that B128.7.13 Δ dlp responds differently to acid stress compared to the WT. When WT was exposed to an acid shock, there were not many proteins that changed in their abundance within the criteria, with only PROKKA_00773 (CMP-N,N'-diacetyllegionaminic acid synthase) and PROKKA_00934 (D-glycero-alpha-D-manno-heptose-1,7-bisphosphate 7-phosphatase) showing a significant p-value as well as a fold change greater than two fold (Table 6.1). CMP-N,N'-diacetyllegionaminic acid synthase has been shown to be a key player in the production of CMP-legionaminic acid and this is thought to assist in the regulation

of O-linked flagellin glycosylation in *Campylobacter* (290). This glycosylation is essential for the assembly of the flagella filament, one of the ways *H. pylori* adapts to an acidic environment (291, 292). Therefore, an increase in its abundance is not a surprise, as it is an initial response by *H. pylori* to begin to prepare for acid stress.

A surprising observation was that the abundance of many of the proteins in the urease pathway in the WT did not change when exposed to acid, with only UreA, the alpha subunit of the urease enzyme showing an increased abundance of approximately two-fold (*Figure 6.4*). It is well-known that *H. pylori* synthesizes constitutively large quantities of the dimeric urease enzyme (UreA and UreB) with optimum production and activity of these proteins at neutral pH (293, 294). Quite staggering is that the amount of the UreA and UreB proteins produced by *H. pylori* accounts for between 10 to 15% of the total protein content of the bacteria (293, 294). Therefore, it is possible that small changes to the abundance of these proteins would not be considered significant in the context of the global proteome or that production of additional enzyme would not result in an increase in enzymatic activity. This could be because other mechanisms of acid tolerance may be preferred under these types of acid conditions. Within the urease gene cluster an essential gene encoding an acid-activated urea channel, UreI is well established now to be the protein that responds to the external acidic environment (137, 295, 296). The neutral pH for optimum production of urease allows the cytoplasm to remain relatively neutral while the organism protects itself against acid by enhancing access of urea to the internal urease via the UreI channel and by doing so achieves buffering of the periplasm (293). Therefore, it could be hypothesised that not much change to the abundance of the proteins in this pathway would be expected, with an acidic shock of pH 5 not harsh enough to cause significant upregulation to this pathway, with a high enough abundance of these proteins already present to combat the acid when supplemented with 10 mM urea.

This characterisation of the WT acid response in B128.7.13 specifically, has not previously been noted in the literature and from this SWATH analysis, under these conditions of pH 5, it seems that there are very small changes to the proteome. Further investigation using

multiple time points and multiple pH media would allow a more comprehensive look at the acidic response of B128.7.13.

6.4.3 THE SHARED ACID RESPONSE

There were only two proteins in common that showed significant changes in both strains when pH 7 and pH 5 were compared. They were a hypothetical protein PROKKA_01248 and an outer membrane protein PROKKA_00299. For both WT and B128.7.13 Δ dlp, PROKKA_01248 showed a higher abundance when the strains were exposed to acid with fold changes of 1.65 and 1.3 respectively. This indicates that this protein, which shows homology to Rho, a transcriptional terminator factor, firstly is not hypothetical and secondly is changing in response to the condition, despite the difference between strains. However, PROKKA_00299, an outer membrane protein, showed opposite responses in the deletion strain. In B128.7.13 Δ dlp PROKKA_00299 showed higher abundance when the strain was exposed to acid with a significant fold change of 1.46 (p-value = 0.038) where as in WT it showed lower abundance when the strain was exposed to same condition with a significant fold change of 0.79 (p-value = 0.026). This is quite interesting and further discussion of OMPs in B128.7.13 Δ dlp is seen below.

6.4.4 THE B128.7.13 Δ DLP MEMBRANE IS SIGNIFICANTLY ALTERED IN RESPONSE TO ACID SHOCK

The key comparison in this SWATH analysis was between the strains under acidic conditions, highlighting the changes in the proteome in the presence and absence of dynamin when the bacteria were put under acid stress. Specifically, it is this comparison that allows speculation of the role of dynamin in acid tolerance of *H. pylori*.

As seen in *Table 6.4*, there were not many notable changes of significance between the strains for proteins with lower abundance in B128.7.13 Δ dlp, with the most interesting findings seen in the proteins that were found to have a higher abundance in the proteome of B128.7.13 Δ dlp after acid shock. This is not surprising as we can expect there to be more proteins changing

in response to this stressor to try and overcome it; as more peptides will be present to perform more accurate quantitation. Sampling more peptides results in a higher intensity of ion signal and a greater number of measurements and this in turn increases confidence in the normalisation of the data and allows a better overall quantification of the proteome. Two important sub categories of these proteins which were found in higher abundance will be discussed; flagella associated proteins and outer membrane proteins (OMPs) (*Table 6.4*). Both of these groups of proteins have been shown previously to be involved in urease-independent mechanisms for acid tolerance of *H. pylori* (122).

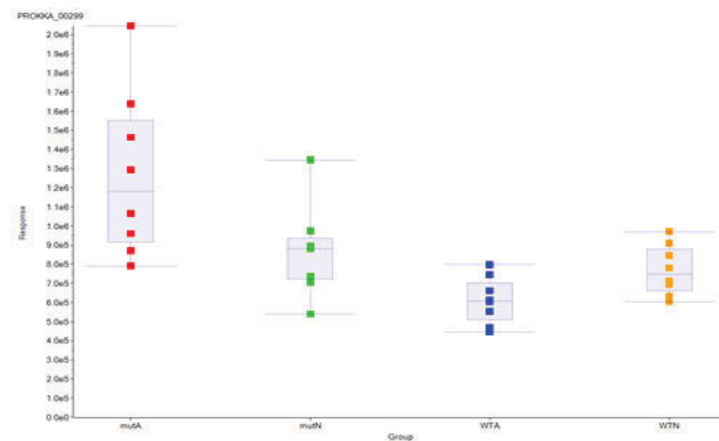
Outer membrane proteins (OMPs) of Gram-negative bacteria are important molecules that form the crucial interface between the cell and the environment. The initial point of contact for *H. pylori* with a host cell is through the outer membrane adhesin, BabA (155, 156). This then triggers an inflammatory cascade, that allows the following adhesin SabA to bind to the host cell to cause further damage (159). Furthermore, both OMPs are known to respond to an acid challenge (154). Unfortunately, in this SWATH analysis SabA was not captured in the ion library and therefore was not able to be investigated. Interestingly, BabA (PROKKA_0434) was significantly lower in abundance in B128.7.13 Δ dlp when compared to WT under both neutral (p-value 0.008, log fold change -0.19) and acid conditions (p-value 0.00073, log fold change -0.11) indicating this reduction was due to the absence of dynamin rather than the acid challenge. Adhesins, particularly BabA are embedded into the membrane and, so it is not surprising to see a reduction of BabA in B128.7.13 Δ dlp, a strain previously characterised with a compromised membrane. In B128.7.13 Δ dlp BabA would thus be expected to have a limited association with the membrane due to the reduction in membrane integrity.

Porins are a specific family of OMPs that form a hydrophilic channel, allowing nonspecific diffusion of small molecules, like urea and antibiotics, across the outer membrane (297, 298). Expression of outer membrane proteins is affected by pH, with the mechanism for pH-dependent porin expression still enigmatic in the literature (87, 299, 300). In *H. pylori*, OMPs play an important role in acid acclimatisation, with the movement of urea into the cell through porins that span the outer membrane for neutralisation of acid in the periplasmic space (8,

295). In addition to this, porins play a vital role in allowing the membrane to be adaptable to combat a range of stressors to ensure the survival of the cell and are therefore important in membrane integrity (8). In contrast to BabA, in this SWATH analysis when exposed to acid two OMPs in B128.7.13 Δ dlp are found in higher abundance, both with a significant fold change of approximately 2 ($\log_{10}0.3$) (Table 6.4). One of these OMPs, PROKKA_00299, as previously stated is the only protein that showed a significantly lower abundance in the WT in response to acid and a significantly higher abundance in B128.7.13 Δ dlp when exposed to the same condition and that can be seen in Figure 6.10. An increase in the overall abundance when B128.7.13 Δ dlp is exposed to acid is seen and interestingly, this same trend is also seen for the other OMP identified, PROKKA_01348. Both proteins show low and reproducible responses in the WT under both conditions, whereas there is a larger and more varied response in the mutant (Figure 6.10). The variation of the protein abundance amongst samples for both OMPs is further exacerbated when the mutant is exposed to acid. The higher abundance of these OMPs could be the result of a stress response that B128.7.13 Δ dlp is implementing to overcome two things; a compromised membrane in the absence of dynamin and a second, even greater increase to overcome the acid stress with the same compromised membrane.

(A)

PROKKA_00299



(B)

PROKKA_01348

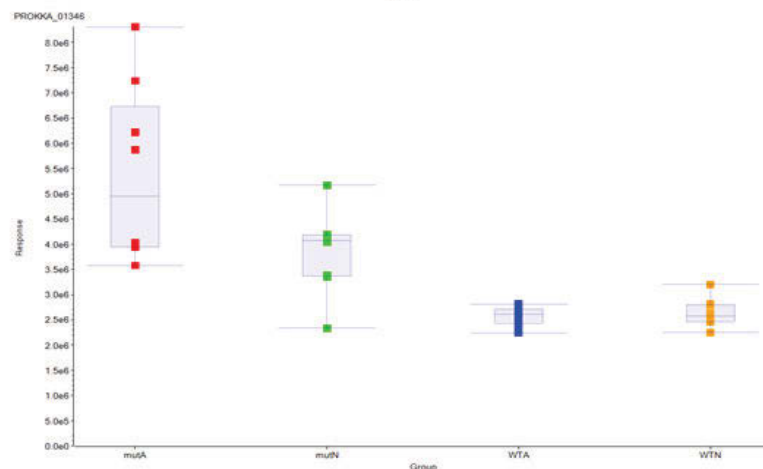


Figure 6.10 The Ion Intensity for OMPs PROKKA_00299) and PROKKA_01348

Box and whisker plots generated by Markerview show the response (ion abundance) for each of the samples for both PROKKA_00299 (A) and PROKKA_01348 (B). Where, *mutA* - B128.7.13 Δ dlp acid treated (red), *mutN* - B128.7.13 Δ dlp no treatment (green), *WTA* - B128.7.13 WT acid treated (blue) and *WTN* - B128.7.13 WT no treatment (orange). There is a clear increase in the mutant groups of the response (ion intensity) detected for both of these OMPs when compared to the WT under the same conditions. As well as an increase in the variability of the ion response, indicating a possible change in expression to these OMPs in the mutant strain especially under acid stress.

The other group of proteins that showed the most significant change to their abundance in the absence of dynamin when under acid stress were related to the flagella machinery (*Table 6.4*). Firstly, there is a strikingly different protein response in B128.7.13 Δ *dlp* compared to WT with the presence of two members of the flagella machinery when the growth in pH 7 and pH 5 was compared for the mutant strain (*Table 6.2*). This was the first evidence that the absence of dynamins from B128.7.13 was having an effect on the way *H. pylori* overcame the acid stress. The most significant change to B128.7.13 Δ *dlp* when neutral and acidic conditions were compared (*Table 6.2*) was the abundance of Flagellin B, with a fold change of 2.1636 and a p-value of 1.4×10^{-5} . Flagellin B (FlaB) is an integral part of the flagellar filament, responsible for motility in *H. pylori* (301). In this dataset along with FlaB, another protein involved in the flagella machinery, FlgE, the flagella hook protein, showed a significantly higher abundance in the proteome when B128.7.13 Δ *dlp* was exposed to acid. Taken together, the significant increase in the abundance of these proteins in B128.7.13 Δ *dlp* suggests that *H. pylori* in the absence of dynamin is attempting to escape the acid by becoming more motile, a known mechanism for acid adaption in *H. pylori*. (122, 252).

Motility has been implicated as a key mechanism that *H. pylori* utilises to escape acid shock as quickly as possible, with the flagella machinery essential for colonisation and subsequent infection by the bacterium (122, 301). In *H. pylori*, like most enteric bacteria, the flagella machinery consists of three structural elements; a basal body, which is internal and contains the proteins needed for rotation and chemotaxis, an external filament, which acts as a propeller, and a hook that links the basal body to the filament. Over 50 proteins are predicted to be involved in the expression, secretion and assembly of this complex with nearly half to be involved in the structural components (143). In this SWATH analysis, 20 of these proteins were identified and quantified, with 6 of them showing a significant change (4 of which showed a fold change greater than 2) when B128.7.13 Δ *dlp* was exposed to acid compared to the WT (*see S7j for all flagella proteins quantified*). These proteins were; Flagellin B (FlaB), Flagellar hook protein E (FlgE), homolog of FlgE, Flagella M-ring protein (FliF), Flagella motor switch protein (FliM) and flagellar assembly protein H (FliH). As it can be seen in *Figure 6.11*, the proteins identified as having a significant fold change when B128.7.13 Δ *dlp* is exposed to acid are spread throughout the flagella machinery, with FlaB and FlgE showing

the greatest fold change of 3.1 and 2.9 respectively (*Table 6.4*). Further interrogation of the 20 flagella proteins quantified by SWATH can classify them into different structural components, basal body, biosynthetic, filament, hook and unknown. From these, the flagella proteins that make up the filament and hook are all found in higher abundance, with mostly the flagella proteins of unknown function found in lower abundance (*see S7j*).

The higher abundance of these proteins strongly supports the notion that B128.7.13 Δ *dlp* is attempting to escape the acidic environment by increasing its ability to be motile, potentially because acid is making its way into the cell, due to a compromised membrane, at a faster rate than it can neutralise it. In order to test the motility concept of B128.7.13 Δ *dlp*, it could be hypothesised that in the presence of acid, the mutant strain will be hyper motile compared to the WT, which would require a separate set of experiments.

In summary, this is the first study investigating acid tolerance in the gerbil adapted strain B128.7.13. It is to date, one of the most comprehensive ion libraries and proteomes established for *H. pylori* with approximately 72% coverage (minimum two peptides per protein). It is also one of the highest proteome quantification studies for the *H. pylori* proteome by a non-labelled strategy with 60% using SWATH-MS. Furthermore, SWATH-MS has demonstrated that dynamins are having significant effects on the global proteome of *H. pylori* both in normal growth as well as under acid stress. Particularly in acidic conditions, the absence of dynamin causes *H. pylori* to seemingly undergo a strong stress response relying on urease-independent acid tolerance mechanisms, such as increased motility, AmiE/AmiF ammonia production and increased production of OMPs. This is potentially to compensate for the increased stress that the cell experiences due to the reduction of membrane integrity that comes with the deletion of the *dlp* operon. SWATH-MS is a powerful and versatile tool with the ability for continual addition of spectra of peptides representing proteins to the ion library, and ongoing testing of multiple new conditions or reanalysis of previously acquired data to address new hypotheses. This study will therefore form a strong foundation for future research in B128.7.13. Additionally, this study represents the beginning of future research into the function of dynamins in *H. pylori* with a more targeted approach to understanding their possible function(s).

CHAPTER VII

GENERAL DISCUSSION

At the commencement of the work presented in this thesis, there was minimal research into the overwhelming presence of putative dynamins in the bacterial genome (6). Furthermore, there was very limited research that had investigated the function of these important membrane remodelling proteins in bacteria (5, 93, 97, 98, 100) and this has not changed significantly at the time of writing. Recently, a previously characterised ETEC virulence factor implicated in toxin release via vesicle formation, LeoA, was identified as a bacterial dynamin-like protein (5). It was also found that upstream of LeoA there was another dynamin homolog encoded by two genes, *leoC* and *leoB* (5). The closest homologs to the *leoCBA* operon were identified in *H. pylori* but to date these have not been investigated.

The overarching goal of this thesis was to investigate the presence and functions of dynamin-like proteins (DLPs) in *H. pylori*. *H. pylori* is an important and surprisingly understudied human pathogen that is known, like many other bacteria, to be dependent on membrane dynamics for virulence (8-10). It is also a pathogen that survives and thrives in one of the harshest environments, the acid environment of the human stomach, where membrane integrity is essential to its protection.

The investigation into the presence of DLPs in *H. pylori* undertaken in this thesis was achieved by addressing a series of aims consisting of a bioinformatic survey to determine the conservation of the DLP genes in the *H. pylori* genome (Aim 1), research into the expression and biochemical features of the DLPs (Aim 2) and finally their potential function in *H. pylori* using knock out strains in both a direct phenotype approach and global proteomic exploration (Aim 3).

By achieving each of these aims, this has been the first study to confirm the presence of dynamins in *H. pylori* as well as being the first investigation into the function of DLPs in *H. pylori*. As such this thesis has contributed significantly to understanding both the role of bacterial dynamins in this important human pathogen, as well as showcasing the diversity of functions that bacterial dynamins possess, similar to their eukaryotic counterparts.

7.1 THE INS AND OUTS OF DYNAMIN-LIKE PROTEINS IN *H. PYLORI*

7.1.1 AIM I: TO INVESTIGATE THE CONSERVATION OF DYNAMIN-LIKE PROTEINS IN *H. PYLORI*

It has been previously noted that the closest homologs to the LeoCBA operon were in *H. pylori*, however the extent of this conservation was not explored (5, 99). Through the use of psi-BLASTp and JackHMMER iterative searches, the level of conservation was demonstrated, with homologs found across the bacterial phyla and specifically, homologs of LeoCBA detected in over 90% of *H. pylori* strains (249 strains) screened (Section 3.3.2). To date, *H. pylori* displays the highest level of reported conservation of DLPs compared to any other bacteria studied, making them the ideal model organism to study these genes. It also may indicate that *H. pylori* has retained these genes through evolution because they serve an important function to the survival of the bacterium. Studying the ETEC LeoCBA homologs in the context of *H. pylori*, may help elucidate the fundamental function they have, with only 3% of ETEC genomes encoding LeoCBA (5). Of the 260 *H. pylori* strains screened, 136 were found to have LeoCBA homologs occurring sequentially in their genome, a previously seen genomic feature of bacterial dynamins that may indicate functional interaction (6).

Overall, this aim of the thesis significantly added to our understanding of the depth of conservation of dynamins in the greater bacterial phyla, which are worth pursuing in the future as well as establishing never before seen conservation of the dynamin superfamily across a wide variety of strains in a single bacterial species.

7.1.2 AIM II: TO CHARACTERISE THE EXPRESSION AND BIOCHEMICAL FEATURES, SUCH AS LOCALISATION, LIPID BINDING AND INTERACTIONS OF DYNAMIN-LIKE PROTEINS FROM H. PYLORI

To embark on the investigation into bacterial dynamins in *H. pylori*, the presence of a gene product needed to first be confirmed as they remain termed ‘hypothetical’ in the NCBI database for the *H. pylori* strain 26695. As seen in Chapters III and IV respectively, RT-PCR and Western blot both confirmed the expression and translation of these genes. From this, the genes were termed *dlp1a*, *dlp1b* and *dlp2*, respective to their ETEC Leo homologs, *leoC*, *leoB* and *leoA*. Intriguingly, a genotypic phase variation was seen in the region of overlap in *dlp1a* and *dlp1b*, with a single nucleotide insertion placing the genes into the same reading-frame, forming one full-length dynamin, termed *dlp1* (Section 3.3.3). Further support of this was seen *in vivo* where evidence of both *dlp1a* and *dlp1* proteins was seen in the whole cell lysate for two *H. pylori* isolates of 26695 as well as another strain, B128.7.13. (Section 4.3.3). If phase variation was confirmed to be a mechanism for regulating *dlp1* function, this would be a first for a dynamin operon. In support of this, there has been a primary transcriptome published for the *H. pylori* strain of interest, 26695, in which they found *dlp1a*, *dlp1b* and *dlp2* to form an operon, confirming what we see in this work (Section 3.3.3 and Section 4.3.3) (192). Since this study, a new genome for 26695 has been published whereby *dlp1a* does not exist and hence since has been removed from the NCBI database (189, 190). This has also been seen in this work (Section 3.3.3 and Section 4.3.3). With these two conflicting genomes in terms of the *dlp* operon being seen in this work, it warrants further investigation to explore the possibility of phase variation regulating the transcription of a full-length dynamin, *dlp1*. In this thesis, although phase variation was not explored experimentally due to it being outside the scope of this initial study, it warrants further investigation using a technique such as RNA-seq, a well-established method for quantifying mRNA transcripts in a whole genome capacity (302, 303). If phase variation was proven in the *H. pylori dlp* operon this firstly, would be a key finding to deepen our understanding of gene regulation within the greater dynamin family, and more importantly it would bring to light that the split Leo-Type dynamins (*dlp1a* and *dlp1b*) are not necessarily the functional proteins. However, this is only

speculation as it has not been explored in any of the other known split Leo-Type bacterial dynamins previously.

With the presence and expression of DLPs in *H. pylori* confirmed, their dynamin features needed to be investigated. There are several fundamental biochemical features that classify a protein as a dynamin superfamily member including; domain architecture, GTPase activity, lipid binding and self-interaction (1, 2, 33). In this thesis, a preliminary investigation was conducted to identify and confirm these genes as encoding genuine dynamin-like proteins and then investigate some of their biochemical properties both *in vitro* and *in vivo*. With the use of bioinformatics tools, predications were made of the structure of Dlp1 and Dlp2, which not only highlighted the likely structural similarities to characterised dynamins, like LeoA and BDLP, but also showed that the sequence contains potential GTP binding motifs and lipid binding capabilities (Chapter III). Using purified polyhistidine fusion proteins a number of features were investigated, such as *in vivo* localisation (Section 4.3.4), effects on liposome structure (Section 4.3.5) as well as protein-protein interactions (Section 4.3.6).

A prediction for the primary cellular location for a membrane remodelling protein, such as dynamin, might logically be the membrane but this is not always the case. LeoA was found in the periplasm of *E. coli* whereas in *B. subtilis* only one of the DynA subunits (DynA1) was membrane associated when studied on its own (3, 5). Interestingly though, the full-length DynA localisation was previously found to be membrane-associated, indicating that one subunit was sufficient to localise the full protein to its site of function (3). A hypothesis could be formed for *H. pylori* dynamins, in which Dlp1 and Dlp2 function together to carry out membrane remodelling as a complex. Dlp1 was predicted to have a transmembrane domain and was identified in the membrane fraction by Western blot (Section 4.3.4). Potential interactions between Dlp1 and Dlp2, might explain the mixed distribution of Dlp2 between the soluble and membrane fractions (Section 4.3.4), given the prediction of Dlp2 to be a soluble protein. In this situation, Dlp1 may act as the predominant membrane-bound form, and reversibly associates with Dlp2, bringing a fraction of the cellular Dlp2 into the membrane fraction.

To begin to investigate such possible interactions directly, the *H. pylori* Dlp1s were purified for a preliminary series of binding and lipid modification studies. However, due to the insolubility of Dlp1a and Dlp1b *in vitro*, lipid binding experiments could only be performed with Dlp2; however, with Dlp1 being the protein with the predicted transmembrane (TM) domain, these experiments need to be performed in the future. Unexpectedly, reproducible liposome aggregation in the presence of Dlp2 and GTP was observed in the absence of Dlp1, despite the lack of sequence markers in Dlp2 that would predict a direct membrane association (Section 4.3.5). These indications of membrane interaction and fusion by Dlp2 are consistent with the *in vivo* subcellular fractionation studies showing a partial membrane association, and represent the first hint that the *H. pylori* DLPs might function in maintaining membrane integrity with a membrane fusion capacity. Future studies should investigate this hypothesis through the further development and utilisation of biochemical methods combined with fluorescence microscopy, such as super resolution microscopy, specifically 3D-Structured Illumination Microscopy. These should also be associated with *in vivo* methods for detecting interactions and visualising these proteins interacting with membranes in pure liposomes and in *H. pylori* cells. The implementation of a GFP-tagged DLP would aid in investigating the *in vivo* location of DLP and would potentially help with co-localisation studies of both Dlp1 and Dlp2 in *H. pylori*. Further studies utilising amino acid substitutions should also be implemented to investigate the precise mechanisms of lipid binding and aggregation. Furthermore, co-localisation studies using a fluorescently tagged DLP with liposomes would confirm that the aggregation and deformations seen are due to the binding of DLP to the lipid and not some other external factor such as disruption to surface tension due to crowding effects with the protein.

It is well established that both eukaryotic and bacterial dynamin self-interact or polymerise to carry out GTP hydrolysis and membrane curvature events (3, 4, 32). One question not addressed in this thesis was whether DLPs in *H. pylori* have intrinsic or stimulated GTPase activity. Previously, LeoA showed a lack of GTP binding and no GTPase activity *in vitro*, suggesting that additional factors may be required to observe these functions (5). Similarly Dlp2 might not show GTP hydrolysis capabilities but this could be potentially remedied with the presence of other factors such as lipids and/or its neighbouring protein Dlp1. Notably,

preliminary gel filtration data demonstrated the presence of a possible interaction between Dlp1a and Dlp2 (*Section 4.3.6*). A search for potential interactions between *H. pylori* DLPs was attempted using the established BacTH system (*Section 4.3.5*). In this system, interaction between the *H. pylori* DLPs was not detected, indicating that either the proteins do not interact as predicted or the system as applied is not suitable and/or sensitive enough to detect the interactions. However, with the overwhelming literature to suggest that dynamins do in fact interact to carry out GTP hydrolysis and to induce membrane curvature, other avenues were explored in a preliminary capacity. Preliminary gel-filtration chromatography as a means to detect complex formation has been tested with promising initial results. Moreover, a plasmid for co-expression of Dlp1 and Dlp2 has been prepared which allows the use of affinity purification to capture a possible complex. This expression system may also be a means to potentially increase the solubility of Dlp1 with its potential interaction partner, Dlp2 present.

This work has been the first to confirm the *in vivo* expression of DLPs in *H. pylori*. Furthermore, biochemical features such as lipid binding and protein-protein interactions have been initiated with future work directed towards investigating the possible functional complex for not only GTP hydrolysis but also membrane remodelling events that DLPs may perform in bacteria.

7.1.3 AIM III: TO EXPLORE THE FUNCTION OF DYNAMIN-LIKE PROTEINS IN *H. PYLORI*

Despite a number of investigations, the function of bacterial DLPs remains poorly understood, with minimal research examining the *in vivo* function of identified dynamins. Of those investigated, the functions display a remarkable level of diversity (5, 93, 97, 98, 100, 102). With over 90% of all *H. pylori* genomes screened in this study showing evidence of DLPs, establishing a putative function for them in this organism may shed light on the potential primitive function(s) of DLPs in bacteria. From this work, there are a number of factors that support the hypothesis that DLPs play a role in maintaining membrane integrity, with increased haemolysis (*Section 5.3.5*) and cell permeability (*Section 5.3.6*) and a reduction in acid tolerance independent of the urease activity noted in B128.7.13 Δ dlp

(Section 5.3.7 and 5.3.8). Using SWATH-MS, one of the most comprehensive quantitative investigations into the *H. pylori* proteome was achieved in this work and highlighted a number of potential pathways affected by the absence of DLPs under acid stress (Section 6.3). Specifically, the higher abundance of cytoplasmic pH homeostasis maintaining enzymes, such as AmiE and AmiF, suggests that without dynamin, *H. pylori* is compensating for this loss by attempting to neutralise the cytoplasmic influx of hydrogen ions. Interestingly, there were multiple components of the flagella machinery found in higher abundance, particularly the filament (FlaB) and the hook (FlgE) proteins under the same acidic conditions, suggesting that, due to an increase in the cytoplasmic pH resulting from a compromised membrane (Section 5.3.6), *H. pylori* may be attempting to increase its capacity to become motile. This is a known chemotactic motile response that *H. pylori* uses in response to acid (122). Furthermore, in one of the most comprehensive quantitative global studies of the proteome changes in *H. pylori* to date, there were several protein identifications that would indicate that without DLPs, *H. pylori* when under acid stress, is trying to ‘escape’ to a more pH neutral environment. In order to confirm this hypothesis based on the results presented here, a closer look at the motility in the B128.7.13 Δ *dlp* strain is needed. The increase in abundance of several components of the motility machinery in response to acid and in the absence of DLPs, would suggest that this strain may have an associated motility phenotype. Motility assays under a range of acid conditions and in the presence of a gradient of urea would help to not only follow up the SWATH-MS data, but to also determine the dependency on the urease pathway.

Along with being able to quantify relative protein abundance efficiently at a proteome-wide scale, the primary advantage of SWATH-MS is the endless possibilities in terms of conditions that can be tested and the timeframe in which they can be tested. By applying the SWATH-MS protocol to a global investigation of the proteomic changes in response to a certain stress, there are a plethora of possibilities for future studies stemming from these datasets (258). Therefore, with the notion that the DLPs are playing a fundamental role in membrane integrity in *H. pylori*, the proteome response in the presence and absence of DLP’s under a range of membrane targeting stresses is expected to lead to important insights. *H. pylori* frequently encounters acid stress (pH 2 – 4), oxidative stress, osmotic stress and

antibiotic exposure in the human stomach and survival is dependent on a membrane response (122, 304-306). The availability of a dynamin deletion strain, with its apparent membrane defect, presents an excellent model to study fundamental survival strategies in *H. pylori*, as some of these events will be amplified and more readily observed. Thus, exposing B128.7.13 WT and B128.7.13 Δ *dlp* strains to each of these stresses and assessing the proteome changes, similar to the acid exposure in this study, will help elucidate whether DLPs respond specifically to acid stress or in a more global membrane remodelling stress response.

One question that is not addressed in this thesis is whether the absence of dynamin from *H. pylori* affects vesicle production. The suggested role of LeoA in vesicle production in ETEC, and the knowledge that *H. pylori* is a vesicle forming pathogen, it could be reasonably hypothesised that the LeoCBA homologs may be playing a similar role in *H. pylori* (5, 7, 111, 173). This could be addressed qualitatively through use of higher resolution microscopy such as electron microscopy (EM) or a more quantitative approach using Western blot detection of vesicle markers like VacA (111, 173). The observation of increased haemolysis seen in the absence of DLPs in *H. pylori* (Section 5.3.5) would also be consistent with an increase in membrane vesicle formation due to a more fragile membrane, which could lead to an increase in the levels of haemolysin in the surrounding environment.

The functional analysis undertaken in this thesis was carried out in the gerbil-adapted *H. pylori* strain B128.7.13, due to the transformation required for deletion only being successful in this strain despite extensive attempts in both 26695 and SS1. However, with the promising results in this strain, establishing a putative function for DLPs in *H. pylori* membrane integrity, future work in more clinically relevant strains would help determine if this function is maintained across other strains of *dlp*-containing *H. pylori* strains. The utilisation of *H. pylori* strains such as SS1, which can colonise mouse models, will allow further *in vivo* study into the effect that dynamin may have on the ability of *H. pylori* to colonise and cause infection in a mouse host. A number of essential factors in *H. pylori* pathogenesis have been shown to be affected by the compromised membrane in B128.7.13 Δ *dlp* including; an inability of the strain to overcome acid exposure (Figure 5.10) and lower abundance of the important adhesin BabA (Section 6.4.4). Therefore, it would be speculated that the ability of

B128.7.13 Δ *dlp* to colonise a mouse would be reduced due to the reduction in membrane integrity (Figure 5.9). Along with complete knockout studies in the other strains, individual *dlp* gene knockout studies (i.e. deletion of only *dlp1*, *dlp2* or only the transmembrane domain) will help establish the role of each gene in the overall operon function in the bacteria. This will elucidate whether one dynamin is dependent on the other, a study which is yet to be done for “Leo-family” dynamin superfamily proteins.

Overall, this work has established a putative role of DLPs in membrane integrity in *H. pylori* under acid stress by utilising gene knock out studies as well as a novel global proteome approach. Interestingly, this function is reminiscent of DynA from *B. subtilis* despite having greater sequence homology to the LeoCBA proteins from ETEC. Future work will help confirm this putative function as well as explore the details of the network/pathways in which dynamins operate within *H. pylori*, which is completely unexplored territory in terms of bacterial DLPs.

7.2 WHAT DOES THIS MEAN FOR FURTHERING OUR UNDERSTANDING OF *H. PYLORI*?

H. pylori is known to be a dynamic and hardy pathogen that can colonise one of the harshest environments in the human body, the stomach. Still to this day, *H. pylori* is an understudied and underappreciated bacterial pathogen that requires much more attention to develop our understanding of how it persists in over half the world’s population. This thesis addresses the previously uncharacterised presence of membrane remodelling dynamin-like proteins in the *H. pylori* genome. Despite its dynamic nature, in terms of shuffling around its genome in response to its environment (229, 307), at least one DLP has been retained across 249 *H. pylori* strains screened in this study. This conservation, the highest noted level across a bacterial genus thus far, emphasises that *H. pylori* is using these genes for a purpose it is fundamentally required to survive and thrive in its native environment.

The importance of the membrane in putting up a defence against the harsh environment in which *H. pylori* thrives is under appreciated, specifically in how its integrity is maintained and repaired when damaged. Bacterial dynamins that have a previously characterised function in membrane surveillance are a worthy candidate for this role in *H. pylori* (100). This work has highlighted that this may well be the function that DLPs have in *H. pylori*, with clear reductions in membrane integrity and acid tolerance in their absence. The inability to tolerate short and longer-term exposures to low pH in the absence of dynamins indicates that these potential membrane remodelling proteins are essential for the successful colonisation of its host niche.

There is a great need for future development of *H. pylori* treatments, a bacterium naturally resistant to many antibiotics with this resistance only increasing around the world (304). There has been a focus on looking for novel vaccine candidates against *H. pylori* recently, but due to the intracellular localisation of DLPs, it is unlikely they will form a useful potential target, which ideally needs to be surface exposed. However, targeting the maintenance of the integrity of the membrane could be a novel approach in *H. pylori* treatment. With there being previously established pharmacological targeting of human dynamin, there may already be a bank of potential therapeutics to be tested against *H. pylori* dynamins (308-310). However, due to the low level of conservation between eukaryotic and bacterial dynamin, more research is needed to explore whether in combination with more conventional drugs, the pharmacological targeting of DLPs in *H. pylori* is a feasible target for the development of novel therapeutics for *H. pylori* treatment.

7.3 CONCLUDING REMARKS

The dynamin superfamily has been well established in eukaryotic cells as a diverse group of proteins that carry out a range of membrane remodelling events essential to many biological processes. Bacterial dynamin-like proteins, now recognised as members of this superfamily and despite being identified in many bacterial species have had minimal research conducted into understanding why they have been maintained throughout evolution. The research to date however begins to highlight that bacterial dynamins are also carrying out a diverse range

of functions in the cell, from ensuring membrane integrity under stress, to vesicle production, to maintaining cell division machinery stability.

This thesis has established that there is a high level of conservation of dynamin-like proteins in *H. pylori* and is the first study to reveal *in vivo* production of these proteins. The *H. pylori* dynamins have been shown to exhibit dynamin-like features such as lipid membrane aggregation. Finally, a putative function in maintaining membrane integrity under niche relevant acidic conditions has been established. This was further explored using a quantitative mass spectrometry approach that, to date has been one of the most comprehensive studies of any *H. pylori* proteome providing direct evidence for the production of 75% of the predicted *H. pylori* proteome. Overall, this thesis provides the first evidence that this bacterium uses dynamin-like proteins to maintain its membrane integrity in the face of acid stress.

H. pylori is highly understudied for a bacterium that is known to cause significant disease in the world including gastric cancer. Developing our understanding of how *H. pylori* not only causes but sustains colonisation in over half the world's population will not only help in improving the treatment, prevention and possible eradication of this devastating bacterium, but it will also contribute to deepening our knowledge of membrane-dependent bacterial survival strategies utilised for protection against host niche environments.

CHAPTER VIII

ADDITIONAL DATA (APPENDIX)

Supplementary Data is provided with this thesis. There are three chapters that have supplementary data (S1-7) to support their results and these are listed below according to their corresponding thesis chapter:

CHAPTER III

- S1 - *H. pylori* (1997) Dlp1a, Dlp1b and Dlp2 as well as *H. pylori* (2013) Dlp1 Sequences
- S2 - Conservation screen for Dynamin-Like proteins in *H. pylori*
- S3 - Transmembrane predictions of selected *H. pylori* Dlp
- S4 - Completely Labelled *H. pylori* dynamin (back-to-back) phylogenetic tree

CHAPTER V

- S5 - *H. pylori* B128.7.13 predicted proteome annotated by PROKKA
- S6 - Membrane Integrity Individual Filter Images from the DeltaVision Elite

CHAPTER VI

- S7 - Protein Pilot Output for SWATH Ion Library and Markerview quantification and statistical analysis (In the form of a tabbed Excel spreadsheet)
- Tab Headings (Left to right)
 - S7a B128.7.13 SWATH Ion Library
 - S7b WT pH 7 verses pH 5 all quantified proteins
 - S7c WT pH 7 verses pH 5 significantly (< 0.05) quantified proteins
 - S7d Mut (Δdlp) pH7 verses pH 5 all quantified proteins
 - S7e Mut (Δdlp) pH 7 verses pH 5 significantly (< 0.05) quantified proteins
 - S7f WT verses Mut (Δdlp) at pH 7 all quantified proteins
 - S7g WT verses Mut (Δdlp) at pH 7 significantly (< 0.05) quantified proteins
 - S7h WT verses Mut (Δdlp) at pH 5 all quantified proteins
 - S7i WT verses Mut (Δdlp) at pH 5 significantly (< 0.05) quantified proteins
 - S7j WT verses Mut (Δdlp) pH 5 Quantified Flagella Proteins

REFERENCES

1. Ferguson SM & De Camilli P (2012) Dynamin, a membrane-remodelling GTPase. *Nat Rev Mol Cell Biol* 13(2):75-88.
2. Praefcke GJ & McMahon HT (2004) The dynamin superfamily: universal membrane tubulation and fission molecules? *Nat Rev Mol Cell Biol* 5(2):133-147.
3. Burmann F, Ebert N, van Baarle S, & Bramkamp M (2011) A bacterial dynamin-like protein mediating nucleotide-independent membrane fusion. *Mol Microbiol* 79(5):1294-1304.
4. Low HH & Lowe J (2006) A bacterial dynamin-like protein. *Nature* 444(7120):766-769.
5. Michie KA, Boysen A, Low HH, Moller-Jensen J, & Lowe J (2014) LeoA, B and C from enterotoxigenic *Escherichia coli* (ETEC) are bacterial dynamins. *PLoS One* 9(9):e107211.
6. Bramkamp M (2012) Structure and function of bacterial dynamin-like proteins. *Biol Chem* 393(11):1203-1214.
7. Brown EA & Hardwidge PR (2007) Biochemical characterization of the enterotoxigenic *Escherichia coli* LeoA protein. *Microbiology* 153(Pt 11):3776-3784.
8. Oleastro M & Menard A (2013) The Role of *Helicobacter pylori* Outer Membrane Proteins in Adherence and Pathogenesis. *Biology (Basel)* 2(3):1110-1134.
9. Roesler BM, Rabelo-Goncalves EM, & Zeitune JM (2014) Virulence Factors of *Helicobacter pylori*: A Review. *Clin Med Insights Gastroenterol* 7:9-17.
10. Salama NR, Hartung ML, & Muller A (2013) Life in the human stomach: persistence strategies of the bacterial pathogen *Helicobacter pylori*. *Nat Rev Microbiol* 11(6):385-399.
11. Doherty GJ & McMahon HT (2009) Mechanisms of endocytosis. *Annu Rev Biochem* 78:857-902.
12. Wu LG, Hamid E, Shin W, & Chiang HC (2014) Exocytosis and endocytosis: modes, functions, and coupling mechanisms. *Annu Rev Physiol* 76:301-331.
13. McNew JA, Sondermann H, Lee T, Stern M, & Brandizzi F (2013) GTP-dependent membrane fusion. *Annu Rev Cell Dev Biol* 29:529-550.

14. Morlot S & Roux A (2013) Mechanics of dynamin-mediated membrane fission. *Annu Rev Biophys* 42:629-649.
15. Di Paolo G & De Camilli P (2006) Phosphoinositides in cell regulation and membrane dynamics. *Nature* 443(7112):651-657.
16. Kozlov MM, McMahon HT, & Chernomordik LV (2010) Protein-driven membrane stresses in fusion and fission. *Trends Biochem Sci* 35(12):699-706.
17. McMahon HT & Boucrot E (2011) Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nat Rev Mol Cell Biol* 12(8):517-533.
18. Wileman T, Harding C, & Stahl P (1985) Receptor-mediated endocytosis. *Biochem J* 232(1):1-14.
19. Xiao G & Gan LS (2013) Receptor-mediated endocytosis and brain delivery of therapeutic biologics. *Int J Cell Biol* 2013:703545.
20. Hinshaw JE & Schmid SL (1995) Dynamin self-assembles into rings suggesting a mechanism for coated vesicle budding. *Nature* 374(6518):190-192.
21. Suzuki DT, Grigliatti T, & Williamson R (1971) Temperature-sensitive mutations in *Drosophila melanogaster* VII. A mutation (para-ts) causing reversible adult paralysis. *Proc Natl Acad Sci U S A* 68(5):890-893.
22. Grigliatti TA, Hall L, Rosenbluth R, & Suzuki DT (1973) Temperature-sensitive mutations in *Drosophila melanogaster* XIV. A selection of immobile adults. *Mol Gen Genet* 120(2):107-114.
23. Koenig JH & Ikeda K (1989) Disappearance and reformation of synaptic vesicle membrane upon transmitter release observed under reversible blockage of membrane retrieval. *J Neurosci* 9(11):3844-3860.
24. van der Bliek AM & Meyerowitz EM (1991) Dynamin-like protein encoded by the *Drosophila* shibire gene associated with vesicular traffic. *Nature* 351(6325):411-414.
25. Chappie JS, Acharya S, Leonard M, Schmid SL, & Dyda F (2010) G domain dimerization controls dynamin's assembly-stimulated GTPase activity. *Nature* 465(7297):435-440.
26. Chappie JS, Acharya S, Liu YW, Leonard M, Pucadyil TJ, & Schmid SL (2009) An intramolecular signaling element that modulates dynamin function *in vitro* and *in vivo*. *Mol Biol Cell* 20(15):3561-3571.

27. Timm D, Salim K, Gout I, Guruprasad L, Waterfield M, & Blundell T (1994) Crystal structure of the pleckstrin homology domain from Dynamin. *Nat Struct Biol* 1(11):782-788.
28. Ferguson KM, Lemmon MA, Schlessinger J, & Sigler PB (1994) Crystal structure at 2.2 Å resolution of the pleckstrin homology domain from human Dynamin. *Cell* 79(2):199-209.
29. Ford MG, Jenni S, & Nunnari J (2011) The crystal structure of Dynamin. *Nature* 477(7366):561-566.
30. Faelber K, Posor Y, Gao S, Held M, Roske Y, Schulze D, Haucke V, Noe F, & Daumke O (2011) Crystal structure of nucleotide-free Dynamin. *Nature* 477(7366):556-560.
31. Anand R, Eschenburg S, & Reubold TF (2016) Crystal structure of the GTPase domain and the bundle signalling element of dynamin in the GDP state. *Biochem Biophys Res Commun* 469(1):76-80.
32. Reubold TF, Faelber K, Plattner N, Posor Y, Ketel K, Curth U, Schlegel J, Anand R, Manstein DJ, Noe F, Haucke V, Daumke O, & Eschenburg S (2015) Crystal structure of the Dynamin tetramer. *Nature* 525(7569):404-408.
33. Antonny B, Burd C, De Camilli P, Chen E, Daumke O, Faelber K, Ford M, Frolov VA, Frost A, Hinshaw JE, Kirchhausen T, Kozlov MM, Lenz M, Low HH, McMahon H, Merrifield C, Pollard TD, Robinson PJ, Roux A, & Schmid S (2016) Membrane fission by Dynamin: what we know and what we need to know. *EMBO J* 35(21):2270-2284.
34. Vallis Y, Wigge P, Marks B, Evans PR, & McMahon HT (1999) Importance of the pleckstrin homology domain of Dynamin in clathrin-mediated endocytosis. *Curr Biol* 9(5):257-260.
35. Morita M, Hamao K, Izumi S, Okumura E, Tanaka K, Kishimoto T, & Hosoya H (2010) Proline-rich domain in Dynamin-2 has a low microtubule-binding activity: how is this activity controlled during mitosis in HeLa cells? *J Biochem* 148(5):533-538.

36. Hamao K, Morita M, & Hosoya H (2009) New function of the proline rich domain in Dynamin-2 to negatively regulate its interaction with microtubules in mammalian cells. *Exp Cell Res* 315(7):1336-1345.
37. Cao S, Yao J, & Shah V (2003) The proline-rich domain of Dynamin-2 is responsible for dynamin-dependent *in vitro* potentiation of endothelial nitric-oxide synthase activity via selective effects on reductase domain function. *J Biol Chem* 278(8):5894-5901.
38. Szaszak M, Gaborik Z, Turu G, McPherson PS, Clark AJ, Catt KJ, & Hunyady L (2002) Role of the proline-rich domain of Dynamin-2 and its interactions with Src homology 3 domains during endocytosis of the AT1 angiotensin receptor. *J Biol Chem* 277(24):21650-21656.
39. Okamoto PM, Herskovits JS, & Vallee RB (1997) Role of the basic, proline-rich region of Dynamin in Src homology 3 domain binding and endocytosis. *J Biol Chem* 272(17):11629-11635.
40. Warnock DE, Hinshaw JE, & Schmid SL (1996) Dynamin self-assembly stimulates its GTPase activity. *J Biol Chem* 271(37):22310-22314.
41. Warnock DE & Schmid SL (1996) Dynamin GTPase, a force-generating molecular switch. *Bioessays* 18(11):885-893.
42. Daumke O & Praefcke GJ (2016) Invited review: Mechanisms of GTP hydrolysis and conformational transitions in the Dynamin superfamily. *Biopolymers* 105(8):580-593.
43. Li G & Zhang XC (2004) GTP hydrolysis mechanism of Ras-like GTPases. *J Mol Biol* 340(5):921-932.
44. Bashkistrov PV, Akimov SA, Evseev AI, Schmid SL, Zimmerberg J, & Frolov VA (2008) GTPase cycle of Dynamin is coupled to membrane squeeze and release, leading to spontaneous fission. *Cell* 135(7):1276-1286.
45. Pucadyil TJ & Schmid SL (2008) Real-time visualization of Dynamin-catalyzed membrane fission and vesicle release. *Cell* 135(7):1263-1275.
46. Chappie JS, Mears JA, Fang S, Leonard M, Schmid SL, Milligan RA, Hinshaw JE, & Dyda F (2011) A pseudoatomic model of the Dynamin polymer identifies a hydrolysis-dependent powerstroke. *Cell* 147(1):209-222.

47. Sontag JM, Fykse EM, Ushkaryov Y, Liu JP, Robinson PJ, & Sudhof TC (1994) Differential expression and regulation of multiple Dynamins. *J Biol Chem* 269(6):4547-4554.
48. Peng J, Wu L, & Zhang CL (2011) Role of functional domains of Dynamin-I in synaptic vesicle endocytosis. *Sheng Li Ke Xue Jin Zhan* 42(2):121-124.
49. Gray NW, Fourgeaud L, Huang B, Chen J, Cao H, Oswald BJ, Hemar A, & McNiven MA (2003) Dynamin 3 is a component of the postsynapse, where it interacts with mGluR5 and Homer. *Curr Biol* 13(6):510-515.
50. Meckler X, Bertandeu E, McNiven MA, Allinquant B, & Hemar A (2006) The cytosolic domain of APP induces the relocalization of Dynamin 3 in hippocampal neurons. *Eur J Neurosci* 24(9):2439-2443.
51. Dong J, Misselwitz R, Welfle H, & Westermann P (2000) Expression and purification of Dynamin II domains and initial studies on structure and function. *Protein Expr Purif* 20(2):314-323.
52. Altschuler Y, Barbas SM, Terlecky LJ, Tang K, Hardy S, Mostov KE, & Schmid SL (1998) Redundant and distinct functions for Dynamin-1 and Dynamin-2 isoforms. *J Cell Biol* 143(7):1871-1881.
53. Zuchner S, Noureddine M, Kennerson M, Verhoeven K, Claeys K, De Jonghe P, Merory J, Oliveira SA, Speer MC, Stenger JE, Walizada G, Zhu D, Pericak-Vance MA, Nicholson G, Timmerman V, & Vance JM (2005) Mutations in the pleckstrin homology domain of Dynamin 2 cause dominant intermediate Charcot-Marie-Tooth disease. *Nat Genet* 37(3):289-294.
54. Claeys KG, Zuchner S, Kennerson M, Berciano J, Garcia A, Verhoeven K, Storey E, Merory JR, Bienfait HM, Lammens M, Nelis E, Baets J, De Vriendt E, Berneman ZN, De Veuster I, Vance JM, Nicholson G, Timmerman V, & De Jonghe P (2009) Phenotypic spectrum of Dynamin 2 mutations in Charcot-Marie-Tooth neuropathy. *Brain* 132(Pt 7):1741-1752.
55. Boumil RM, Letts VA, Roberts MC, Lenz C, Mahaffey CL, Zhang ZW, Moser T, & Frankel WN (2010) A missense mutation in a highly conserved alternate exon of Dynamin-1 causes epilepsy in fitful mice. *PLoS Genet* 6(8).

56. Li YY, Chen XN, Fan XX, Zhang YJ, Gu J, Fu XW, Wang ZH, Wang XF, & Xiao Z (2015) Upregulated Dynamin 1 in an acute seizure model and in epileptic patients. *Synapse* 69(2):67-77.
57. Miyagishima SY, Nishida K, Mori T, Matsuzaki M, Higashiyama T, Kuroiwa H, & Kuroiwa T (2003) A plant-specific Dynamin-related protein forms a ring at the chloroplast division site. *Plant Cell* 15(3):655-665.
58. Gao H, Kadirjan-Kalbach D, Froehlich JE, & Osteryoung KW (2003) ARC5, a cytosolic Dynamin-like protein from plants, is part of the chloroplast division machinery. *Proc Natl Acad Sci U S A* 100(7):4328-4333.
59. Ishida N, Nakamura Y, Tanabe K, Li SA, & Takei K (2011) Dynamin 2 associates with microtubules at mitosis and regulates cell cycle progression. *Cell Struct Funct* 36(2):145-154.
60. Staeheli P, Danielson P, Haller O, & Sutcliffe JG (1986) Transcriptional activation of the mouse Mx gene by type I interferon. *Mol Cell Biol* 6(12):4770-4774.
61. Staeheli P, Haller O, Boll W, Lindenmann J, & Weissmann C (1986) Mx protein: constitutive expression in 3T3 cells transformed with cloned Mx cDNA confers selective resistance to influenza virus. *Cell* 44(1):147-158.
62. Janzen C, Kochs G, & Haller O (2000) A monomeric GTPase-negative MxA mutant with antiviral activity. *J Virol* 74(17):8202-8206.
63. Kochs G & Haller O (1999) Interferon-induced human MxA GTPase blocks nuclear import of Thogoto virus nucleocapsids. *Proc Natl Acad Sci U S A* 96(5):2082-2086.
64. Kochs G & Haller O (1999) GTP-bound human MxA protein interacts with the nucleocapsids of Thogoto virus (Orthomyxoviridae). *J Biol Chem* 274(7):4370-4376.
65. Flohr F, Schneider-Schaulies S, Haller O, & Kochs G (1999) The central interactive region of human MxA GTPase is involved in GTPase activation and interaction with viral target structures. *FEBS Lett* 463(1-2):24-28.
66. Smirnova E, Shurland DL, Ryazantsev SN, & van der Blik AM (1998) A human Dynamin-related protein controls the distribution of mitochondria. *J Cell Biol* 143(2):351-358.

67. Labrousse AM, Zappaterra MD, Rube DA, & van der Blik AM (1999) C. elegans Dynamin-related protein DRP-1 controls severing of the mitochondrial outer membrane. *Mol Cell* 4(5):815-826.
68. Kamimoto T, Nagai Y, Onogi H, Muro Y, Wakabayashi T, & Hagiwara M (1998) Dymple, a novel Dynamin-like high molecular weight GTPase lacking a proline-rich carboxyl-terminal domain in mammalian cells. *J Biol Chem* 273(2):1044-1051.
69. Shin HW, Shinotsuka C, Torii S, Murakami K, & Nakayama K (1997) Identification and subcellular localization of a novel mammalian Dynamin-related protein homologous to yeast Vps1p and Dnm1p. *J Biochem* 122(3):525-530.
70. Satoh M, Hamamoto T, Seo N, Kagawa Y, & Endo H (2003) Differential sublocalization of the Dynamin-related protein OPA1 isoforms in mitochondria. *Biochem Biophys Res Commun* 300(2):482-493.
71. Olichon A, Baricault L, Gas N, Guillou E, Valette A, Belenguer P, & Lenaers G (2003) Loss of OPA1 perturbs the mitochondrial inner membrane structure and integrity, leading to cytochrome c release and apoptosis. *J Biol Chem* 278(10):7743-7746.
72. Olichon A, Emorine LJ, Descoins E, Pelloquin L, Brichese L, Gas N, Guillou E, Delettre C, Valette A, Hamel CP, Ducommun B, Lenaers G, & Belenguer P (2002) The human Dynamin-related protein OPA1 is anchored to the mitochondrial inner membrane facing the inter-membrane space. *FEBS Lett* 523(1-3):171-176.
73. Hermann GJ, Thatcher JW, Mills JP, Hales KG, Fuller MT, Nunnari J, & Shaw JM (1998) Mitochondrial fusion in yeast requires the transmembrane GTPase Fzo1p. *J Cell Biol* 143(2):359-373.
74. Santel A & Fuller MT (2001) Control of mitochondrial morphology by a human mitofusin. *J Cell Sci* 114(Pt 5):867-874.
75. Rapaport D, Brunner M, Neupert W, & Westermann B (1998) Fzo1p is a mitochondrial outer membrane protein essential for the biogenesis of functional mitochondria in *Saccharomyces cerevisiae*. *J Biol Chem* 273(32):20150-20155.
76. Prakash B, Praefcke GJ, Renault L, Wittinghofer A, & Herrmann C (2000) Structure of human guanylate-binding protein 1 representing a unique class of GTP-binding proteins. *Nature* 403(6769):567-571.

77. Zhu PP, Patterson A, Lavoie B, Stadler J, Shoeb M, Patel R, & Blackstone C (2003) Cellular localization, oligomerization, and membrane association of the hereditary spastic paraplegia 3A (SPG3A) protein Atlastin. *J Biol Chem* 278(49):49063-49071.
78. Uthaiiah RC, Praefcke GJ, Howard JC, & Herrmann C (2003) IIGP1, an interferon-gamma-inducible 47-kDa GTPase of the mouse, showing cooperative enzymatic activity and GTP-dependent multimerization. *J Biol Chem* 278(31):29336-29343.
79. Anderson SL, Carton JM, Lou J, Xing L, & Rubin BY (1999) Interferon-induced guanylate binding protein-1 (GBP-1) mediates an antiviral effect against vesicular stomatitis virus and encephalomyocarditis virus. *Virology* 256(1):8-14.
80. Silhavy TJ, Kahne D, & Walker S (2010) The bacterial cell envelope. *Cold Spring Harb Perspect Biol* 2(5):a000414.
81. Vollmer W, Blanot D, & de Pedro MA (2008) Peptidoglycan structure and architecture. *FEMS Microbiol Rev* 32(2):149-167.
82. Heptinstall S, Archibald AR, & Baddiley J (1970) Teichoic acids and membrane function in bacteria. *Nature* 225(5232):519-521.
83. Brown S, Santa Maria JP, Jr., & Walker S (2013) Wall teichoic acids of Gram-positive bacteria. *Annu Rev Microbiol* 67:313-336.
84. Aderem A & Ulevitch RJ (2000) Toll-like receptors in the induction of the innate immune response. *Nature* 406(6797):782-787.
85. Bohuszewicz O, Liu J, & Low HH (2016) Membrane remodelling in bacteria. *J Struct Biol* 196(1):3-14.
86. Siegel SD, Liu J, & Ton-That H (2016) Biogenesis of the Gram-positive bacterial cell envelope. *Curr Opin Microbiol* 34:31-37.
87. Yen MR, Peabody CR, Partovi SM, Zhai Y, Tseng YH, & Saier MH (2002) Protein-translocating outer membrane porins of Gram-negative bacteria. *Biochim Biophys Acta* 1562(1-2):6-31.
88. Bos MP, Robert V, & Tommassen J (2007) Biogenesis of the Gram-negative bacterial outer membrane. *Annu Rev Microbiol* 61:191-214.
89. Tripathi P, Beaussart A, Andre G, Rolain T, Lebeer S, Vanderleyden J, Hols P, & Dufrene YF (2012) Towards a nanoscale view of lactic acid bacteria. *Micron* 43(12):1323-1330.

90. Rojas E, Theriot JA, & Huang KC (2014) Response of *Escherichia coli* growth rate to osmotic shock. *Proc Natl Acad Sci U S A* 111(21):7807-7812.
91. Low HH, Sachse C, Amos LA, & Lowe J (2009) Structure of a bacterial Dynamin-like protein lipid tube provides a mechanism for assembly and membrane curving. *Cell* 139(7):1342-1352.
92. Roux A, Uyhazi K, Frost A, & De Camilli P (2006) GTP-dependent twisting of Dynamin implicates constriction and tension in membrane fission. *Nature* 441(7092):528-531.
93. Ozaki S, Matsuda Y, Keyamura K, Kawakami H, Noguchi Y, Kasho K, Nagata K, Masuda T, Sakiyama Y, & Katayama T (2013) A replicase clamp-binding Dynamin-like protein promotes colocalization of nascent DNA strands and equipartitioning of chromosomes in *E. coli*. *Cell Rep* 4(5):985-995.
94. Hedglin M, Kumar R, & Benkovic SJ (2013) Replication clamps and clamp loaders. *Cold Spring Harb Perspect Biol* 5(4):a010165.
95. Naqvi AA, Anjum F, Khan FI, Islam A, Ahmad F, & Hassan MI (2016) Sequence Analysis of Hypothetical Proteins from *Helicobacter pylori* 26695 to Identify Potential Virulence Factors. *Genomics Inform* 14(3):125-135.
96. Alland D, Steyn AJ, Weisbrod T, Aldrich K, & Jacobs WR, Jr. (2000) Characterization of the *Mycobacterium tuberculosis* iniBAC promoter, a promoter that responds to cell wall biosynthesis inhibition. *J Bacteriol* 182(7):1802-1811.
97. Colangeli R, Helb D, Sridharan S, Sun J, Varma-Basil M, Hazbon MH, Harbacheuski R, Megjugorac NJ, Jacobs WR, Jr., Holzenburg A, Sacchettini JC, & Alland D (2005) The *Mycobacterium tuberculosis* iniA gene is essential for activity of an efflux pump that confers drug tolerance to both isoniazid and ethambutol. *Mol Microbiol* 55(6):1829-1840.
98. Schlimpert S, Wasserstrom S, Chandra G, Bibb MJ, Findlay KC, Flardh K, & Buttner MJ (2017) Two Dynamin-like proteins stabilize FtsZ rings during *Streptomyces* sporulation. *Proc Natl Acad Sci U S A*.

99. Fleckenstein JM, Lindler LE, Elsinghorst EA, & Dale JB (2000) Identification of a Gene within a Pathogenicity Island of Enterotoxigenic *Escherichia coli* H10407 Required for Maximal Secretion of the Heat-Labile Enterotoxin. *Infect Immun* 68(5):2766-2774.
100. Sawant P, Eissenberger K, Karier L, Mascher T, & Bramkamp M (2015) A Dynamin-like protein involved in bacterial cell membrane surveillance under environmental stress. *Environ Microbiol*.
101. Patrick JE & Kearns DB (2008) MinJ (YvjD) is a topological determinant of cell division in *Bacillus subtilis*. *Mol Microbiol* 70(5):1166-1179.
102. Dempwolff F, Wischhusen HM, Specht M, & Graumann PL (2012) The deletion of bacterial Dynamin and flotillin genes results in pleiotrophic effects on cell division, cell growth and in cell shape maintenance. *BMC Microbiol* 12:298.
103. Burmann F, Sawant P, & Bramkamp M (2012) Identification of interaction partners of the Dynamin-like protein DynA from *Bacillus subtilis*. *Commun Integr Biol* 5(4):362-369.
104. Mansilla MC, Cybulski LE, Albanesi D, & de Mendoza D (2004) Control of membrane lipid fluidity by molecular thermosensors. *J Bacteriol* 186(20):6681-6688.
105. Li J, Chikindas ML, Ludescher RD, & Montville TJ (2002) Temperature- and surfactant-induced membrane modifications that alter *Listeria monocytogenes* nisin sensitivity by different mechanisms. *Appl Environ Microbiol* 68(12):5904-5910.
106. Fleckenstein JM, Kopecko DJ, Warren RL, & Elsinghorst EA (1996) Molecular characterization of the tia invasion locus from enterotoxigenic *Escherichia coli*. *Infect Immun* 64(6):2256-2265.
107. Nilsson C, Sillen A, Eriksson L, Strand ML, Enroth H, Normark S, Falk P, & Engstrand L (2003) Correlation between cag pathogenicity island composition and *Helicobacter pylori* associated gastroduodenal disease. *Infect Immun* 71(11):6573-6581.
108. Horstman AL & Kuehn MJ (2000) Enterotoxigenic *Escherichia coli* secretes active heat-labile enterotoxin via outer membrane vesicles. *J Biol Chem* 275(17):12489-12496.

109. Schwechheimer C & Kuehn MJ (2015) Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. *Nat Rev Microbiol* 13(10):605-619.
110. Ellis TN & Kuehn MJ (2010) Virulence and immunomodulatory roles of bacterial outer membrane vesicles. *Microbiol Mol Biol Rev* 74(1):81-94.
111. Fiocca R, Necchi V, Sommi P, Ricci V, Telford J, Cover TL, & Solcia E (1999) Release of *Helicobacter pylori* vacuolating cytotoxin by both a specific secretion pathway and budding of outer membrane vesicles. Uptake of released toxin and vesicles by gastric epithelium. *J Pathol* 188(2):220-226.
112. Devoe IW & Gilchrist JE (1973) Release of endotoxin in the form of cell wall blebs during in vitro growth of *Neisseria meningitidis*. *J Exp Med* 138(5):1156-1167.
113. Chatterjee SN & Das J (1967) Electron microscopic observations on the excretion of cell-wall material by *Vibrio cholerae*. *J Gen Microbiol* 49(1):1-11.
114. Yoon H, Ansong C, Adkins JN, & Heffron F (2011) Discovery of *Salmonella* virulence factors translocated via outer membrane vesicles to murine macrophages. *Infect Immun* 79(6):2182-2192.
115. Elmi A, Watson E, Sandu P, Gundogdu O, Mills DC, Inglis NF, Manson E, Imrie L, Bajaj-Elliott M, Wren BW, Smith DG, & Dorrell N (2012) *Campylobacter jejuni* outer membrane vesicles play an important role in bacterial interactions with human intestinal epithelial cells. *Infect Immun* 80(12):4089-4098.
116. Jan AT (2017) Outer Membrane Vesicles (OMVs) of Gram-negative Bacteria: A Perspective Update. *Front Microbiol* 8:1053.
117. Miller VL, Beer KB, Heusipp G, Young BM, & Wachtel MR (2001) Identification of regions of Ail required for the invasion and serum resistance phenotypes. *Mol Microbiol* 41(5):1053-1062.
118. de Kort G, Bolton A, Martin G, Stephen J, & van de Klundert JA (1994) Invasion of rabbit ileal tissue by *Enterobacter cloacae* varies with the concentration of OmpX in the outer membrane. *Infect Immun* 62(11):4722-4726.
119. Prasadarao NV, Wass CA, Weiser JN, Stins MF, Huang SH, & Kim KS (1996) Outer membrane protein A of *Escherichia coli* contributes to invasion of brain microvascular endothelial cells. *Infect Immun* 64(1):146-153.

120. Aldridge PD, Karlinsey JE, Aldridge C, Birchall C, Thompson D, Yagasaki J, & Hughes KT (2006) The flagellar-specific transcription factor, sigma28, is the Type III secretion chaperone for the flagellar-specific anti-sigma28 factor FlgM. *Genes Dev* 20(16):2315-2326.
121. Brown LM (2000) *Helicobacter pylori*: epidemiology and routes of transmission. *Epidemiol Rev* 22(2):283-297.
122. Keilberg D & Ottemann KM (2016) How *Helicobacter pylori* senses, targets, and interacts with the gastric epithelium. *Environ Microbiol*.
123. Rolig AS, Shanks J, Carter JE, & Ottemann KM (2012) *Helicobacter pylori* requires TlpD-driven chemotaxis to proliferate in the antrum. *Infect Immun* 80(10):3713-3720.
124. Warren JR & Marshall B (1983) Unidentified curved bacilli on gastric epithelium in active chronic gastritis. *Lancet* 1(8336):1273-1275.
125. Marshall BJ (2001) One Hundred Years of Discovery and Rediscovery of *Helicobacter pylori* and Its Association with Peptic Ulcer Disease. *Helicobacter pylori: Physiology and Genetics*, eds Mobley HLT, Mendz GL, & Hazell SL (Washington (DC)).
126. Jyotheeswaran S, Shah AN, Jin HO, Potter GD, Ona FV, & Chey WY (1998) Prevalence of *Helicobacter pylori* in peptic ulcer patients in greater Rochester, NY: is empirical triple therapy justified? *Am J Gastroenterol* 93(4):574-578.
127. Vogiati P, Cassone M, Luzzi I, Lucchetti C, Otvos L, Jr., & Giordano A (2007) *Helicobacter pylori* as a class I carcinogen: physiopathology and management strategies. *J Cell Biochem* 102(2):264-273.
128. Atherton JC & Blaser MJ (2009) Coadaptation of *Helicobacter pylori* and humans: ancient history, modern implications. *J Clin Invest* 119(9):2475-2487.
129. Kandulski A, Selgrad M, & Malfertheiner P (2008) *Helicobacter pylori* infection: a clinical overview. *Dig Liver Dis* 40(8):619-626.
130. Kusters JG, van Vliet AH, & Kuipers EJ (2006) Pathogenesis of *Helicobacter pylori* infection. *Clin Microbiol Rev* 19(3):449-490.

131. Loh JT, Gaddy JA, Algood HM, Gaudieri S, Mallal S, & Cover TL (2015) *Helicobacter pylori* adaptation *in vivo* in response to a high-salt diet. *Infect Immun* 83(12):4871-4883.
132. Gaddy JA, Radin JN, Loh JT, Zhang F, Washington MK, Peek RM, Jr., Algood HM, & Cover TL (2013) High dietary salt intake exacerbates *Helicobacter pylori*-induced gastric carcinogenesis. *Infect Immun* 81(6):2258-2267.
133. Cover TL & Peek RM, Jr. (2013) Diet, microbial virulence, and *Helicobacter pylori*-induced gastric cancer. *Gut Microbes* 4(6):482-493.
134. Merrell DS, Goodrich ML, Otto G, Tompkins LS, & Falkow S (2003) pH-regulated gene expression of the gastric pathogen *Helicobacter pylori*. *Infect Immun* 71(6):3529-3539.
135. Pflock M, Finsterer N, Joseph B, Mollenkopf H, Meyer TF, & Beier D (2006) Characterization of the ArsRS regulon of *Helicobacter pylori*, involved in acid adaptation. *J Bacteriol* 188(10):3449-3462.
136. Bauerfeind P, Garner R, Dunn BE, & Mobley HL (1997) Synthesis and activity of *Helicobacter pylori* urease and catalase at low pH. *Gut* 40(1):25-30.
137. Scott DR, Marcus EA, Weeks DL, Lee A, Melchers K, & Sachs G (2000) Expression of the *Helicobacter pylori* ureI gene is required for acidic pH activation of cytoplasmic urease. *Infect Immun* 68(2):470-477.
138. Vollan HS, Tannaes T, Caugant DA, Vriend G, & Bukholm G (2017) Outer membrane phospholipase A's roles in *Helicobacter pylori* acid adaptation. *Gut Pathog* 9:36.
139. Athmann C, Zeng N, Kang T, Marcus EA, Scott DR, Rektorschek M, Buhmann A, Melchers K, & Sachs G (2000) Local pH elevation mediated by the intrabacterial urease of *Helicobacter pylori* cocultured with gastric cells. *J Clin Invest* 106(3):339-347.
140. Montecucco C & Rappuoli R (2001) Living dangerously: how *Helicobacter pylori* survives in the human stomach. *Nat Rev Mol Cell Biol* 2(6):457-466.
141. Andermann TM, Chen YT, & Ottemann KM (2002) Two predicted chemoreceptors of *Helicobacter pylori* promote stomach infection. *Infect Immun* 70(10):5877-5881.

142. Foynes S, Dorrell N, Ward SJ, Stabler RA, McColm AA, Rycroft AN, & Wren BW (2000) *Helicobacter pylori* possesses two CheY response regulators and a histidine kinase sensor, CheA, which are essential for chemotaxis and colonization of the gastric mucosa. *Infect Immun* 68(4):2016-2023.
143. Lertsethtakarn P, Ottemann KM, & Hendrixson DR (2011) Motility and chemotaxis in *Campylobacter* and *Helicobacter*. *Annu Rev Microbiol* 65:389-410.
144. Mizote T, Yoshiyama H, & Nakazawa T (1997) Urease-independent chemotactic responses of *Helicobacter pylori* to urea, urease inhibitors, and sodium bicarbonate. *Infect Immun* 65(4):1519-1521.
145. Sanders L, Andermann TM, & Ottemann KM (2013) A supplemented soft agar chemotaxis assay demonstrates the *Helicobacter pylori* chemotactic response to zinc and nickel. *Microbiology* 159(Pt 1):46-57.
146. Gao B, Lara-Tejero M, Lefebvre M, Goodman AL, & Galan JE (2014) Novel components of the flagellar system in epsilonproteobacteria. *MBio* 5(3):e01349-01314.
147. Spohn G & Scarlato V (2001) Motility, Chemotaxis, and Flagella. *Helicobacter pylori: Physiology and Genetics*, eds Mobley HLT, Mendz GL, & Hazell SL Washington (DC)).
148. Ottemann KM & Lowenthal AC (2002) *Helicobacter pylori* uses motility for initial colonization and to attain robust infection. *Infect Immun* 70(4):1984-1990.
149. Sachs G, Weeks DL, Wen Y, Marcus EA, Scott DR, & Melchers K (2005) Acid acclimation by *Helicobacter pylori*. *Physiology (Bethesda)* 20:429-438.
150. Baker MD, Wolanin PM, & Stock JB (2006) Signal transduction in bacterial chemotaxis. *Bioessays* 28(1):9-22.
151. Sycuro LK, Wyckoff TJ, Biboy J, Born P, Pincus Z, Vollmer W, & Salama NR (2012) Multiple peptidoglycan modification networks modulate *Helicobacter pylori*'s cell shape, motility, and colonization potential. *PLoS Pathog* 8(3):e1002603.
152. Sycuro LK, Pincus Z, Gutierrez KD, Biboy J, Stern CA, Vollmer W, & Salama NR (2010) Peptidoglycan crosslinking relaxation promotes *Helicobacter pylori*'s helical shape and stomach colonization. *Cell* 141(5):822-833.

153. Bonis M, Ecobichon C, Guadagnini S, Prevost MC, & Boneca IG (2010) A M23B family metallopeptidase of *Helicobacter pylori* required for cell shape, pole formation and virulence. *Mol Microbiol* 78(4):809-819.
154. Backert S, Clyne M, & Tegtmeyer N (2011) Molecular mechanisms of gastric epithelial cell adhesion and injection of CagA by *Helicobacter pylori*. *Cell Commun Signal* 9:28.
155. Yamaoka Y (2008) Roles of *Helicobacter pylori* BabA in gastroduodenal pathogenesis. *World J Gastroenterol* 14(27):4265-4272.
156. Aspholm-Hurtig M, Dailide G, Lahmann M, Kalia A, Ilver D, Roche N, Vikstrom S, Sjostrom R, Linden S, Backstrom A, Lundberg C, Arnqvist A, Mahdavi J, Nilsson UJ, Velapatino B, Gilman RH, Gerhard M, Alarcon T, Lopez-Brea M, Nakazawa T, Fox JG, Correa P, Dominguez-Bello MG, Perez-Perez GI, Blaser MJ, Normark S, Carlstedt I, Oscarson S, Teneberg S, Berg DE, & Boren T (2004) Functional adaptation of BabA, the *H. pylori* ABO blood group antigen binding adhesin. *Science* 305(5683):519-522.
157. Ansari S & Yamaoka Y (2017) *Helicobacter pylori* BabA in adaptation for gastric colonization. *World J Gastroenterol* 23(23):4158-4169.
158. Pachathundikandi SK, Tegtmeyer N, & Backert S (2013) Signal transduction of *Helicobacter pylori* during interaction with host cell protein receptors of epithelial and immune cells. *Gut Microbes* 4(6):454-474.
159. Mahdavi J, Sonden B, Hurtig M, Olfat FO, Forsberg L, Roche N, Angstrom J, Larsson T, Teneberg S, Karlsson KA, Altraja S, Wadstrom T, Kersulyte D, Berg DE, Dubois A, Petersson C, Magnusson KE, Norberg T, Lindh F, Lundskog BB, Arnqvist A, Hammarstrom L, & Boren T (2002) *Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation. *Science* 297(5581):573-578.
160. Gerhard M, Lehn N, Neumayer N, Boren T, Rad R, Schepp W, Miehle S, Classen M, & Prinz C (1999) Clinical relevance of the *Helicobacter pylori* gene for blood-group antigen-binding adhesin. *Proc Natl Acad Sci U S A* 96(22):12778-12783.

161. Backert S, Schwarz T, Miehlke S, Kirsch C, Sommer C, Kwok T, Gerhard M, Goebel UB, Lehn N, Koenig W, & Meyer TF (2004) Functional analysis of the cag pathogenicity island in *Helicobacter pylori* isolates from patients with gastritis, peptic ulcer, and gastric cancer. *Infect Immun* 72(2):1043-1056.
162. Noto JM & Peek RM, Jr. (2012) The *Helicobacter pylori* cag Pathogenicity Island. *Methods Mol Biol* 921:41-50.
163. Lai YP, Yang JC, Lin TZ, Lin JT, & Wang JT (2006) *Helicobacter pylori* infection and CagA protein translocation in human primary gastric epithelial cell culture. *Helicobacter* 11(5):451-459.
164. Odenbreit S, Puls J, Sedlmaier B, Gerland E, Fischer W, & Haas R (2000) Translocation of *Helicobacter pylori* CagA into gastric epithelial cells by type IV secretion. *Science* 287(5457):1497-1500.
165. Lee IO, Kim JH, Choi YJ, Pillinger MH, Kim SY, Blaser MJ, & Lee YC (2010) *Helicobacter pylori* CagA phosphorylation status determines the gp130-activated SHP2/ERK and JAK/STAT signal transduction pathways in gastric epithelial cells. *J Biol Chem* 285(21):16042-16050.
166. Song X, Chen HX, Wang XY, Deng XY, Xi YX, He Q, Peng TL, Chen J, Chen W, Wong BC, & Chen MH (2013) *H. pylori*-encoded CagA disrupts tight junctions and induces invasiveness of AGS gastric carcinoma cells via Cdx2-dependent targeting of Claudin-2. *Cell Immunol* 286(1-2):22-30.
167. Saadat I, Higashi H, Obuse C, Umeda M, Murata-Kamiya N, Saito Y, Lu H, Ohnishi N, Azuma T, Suzuki A, Ohno S, & Hatakeyama M (2007) *Helicobacter pylori* CagA targets PAR1/MARK kinase to disrupt epithelial cell polarity. *Nature* 447(7142):330-333.
168. Smoot DT, Wynn Z, Elliott TB, Allen CR, Mekasha G, Naab T, & Ashktorab H (1999) Effects of *Helicobacter pylori* on proliferation of gastric epithelial cells in vitro. *Am J Gastroenterol* 94(6):1508-1511.
169. Jones KR, Whitmire JM, & Merrell DS (2010) A Tale of Two Toxins: *Helicobacter pylori* CagA and VacA Modulate Host Pathways that Impact Disease. *Front Microbiol* 1:115.

170. Yahiro K, Akazawa Y, Nakano M, Suzuki H, Hisatune J, Isomoto H, Sap J, Noda M, Moss J, & Hirayama T (2015) *Helicobacter pylori* VacA induces apoptosis by accumulation of connexin 43 in autophagic vesicles via a Rac1/ERK-dependent pathway. *Cell Death Discov* 1:15035.
171. Cover TL & Blanke SR (2005) *Helicobacter pylori* VacA, a paradigm for toxin multifunctionality. *Nat Rev Microbiol* 3(4):320-332.
172. Cover TL & Blaser MJ (1992) Purification and characterization of the vacuolating toxin from *Helicobacter pylori*. *J Biol Chem* 267(15):10570-10575.
173. Parker H & Keenan JI (2012) Composition and function of *Helicobacter pylori* outer membrane vesicles. *Microbes Infect* 14(1):9-16.
174. Pathirana RD & Kaparakis-Liaskos M (2016) Bacterial membrane vesicles: Biogenesis, immune regulation and pathogenesis. *Cell Microbiol* 18(11):1518-1524.
175. Mashburn LM & Whiteley M (2005) Membrane vesicles traffic signals and facilitate group activities in a prokaryote. *Nature* 437(7057):422-425.
176. Sturgis JN (2001) Organisation and evolution of the *tol-pal* gene cluster. *J Mol Microbiol Biotechnol* 3(1):113-122.
177. Dubuisson JF, Vianney A, Hugouvieux-Cotte-Pattat N, & Lazzaroni JC (2005) *Tol-Pal* proteins are critical cell envelope components of *Erwinia chrysanthemi* affecting cell morphology and virulence. *Microbiology* 151(Pt 10):3337-3347.
178. Gerding MA, Ogata Y, Pecora ND, Niki H, & de Boer PA (2007) The trans-envelope *Tol-Pal* complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in *E. coli*. *Mol Microbiol* 63(4):1008-1025.
179. Turner L, Praszquier J, Hutton ML, Steer D, Ramm G, Kaparakis-Liaskos M, & Ferrero RL (2015) Increased Outer Membrane Vesicle Formation in a *Helicobacter pylori* tolB Mutant. *Helicobacter*.
180. Yeh YC, Comolli LR, Downing KH, Shapiro L, & McAdams HH (2010) The *caulobacter Tol-Pal* complex is essential for outer membrane integrity and the positioning of a polar localization factor. *J Bacteriol* 192(19):4847-4858.

181. Waterhouse AM, Procter JB, Martin DM, Clamp M, & Barton GJ (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25(9):1189-1191.
182. Heymann JA & Hinshaw JE (2009) Dynamins at a glance. *J Cell Sci* 122(Pt 19):3427-3431.
183. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792-1797.
184. Tamura K, Stecher G, Peterson D, Filipski A, & Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30(12):2725-2729.
185. Finn RD, Clements J, Arndt W, Miller BL, Wheeler TJ, Schreiber F, Bateman A, & Eddy SR (2015) HMMER web server: 2015 update. *Nucleic Acids Res* 43(W1):W30-38.
186. Sonnhammer EL, von Heijne G, & Krogh A (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* 6:175-182.
187. Kelley LA, Mezulis S, Yates CM, Wass MN, & Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10(6):845-858.
188. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K, Quackenbush J, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glodek A, McKenney K, Fitzgerald LM, Lee N, Adams MD, Hickey EK, Berg DE, Gocayne JD, Utterback TR, Peterson JD, Kelley JM, Cotton MD, Weidman JM, Fujii C, Bowman C, Watthey L, Wallin E, Hayes WS, Borodovsky M, Karp PD, Smith HO, Fraser CM, & Venter JC (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388(6642):539-547.
189. Binh TT, Suzuki R, Shiota S, Kwon DH, & Yamaoka Y (2013) Complete Genome Sequences of *Helicobacter pylori* Clarithromycin-Resistant Strains. *Genome Announc* 1(6).
190. Binh TT, Suzuki R, Kwon DH, & Yamaoka Y (2015) Complete Genome Sequence of a Metronidazole-Resistant *Helicobacter pylori* Strain. *Genome Announc* 3(2).

191. Anand B, Verma SK, & Prakash B (2006) Structural stabilization of GTP-binding domains in circularly permuted GTPases: implications for RNA binding. *Nucleic Acids Res* 34(8):2196-2205.
192. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, Chabas S, Reiche K, Hackermuller J, Reinhardt R, Stadler PF, & Vogel J (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 464(7286):250-255.
193. Pernitzsch SR, Tirier SM, Beier D, & Sharma CM (2014) A variable homopolymeric G-repeat defines small RNA-mediated posttranscriptional regulation of a chemotaxis receptor in *Helicobacter pylori*. *Proc Natl Acad Sci U S A* 111(4):E501-510.
194. Torres-Cruz J & van der Woude MW (2003) Slipped-strand mispairing can function as a phase variation mechanism in *Escherichia coli*. *J Bacteriol* 185(23):6990-6994.
195. Bergman M, Del Prete G, van Kooyk Y, & Appelmek B (2006) *Helicobacter pylori* phase variation, immune modulation and gastric autoimmunity. *Nat Rev Microbiol* 4(2):151-159.
196. de Vries N, Duinsbergen D, Kuipers EJ, Pot RGJ, Wiesenekker P, Penn CW, van Vliet AHM, Vandenbroucke-Grauls CMJE, & Kusters JG (2002) Transcriptional Phase Variation of a Type III Restriction-Modification System in *Helicobacter pylori*. *Journal of Bacteriology* 184(23):6615-6623.
197. Moran AP, Knirel YA, Senchenkova SN, Widmalm G, Hynes SO, & Jansson PE (2002) Phenotypic variation in molecular mimicry between *Helicobacter pylori* lipopolysaccharides and human gastric epithelial cell surface glycoforms. Acid-induced phase variation in Lewis(x) and Lewis(y) expression by H. Pylori lipopolysaccharides. *J Biol Chem* 277(8):5785-5795.
198. Saunders NJ, Peden JF, Hood DW, & Moxon ER (1998) Simple sequence repeats in the *Helicobacter pylori* genome. *Mol Microbiol* 27(6):1091-1098.
199. Snyder LA, Loman NJ, Linton JD, Langdon RR, Weinstock GM, Wren BW, & Pallen MJ (2010) Simple sequence repeats in *Helicobacter canadensis* and their role in phase variable expression and C-terminal sequence switching. *BMC Genomics* 11:67.

200. Tannaes T, Dekker N, Bukholm G, Bijlsma JJ, & Appelmelk BJ (2001) Phase variation in the *Helicobacter pylori* phospholipase A gene and its role in acid adaptation. *Infect Immun* 69(12):7334-7340.
201. Faelber K, Held M, Gao S, Posor Y, Haucke V, Noe F, & Daumke O (2012) Structural insights into Dynamin-mediated membrane fission. *Structure* 20(10):1621-1628.
202. Frohlich C, Grabiger S, Schwefel D, Faelber K, Rosenbaum E, Mears J, Rocks O, & Daumke O (2013) Structural insights into oligomerization and mitochondrial remodelling of Dynamin 1-like protein. *EMBO J* 32(9):1280-1292.
203. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, & Hochstrasser DF (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* 112:531-552.
204. Szwedziak P, Wang Q, Freund SM, & Lowe J (2012) FtsA forms actin-like protofilaments. *EMBO J* 31(10):2249-2260.
205. Villaverde A & Carrio MM (2003) Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol Lett* 25(17):1385-1395.
206. Tsumoto K, Ejima D, Kumagai I, & Arakawa T (2003) Practical considerations in refolding proteins from inclusion bodies. *Protein Expression and Purification* 28(1):1-8.
207. Chappie JS & Dyda F (2013) Building a fission machine--structural insights into Dynamin assembly and activation. *J Cell Sci* 126(Pt 13):2773-2784.
208. Ventura S & Villaverde A (2006) Protein quality in bacterial inclusion bodies. *Trends Biotechnol* 24(4):179-185.
209. Backert S, Kwok T, Schmid M, Selbach M, Moese S, Peek RM, Jr., Konig W, Meyer TF, & Jungblut PR (2005) Subproteomes of soluble and structure-bound *Helicobacter pylori* proteins analyzed by two-dimensional gel electrophoresis and mass spectrometry. *Proteomics* 5(5):1331-1345.
210. Bumann D, Jungblut PR, & Meyer TF (2004) *Helicobacter pylori* vaccine development based on combined subproteome analysis. *Proteomics* 4(10):2843-2848.

211. Carlsohn E, Nystrom, J., Karlsson, H., Svennerholm, A., Nilsson, C.L (2006) Characterization of the Outer Membrane Protein Profile from Disease-Related Isolates by Subcellular Fractionation and Nano-LC FT-ICR MS Analysis. *J Proteome Res* 5:3197-3204.
212. Attaran B, Falsafi T, & Moghaddam AN (2016) Study of biofilm formation in C57Bl/6J mice by clinical isolates of *Helicobacter pylori*. *Saudi J Gastroenterol* 22(2):161-168.
213. Gopal GJ, Kumar A, Pal J, & Mukhopadhyay G (2014) Molecular characterization and polyclonal antibody generation against core component CagX protein of *Helicobacter pylori* type IV secretion system. *Bioengineered* 5(2):107-113.
214. Boyanova L, Panov V, Yordanov D, Gergova G, & Mitov I (2013) Characterization of oral *Helicobacter pylori* strain by 4 methods. *Diagn Microbiol Infect Dis* 77(4):287-288.
215. Faulde M, Putzker M, Mertes T, & Sobe D (1991) Evaluation of an immunofluorescence assay for specific detection of immunoglobulin G antibodies directed against *Helicobacter pylori*, and antigenic cross-reactivity between H. pylori and Campylobacter jejuni. *J Clin Microbiol* 29(2):323-327.
216. Rivera E, Lopez-Vidal Y, Luqueno V, & Ruiz-Palacios GM (1991) Indirect immunofluorescence assay for detection of *Helicobacter pylori* in human gastric mucosal biopsies. *J Clin Microbiol* 29(8):1748-1751.
217. Hirai Y, Haque M, Yoshida T, Yokota K, Yasuda T, & Oguma K (1995) Unique cholesteryl glucosides in *Helicobacter pylori*: composition and structural analysis. *Journal of Bacteriology* 177(18):5327-5333.
218. Low HH & Lowe J (2010) Dynamin architecture--from monomer to polymer. *Curr Opin Struct Biol* 20(6):791-798.
219. Tauschek M, Gorrell RJ, Strugnell RA, & Robins-Browne RM (2002) Identification of a protein secretory pathway for the secretion of heat-labile enterotoxin by an enterotoxigenic strain of *Escherichia coli*. *Proc Natl Acad Sci U S A* 99(10):7066-7071.

220. Copass M, Grandi G, & Rappuoli R (1997) Introduction of unmarked mutations in the *Helicobacter pylori* vacA gene with a sucrose sensitivity marker. *Infect Immun* 65(5):1949-1952.
221. Asim M, Chikara SK, Ghosh A, Vudathala S, Romero-Gallo J, Krishna US, Wilson KT, Israel DA, Peek RM, Jr., & Chaturvedi R (2015) Draft Genome Sequence of Gerbil-Adapted Carcinogenic *Helicobacter pylori* Strain 7.13. *Genome Announc* 3(3).
222. Farnbacher M, Jahns T, Willrodt D, Daniel R, Haas R, Goesmann A, Kurtz S, & Rieder G (2010) Sequencing, annotation, and comparative genome analysis of the gerbil-adapted *Helicobacter pylori* strain B8. *BMC Genomics* 11:335.
223. Darling AC, Mau B, Blattner FR, & Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14(7):1394-1403.
224. Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30(14):2068-2069.
225. Langdon WB (2015) Performance of genetic programming optimised Bowtie2 on genome comparison and analytic testing (GCAT) benchmarks. *BioData Min* 8(1):1.
226. Carver T, Harris SR, Berriman M, Parkhill J, & McQuillan JA (2012) Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28(4):464-469.
227. Alikhan NF, Petty NK, Ben Zakour NL, & Beatson SA (2011) BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402.
228. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, & Genome Project Data Processing S (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078-2079.
229. Draper JL, Hansen LM, Bernick DL, Abedrabbo S, Underwood JG, Kong N, Huang BC, Weis AM, Weimer BC, van Vliet AH, Pourmand N, Solnick JV, Karplus K, & Ottemann KM (2017) Fallacy of the Unique Genome: Sequence Diversity within Single *Helicobacter pylori* Strains. *MBio* 8(1).
230. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, & Madden TL (2009) BLAST+: architecture and applications. *BMC Bioinformatics* 10:421.

231. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, & Yeh LS (2004) UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res* 32(Database issue):D115-119.
232. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer EL, Eddy SR, Bateman A, & Finn RD (2012) The Pfam protein families database. *Nucleic Acids Res* 40(Database issue):D290-301.
233. Haft DH, Selengut JD, Richter RA, Harkins D, Basu MK, & Beck E (2013) TIGRFAMs and Genome Properties in 2013. *Nucleic Acids Res* 41(Database issue):D387-395.
234. Drazek ES, Dubois A, Holmes RK, Kersulyte D, Akopyants NS, Berg DE, & Warren RL (1995) Cloning and characterization of hemolytic genes from *Helicobacter pylori*. *Infect Immun* 63(11):4345-4349.
235. Rektorschek M, Weeks D, Sachs G, & Melchers K (1998) Influence of pH on metabolism and urease activity of *Helicobacter pylori*. *Gastroenterology* 115(3):628-641.
236. Park SJ, Son WS, & Lee BJ (2012) Structural Analysis of Hypothetical Proteins from *Helicobacter pylori*: An Approach to Estimate Functions of Unknown or Hypothetical Proteins. *Int J Mol Sci* 13(6):7109-7137.
237. Corbinais C, Mathieu A, Damke PP, Kortulewski T, Busso D, Prado-Acosta M, Radicella JP, & Marsin S (2017) ComB proteins expression levels determine *Helicobacter pylori* competence capacity. *Sci Rep* 7:41495.
238. de Vos WM, Venema G, Canosi U, & Trautner TA (1981) Plasmid transformation in *Bacillus subtilis*: fate of plasmid DNA. *Mol Gen Genet* 181(4):424-433.
239. Javadi MB & Katzenmeier G (2016) The Forgotten Virulence Factor: The 'non-conventional' Hemolysin TlyA And Its Role in *Helicobacter pylori* Infection. *Curr Microbiol* 73(6):930-937.
240. Lata K, Paul K, & Chattopadhyay K (2014) Functional characterization of *Helicobacter pylori* TlyA: pore-forming hemolytic activity and cytotoxic property of the protein. *Biochem Biophys Res Commun* 444(2):153-157.

241. Muir S, Koopman MB, Libby SJ, Joens LA, Heffron F, & Kusters JG (1992) Cloning and expression of a *Serpula (Treponema) hyodysenteriae* hemolysin gene. *Infect Immun* 60(2):529-535.
242. Wren BW, Stabler RA, Das SS, Butcher PD, Mangan JA, Clarke JD, Casali N, Parish T, & Stoker NG (1998) Characterization of a haemolysin from *Mycobacterium tuberculosis* with homology to a virulence factor of *Serpulina hyodysenteriae*. *Microbiology* 144 (Pt 5):1205-1211.
243. Feder M, Pas J, Wyrwicz LS, & Bujnicki JM (2003) Molecular phylogenetics of the RrmJ/fibrillarin superfamily of ribose 2'-O-methyltransferases. *Gene* 302(1-2):129-138.
244. Segal ED & Tompkins LS (1993) Identification and characterization of a *Helicobacter pylori* hemolysin. *Infect Agents Dis* 2(4):178-182.
245. Lata K & Chattopadhyay K (2014) *Helicobacter pylori* TlyA agglutinates liposomes and induces fusion and permeabilization of the liposome membranes. *Biochemistry* 53(22):3553-3563.
246. Eaton KA & Krakowka S (1994) Effect of gastric pH on urease-dependent colonization of gnotobiotic piglets by *Helicobacter pylori*. *Infect Immun* 62(9):3604-3607.
247. Ferrero RL, Cussac V, Courcoux P, & Labigne A (1992) Construction of isogenic urease-negative mutants of *Helicobacter pylori* by allelic exchange. *J Bacteriol* 174(13):4212-4217.
248. Tsuda M, Karita M, Mizote T, Morshed MG, Okita K, & Nakazawa T (1994) Essential role of *Helicobacter pylori* urease in gastric colonization: definite proof using a urease-negative mutant constructed by gene replacement. *Eur J Gastroenterol Hepatol* 6 Suppl 1:S49-52.
249. Tsuda M, Karita M, Morshed MG, Okita K, & Nakazawa T (1994) A urease-negative mutant of *Helicobacter pylori* constructed by allelic exchange mutagenesis lacks the ability to colonize the nude mouse stomach. *Infect Immun* 62(8):3586-3589.
250. Weeks DL, Eskandari S, Scott DR, & Sachs G (2000) A H⁺-gated urea channel: the link between *Helicobacter pylori* urease and gastric colonization. *Science* 287(5452):482-485.

251. Pflock M, Kennard S, Delany I, Scarlato V, & Beier D (2005) Acid-induced activation of the urease promoters is mediated directly by the ArsRS two-component system of *Helicobacter pylori*. *Infect Immun* 73(10):6437-6445.
252. Croxen MA, Sisson G, Melano R, & Hoffman PS (2006) The *Helicobacter pylori* chemotaxis receptor TlpB (HP0103) is required for pH taxis and for colonization of the gastric mucosa. *J Bacteriol* 188(7):2656-2665.
253. Larhed AW, Artursson P, Grasjo J, & Bjork E (1997) Diffusion of drugs in native and purified gastrointestinal mucus. *J Pharm Sci* 86(6):660-665.
254. Schubert OT, Gillet LC, Collins BC, Navarro P, Rosenberger G, Wolski WE, Lam H, Amodei D, Mallick P, MacLean B, & Aebersold R (2015) Building high-quality assay libraries for targeted analysis of SWATH MS data. *Nat Protoc* 10(3):426-441.
255. Wiktorowicz JE, English RD, Wu Z, & Kurosky A (2012) Model studies on iTRAQ modification of peptides: sequence-dependent reaction specificity. *J Proteome Res* 11(3):1512-1520.
256. Ow SY, Salim M, Noirel J, Evans C, & Wright PC (2011) Minimising iTRAQ ratio compression through understanding LC-MS elution dependence and high-resolution HILIC fractionation. *Proteomics* 11(11):2341-2346.
257. Ow SY, Salim M, Noirel J, Evans C, Rehman I, & Wright PC (2009) iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly". *J Proteome Res* 8(11):5347-5355.
258. Huang Q, Yang L, Luo J, Guo L, Wang Z, Yang X, Jin W, Fang Y, Ye J, Shan B, & Zhang Y (2015) SWATH enables precise label-free quantification on proteome scale. *Proteomics* 15(7):1215-1223.
259. Perkins DN, Pappin DJ, Creasy DM, & Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20(18):3551-3567.
260. Li B, Tian X, Wang C, Zeng X, Xing Y, Ling H, Yin W, Tian L, Meng Z, Zhang J, & Guo S (2017) SWATH label-free proteomics analyses revealed the roles of oxidative stress and antioxidant defending system in sclerotia formation of *Polyporus umbellatus*. *Sci Rep* 7:41283.

261. Grosseholz R, Koh CC, Veith N, Fiedler T, Strauss M, Olivier B, Collins BC, Schubert OT, Bergmann F, Kreikemeyer B, Aebersold R, & Kummer U (2016) Integrating highly quantitative proteomics and genome-scale metabolic modeling to study pH adaptation in the human pathogen *Enterococcus faecalis*. *NPJ Syst Biol Appl* 2:16017.
262. Kamath KS, Krisp C, Chick J, Pascovici D, Gygi SP, & Molloy MP (2017) *Pseudomonas aeruginosa* Proteome under Hypoxic Stress Conditions Mimicking the Cystic Fibrosis Lung. *J Proteome Res*.
263. Loke MF, Ng CG, Vilashni Y, Lim J, & Ho B (2016) Understanding the dimorphic lifestyles of human gastric pathogen *Helicobacter pylori* using the SWATH-based proteomics approach. *Sci Rep* 6:26784.
264. Appelmelk BJ, Shiberu B, Trinks C, Tapsi N, Zheng PY, Verboom T, Maaskant J, Hokke CH, Schiphorst WE, Blanchard D, Simoons-Smit IM, van den Eijnden DH, & Vandenbroucke-Grauls CM (1998) Phase variation in *Helicobacter pylori* lipopolysaccharide. *Infect Immun* 66(1):70-76.
265. Richter C, Mukherjee O, Ermert D, Singh B, Su YC, Agarwal V, Blom AM, & Riesbeck K (2016) Moonlighting of *Helicobacter pylori* catalase protects against complement-mediated killing by utilising the host molecule vitronectin. *Sci Rep* 6:24391.
266. Tacchi JL, Raymond BB, Jarocki VM, Berry IJ, Padula MP, & Djordjevic SP (2014) Cilium adhesin P216 (MHJ_0493) is a target of ectodomain shedding and aminopeptidase activity on the surface of *Mycoplasma hyopneumoniae*. *J Proteome Res* 13(6):2920-2930.
267. Damke PP, Dhanaraju R, Marsin S, Radicella JP, & Rao DN (2016) Mutations in the nucleotide binding and hydrolysis domains of *Helicobacter pylori* MutS2 lead to altered biochemical activities and inactivation of its in vivo function. *BMC Microbiol* 16:14.
268. Aspinall GO, Monteiro MA, Shaver RT, Kurjanczyk LA, & Penner JL (1997) Lipopolysaccharides of *Helicobacter pylori* serogroups O:3 and O:6--structures of a class of lipopolysaccharides with reference to the location of oligomeric units of D-glycero-alpha-D-manno-heptose residues. *Eur J Biochem* 248(2):592-601.

269. Benoit S, Mehta N, Wang G, Gatlin M, & Maier RJ (2004) Requirement of hydD, hydE, hypC and hypE genes for hydrogenase activity in *Helicobacter pylori*. *Microb Pathog* 36(3):153-157.
270. Kumari R, Shariq M, Kumar N, & Mukhopadhyay G (2017) Biochemical characterization of the *Helicobacter pylori* Cag-type IV secretion system unique component CagU. *FEBS Lett* 591(3):500-512.
271. Lower M, Weydig C, Metzler D, Reuter A, Starzinski-Powitz A, Wessler S, & Schneider G (2008) Prediction of extracellular proteases of the human pathogen *Helicobacter pylori* reveals proteolytic activity of the Hp1018/19 protein HtrA. *PLoS One* 3(10):e3510.
272. Zhang H & Au SWN (2017) *Helicobacter pylori* does not use spermidine synthase to produce spermidine. *Biochem Biophys Res Commun* 490(3):861-867.
273. Bertini I, Cavallaro G, & Rosato A (2006) Cytochrome c: occurrence and functions. *Chem Rev* 106(1):90-115.
274. Huang CH, Lee IL, Yeh IJ, Liao JH, Ni CL, Wu SH, & Chiou SH (2010) Upregulation of a non-heme iron-containing ferritin with dual ferroxidase and DNA-binding activities in *Helicobacter pylori* under acid stress. *J Biochem* 147(4):535-543.
275. Padula MP, Berry IJ, MB OR, Raymond BB, Santos J, & Djordjevic SP (2017) A Comprehensive Guide for Performing Sample Preparation and Top-Down Protein Analysis. *Proteomes* 5(2).
276. Muller SA, Pernitzsch SR, Haange SB, Uetz P, von Bergen M, Sharma CM, & Kalkhof S (2015) Stable isotope labeling by amino acids in cell culture based proteomics reveals differences in protein abundances between spiral and coccoid forms of the gastric pathogen *Helicobacter pylori*. *J Proteomics* 126:34-45.
277. Pitcher RS & Watmough NJ (2004) The bacterial cytochrome cbb3 oxidases. *Biochim Biophys Acta* 1655(1-3):388-399.
278. Pinto AV, Mathieu A, Marsin S, Veaute X, Ielpi L, Labigne A, & Radicella JP (2005) Suppression of homologous and homeologous recombination by the bacterial MutS2 protein. *Mol Cell* 17(1):113-120.

279. Waidner B, Greiner S, Odenbreit S, Kavermann H, Velayudhan J, Stahler F, Guhl J, Bisse E, van Vliet AH, Andrews SC, Kusters JG, Kelly DJ, Haas R, Kist M, & Bereswill S (2002) Essential role of ferritin Pfr in *Helicobacter pylori* iron metabolism and gastric colonization. *Infect Immun* 70(7):3923-3929.
280. Bereswill S, Waidner U, Odenbreit S, Lichte F, Fassbinder F, Bode G, & Kist M (1998) Structural, functional and mutational analysis of the pfr gene encoding a ferritin from *Helicobacter pylori*. *Microbiology* 144 (Pt 9):2505-2516.
281. Bereswill S, Greiner S, van Vliet AH, Waidner B, Fassbinder F, Schiltz E, Kusters JG, & Kist M (2000) Regulation of ferritin-mediated cytoplasmic iron storage by the ferric uptake regulator homolog (Fur) of *Helicobacter pylori*. *J Bacteriol* 182(21):5948-5953.
282. Bury-Mone S, Thiberge JM, Contreras M, Maitournam A, Labigne A, & De Reuse H (2004) Responsiveness to acidity via metal ion regulators mediates virulence in the gastric pathogen *Helicobacter pylori*. *Mol Microbiol* 53(2):623-638.
283. Scott DR, Marcus EA, Wen Y, Oh J, & Sachs G (2007) Gene expression in vivo shows that *Helicobacter pylori* colonizes an acidic niche on the gastric surface. *Proc Natl Acad Sci U S A* 104(17):7235-7240.
284. Pflock M, Kennard S, Finsterer N, & Beier D (2006) Acid-responsive gene regulation in the human pathogen *Helicobacter pylori*. *J Biotechnol* 126(1):52-60.
285. Leyer GJ & Johnson EA (1993) Acid adaptation induces cross-protection against environmental stresses in *Salmonella typhimurium*. *Appl Environ Microbiol* 59(6):1842-1847.
286. Leyer GJ & Johnson EA (1992) Acid adaptation promotes survival of *Salmonella spp.* in cheese. *Appl Environ Microbiol* 58(6):2075-2080.
287. Leyer GJ, Wang LL, & Johnson EA (1995) Acid adaptation of *Escherichia coli* O157:H7 increases survival in acidic foods. *Appl Environ Microbiol* 61(10):3752-3755.
288. Liu Y, Tang H, Lin Z, & Xu P (2015) Mechanisms of acid tolerance in bacteria and prospects in biotechnology and bioremediation. *Biotechnol Adv* 33(7):1484-1492.
289. Cotter PD & Hill C (2003) Surviving the Acid Test: Responses of Gram-Positive Bacteria to Low pH. *Microbiology and Molecular Biology Reviews* 67(3):429-453.

290. Schoenhofen IC, Vinogradov E, Whitfield DM, Brisson JR, & Logan SM (2009) The CMP-legionaminic acid pathway in *Campylobacter*: biosynthesis involving novel GDP-linked precursors. *Glycobiology* 19(7):715-725.
291. Schirm M, Schoenhofen IC, Logan SM, Waldron KC, & Thibault P (2005) Identification of unusual bacterial glycosylation by tandem mass spectrometry analyses of intact proteins. *Anal Chem* 77(23):7774-7782.
292. Schirm M, Soo EC, Aubry AJ, Austin J, Thibault P, & Logan SM (2003) Structural, genetic and functional characterization of the flagellin glycosylation process in *Helicobacter pylori*. *Mol Microbiol* 48(6):1579-1592.
293. Sachs G, Scott DR, Weeks DL, Rektorscheck M, & Melchers K (2001) Regulation of Urease for Acid Habitation. *Helicobacter pylori: Physiology and Genetics*, eds Mobley HLT, Mendz GL, & Hazell SL Washington (DC)).
294. Mobley HLT (2001) Urease. *Helicobacter pylori: Physiology and Genetics*, eds Mobley HLT, Mendz GL, & Hazell SL Washington (DC)).
295. Rektorschek M, Buhmann A, Weeks D, Schwan D, Bensch KW, Eskandari S, Scott D, Sachs G, & Melchers K (2000) Acid resistance of *Helicobacter pylori* depends on the UreI membrane protein and an inner membrane proton barrier. *Mol Microbiol* 36(1):141-152.
296. Scott DR, Marcus EA, Wen Y, Singh S, Feng J, & Sachs G (2010) Cytoplasmic histidine kinase (HP0244)-regulated assembly of urease with UreI, a channel for urea and its metabolites, CO₂, NH₃, and NH₄(+), is necessary for acid survival of *Helicobacter pylori*. *J Bacteriol* 192(1):94-103.
297. Delcour AH (2009) Outer membrane permeability and antibiotic resistance. *Biochim Biophys Acta* 1794(5):808-816.
298. Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* 67(4):593-656.
299. Heyde M & Portalier R (1987) Regulation of major outer membrane porin proteins of *Escherichia coli* K 12 by pH. *Mol Gen Genet* 208(3):511-517.
300. Thomas AD & Booth IR (1992) The regulation of expression of the porin gene *ompC* by acid pH. *J Gen Microbiol* 138(9):1829-1835.

301. Gu H (2017) Role of Flagella in the Pathogenesis of *Helicobacter pylori*. *Curr Microbiol* 74(7):863-869.
302. Garber M, Grabherr MG, Guttman M, & Trapnell C (2011) Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat Methods* 8(6):469-477.
303. Wang Z, Gerstein M, & Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10(1):57-63.
304. Graham DY & Fischbach L (2010) *Helicobacter pylori* treatment in the era of increasing antibiotic resistance. *Gut* 59(8):1143-1153.
305. Gancz H, Jones KR, & Merrell DS (2008) Sodium chloride affects *Helicobacter pylori* growth and gene expression. *J Bacteriol* 190(11):4100-4105.
306. Cooksley C, Jenks PJ, Green A, Cockayne A, Logan RP, & Hardie KR (2003) NapA protects *Helicobacter pylori* from oxidative stress damage, and its production is influenced by the ferric uptake regulator. *J Med Microbiol* 52(Pt 6):461-469.
307. Ahmed N, Loke MF, Kumar N, & Vadivelu J (2013) *Helicobacter pylori* in 2013: multiplying genomes, emerging insights. *Helicobacter* 18 Suppl 1:1-4.
308. Muller-Deile J, Teng B, Schenk H, Haller H, Reiser J, Sever S, & Schiffer M (2016) Drugs targeting Dynamin can restore cytoskeleton and focal contact alterations of urinary podocytes derived from patients with nephrotic syndrome. *Ann Transl Med* 4(21):439.
309. Schiffer M, Teng B, Gu C, Shchedrina VA, Kasaikina M, Pham VA, Hanke N, Rong S, Gueler F, Schroder P, Tossidou I, Park JK, Staggs L, Haller H, Erschow S, Hilfiker-Kleiner D, Wei C, Chen C, Tardi N, Hakrrouch S, Selig MK, Vasilyev A, Merscher S, Reiser J, & Sever S (2015) Pharmacological targeting of actin-dependent Dynamin oligomerization ameliorates chronic kidney disease in diverse animal models. *Nat Med* 21(6):601-609.
310. Chircop M, Perera S, Mariana A, Lau H, Ma MP, Gilbert J, Jones NC, Gordon CP, Young KA, Morokoff A, Sakoff J, O'Brien TJ, McCluskey A, & Robinson PJ (2011) Inhibition of Dynamin by dynole 34-2 induces cell death following cytokinesis failure in cancer cells. *Mol Cancer Ther* 10(9):1553-1562.

The Characterisation of Dynamin-Like Proteins in the Gastric Pathogen *Helicobacter pylori*

Emma Mary Dawson

S1: *H. pylori* (1997) Dlp1a, Dlp1b and Dlp2 as well as *H. pylori* (2013) Dlp1 Sequences

In the following sequences, the sequence anomaly noted in *dlp1a* and *dlpb* from the earlier *H. pylori* sequence (1997) has been highlighted. Also seen below is the more recent *H. pylori* sequence for the ‘full dynamin’ dlp1, with the single nucleotide insertion of G highlighted.

>gi|15644634:c787397-785676 *Helicobacter pylori* 26695 chromosome, complete genome (Dlp2)

```
ATGAAGAATGAAACGCTAGAACAAATTTAAAAGAAATCAAAAGAGGAATCAAGAAATTCCTTAAAAAATTGC
TTGATTTTGTCCATGCTGGAGAGAAATACGGCATCCAGATTGAAGAGTCCCTCAAAGACAAAATCCGTAA
TGCGATAAAAAATGTTGCCGATCAAAACTGAAAGTGGCGTTGGTGGGAGGTTTTTCTGAAGGAAAAACA
TCTATAGCGGCAGCGTGGATCGATCGTTTGGATGAGGGCATGAAGATAGACCATCAAGAATCAAGCGATG
CAGTTAAAATCTATGACATAGACAATGAAATGGAGCTGGTTGACACCCCGGGACTATTCGGGTTTAAAGA
AAAAATAACCGATAGCGGCAAAATAGAACGCTATAAAGATATTACGAAAAAATATATCAGCGAAGCCCAC
CTCATTTTATACGCGCTCAACCCGTCAAACCCCATCAAAGACAGCCATAAGGACGATTTGAACTGGTTGT
TCAGGACGCTCAATCTTTTATCCAGGACGATCTTTGTCATCAGCCGTTTTTGATGAAGAAGCGGATATTGA
AGATGAAGAAGATTACAACAAAAGATTTGAAATCAAAAAAGAAAACGTTCAAACCGACTGAATGATCTC
ATTTCTTTAAGCGAAAAAGAAAAAGAGGGTTTGAGTATTGTCGCTGTCGCTGCAAACCCTTATGATTTGG
GGGTTGAACACTGGCTAAAGCATAAAGAAGAATTCCAAAAGCTTTCGCACATCAAACTTTGCAAGATGC
GACGCAAAAAAAGATTAAAGAAAATGGCGGGAAATTAACCATCATAGAAGAAGCCAAAAAAGCGTCATT
CAAGATGTGGTCTATAGGCAAATGCCGCCTGCAAAACAAGCGTTGCAAGATATTAATAGAGAAATGGAAT
ATTTGAATAAAACGCTTGAAAAAAGGCGAAAAGAGATTCAAATTTAAATAGTGAAATTTCTCAAGCCCG
CATTAAATTTAAGAGAATTTATAACGAGGTATTTTAGCGATCTCATCCGTCAAGTCCTTGGCACTAGCCTA
```

GAAACCTTTAATGATTTTGTATTAGAGAAATAGGCGATAAAGGTACCAATATTGACACAAGAGTCCAAA
ATGAATTTGAAAGACAAACGCAAGGGATTGTAAATGAAATTTCTAAAATTGAAACCGGTTTTAACGCCGA
TAGGAGTTTTTTTGAAAAGCATGCCGGAACATTGGGGAAAATTGGAATCAATTTTTTAAACAAAAGCGGG
TTTATCAATGCGACTAACATCAAACCTGGCTAGAGACACGATAGCGGCTGCGGGAAAATTTGTGGGCGTAG
ATTTGGCTTTTTAAATTCAAACCTTGGGGCGCTGTAACTTGGCAGGCAACTTGAACAAAGGTTTACCGCT
TATCGGTCTTGCTTTTGAAGTGTGGGATTCTTGAAAGAAAGCCAAAAATAGAAAACTTGAAAAAGCT
AAAAGAGAAATGGAATCCGATTTTAAACAATCAAAAACAAGAAATCTTAGATCTCATCAATAATGAAACCA
GATTCAAACAAACATGTTTCCCAATGGCGTTGGAAGTGGAAAAATGCATTCAAGAATGCGAAGAAAGCGT
TAAAAAACGCAAGAGTGCACCTCAAGGTCTGGAAAAATGGATTCAAACCGGCGAAGATTTGATCAAGGGC
GAAGATTTTATAGAAGTTGAGCCTGAAAGAGGCATGAACTGA

>gi|15644634:c787743-787387 *Helicobacter pylori* 26695 chromosome, complete
genome (Dlp1b)

TTGGAATCCTATGGGTTGGCTTGCTCTGACCGCA**GG**ATTATTGCAGGATTAGTTGGAATAGCCAGCTCAA
TCTGGAAATTTTTTAGTTCAAGATATCAAAGATCCCAACAAAGAAAAGAAGTGGATAAGAATTTGCATCA
AATTTGTGAAAAAATTGTGCAGGATGTGAAAAGCCGACTTGAAAGTCGAAAAAGACATATGGGAAAAG
ATTGAAAACTCAAAGCCAATCTTAGACCTGTTGACAATTACAAACGCATGAAAAGACAATTGAAAGAAG
CCCATGAAAGATTAAGATACATCTCTCATAGTATCCATCTAATAATATCAAAACAAGGAGCATGCAATGA
AGAATGA

>gi|15644634:c789257-787692 *Helicobacter pylori* 26695 chromosome, complete
genome (Dlp1a)

ATGAAAAACATTTATCTTGATGTGAAAGCCAGCATTGAAAATCTCCAAAATATTTTTAAAAACACTGATA
ATGAAAATGAAAGACTAAAAAATTCAACCAAGAAGCGTTGGAGGTGTTTCAAAAATTAGAGCGTGAAAG
TTTAAAAGAGCTTGAAAGCTTAAAAAATAATGAGGAGTGGGAAAATTTTACTATCGCTTTTTATGGGGAA
ACCGGTGCGGGGAAATCAACCTTCATTGAATGTTTGAGAATGTTTTTTAAAGAACAAAGTAAAGTAGTTC
AACAAGAACGATTCAAGCGGCTTTATTCCAATTACCAAAACAACCTATCAAATGATGAATGCAAAAAGCA
AGCTATTTTAAACGAACCTTCATTGCAAGATGGAGCGATCATAGGCGATGGGAGGAGCGATTTCACT
TTAAAAACACGATCTTATTCTTTCCAATACAACCATCAAACTTTACTTTGCTTGATGTTCCAGGGATAG

AAGGCGACGAAAAAAGTGATCGATCAGATTTCTAACGCAACGCAAAAAGCCCATGCTATTTTTTATGT
TACCAAACGCCTAATCCTCCGCAAAAAGGAGAAGAGAAAAAGAAGGGACGATTGAAAAATCCAAAA
CAACTTGATTCGCAACAGAGGTATGGACGATTTTTTAACAAACCGATTAACAACCCAAGAGCTTTCAAAG
ATGGGCTTATTGATGGAAGCGAAAAAGAAAGCTTAAAAATTTTAAATAAGAAATGAAAAACATTTTAGG
CAAACACTACAAGGGCTATAAAGCAGTGAGCGCCCAAGTGGCTTTTTATGGTCTTTCATCGGCTTTGATC
CCAGGGACTGATTTTGATAAAAACAAACAAAAATTTTTAAAGATTTTAAAGCAAGAGAATTATTGTATC
AATCCCATTTCCAACAATTAGGAGAATTTATAGCCGAAGAGCTTATTAAAACTCGCGTGCCAAAATCAT
TCAATCAAACGCAATAAAGCCTTAAAGTGGTAGAACAATTGCAAAAGGCGATCGAAATTACGATTGAA
AAACGGATCGATCCAATGATTAAAGAAGCACAAGAATACCAACACGAAGCCCGCTATAATCTGGATCGTT
CTACAGATAAATTTATATTAAATTTAACCAATTCAGCGTTCTACGAAATCGATCAATTCAAATCTGACTT
GAGAGAAAAAATGTATGCGCATATTAACAAAAATATTGAAGATGAGGAATGTAAAGAAATTTTTAAAAAT
GAACTCATTCAAGGAATTGAAACATTGCATGAAGACATAAAATGGCGGTTTAGAGAATGTGAGAAACGAT
TTGATGGAGAGATAAAAGAAGCTATTAAACAACCTGAATACAGAATTAAAGATTCTCTAGCAATGTTAGA
GCGCATCAGTATTGATAGAGACTTTAATCTTAATTTTGATACTGATAGCGGTATTGATGGAACAAAATTA
GCCACTTCAATAGGAGGTTTGGGTTTGCTTGGGATATTTAACGCTTGGAATCCTATGGGTTGGCTTGCTC
TGACCGCAGGATTATTGCAGGATTAG

>CP010436.1:c788877-787006 *Helicobacter pylori* strain 26695-1MET, complete
genome (Dlp1)

ATGAAAAACATTTATCTTGATGTGAAAGCCAGCATTGAAAATCTCCAAAATATTTTTAAAAACACTGATA
ATGAAAATGAAAGACTAAAAAATTCAACCAAGAAGCGTTGGAGGTGTTTCAAAAATTAGAGCGTGAAAG
TTTAAAAGAGCTTGAAAGCTTAAAAAATAATGAGGAGTGGGAAAATTTTACTATCGCTTTTTATGGGGAA
ACCGGTGCGGGGAAATCAACCTTCATTGAATGTTTGAGAATGTTTTTTAAAGAACAAAGTAAAGTAGTTC
AACAAGAACGATTCAAGCGGCTTTATTCCAATTACCAAAACAACTATCAAAATGATGAATGCAAAAAGCA
AGCTATTTTAAACGAACCTTCATTGCAAGATGGAGCGATCATAGGCGATGGGAGGAGCGATTTCACT
TTAAAAACACGATCTTATTCTTTCCAATACAACCATCAAACTTTACTTTGCTTGATGTTCCAGGGATAG
AAGGCGACGAAAAAAGTGATCGATCAGATTTCTAACGCAACGCAAAAAGCCCATGCTATTTTTTATGT
TACCAAACGCCTAATCCTCCGCAAAAAGGAGAAGAGAAAAAGAAGGGACGATTGAAAAATCCAAAA
CAACTTGATTCGCAACAGAGGTATGGACGATTTTTTAACAAACCGATTAACAACCCAAGAGCTTTCAAAG

ATGGGCTTATTGATGGAAGCGAAAAAGAAAGCTTAAAAATTTTAAATAAGAAATGAAAAACATTTTAGG
CAAACACTACAAGGGCTATAAAGCAGTGAGCGCCCAAGTGGCTTTTTATGGTCTTTCATCGGCTTTGATC
CCAGGGACTGATTTTGATAAAAACAAACAAAAATTTTAAAGATTTTAAAGCAAGAGAATTATTGTATC
AATCCCATTTCCAACAATTAGGAGAATTTATAGCCGAAGAGCTTATTA AAAACTCGCGTGCCAAAATCAT
TCAATCAA CTGCAATAAAGCCTTAAAGTGGTAGAACAATTGCAAAGGCGATCGAAATTACGATTGAA
AAACGGATCGATCCAATGATTAAAGAAGCACAGAATACCAACACGAAGCCCGCTATAATCTGGATCGTT
CTACAGATAAATTTATATTAAATTTAACCAATTCAGCGTTCTACGAAATCGATCAATTCAAATCTGACTT
GAGAGAAAAAATGTATGCGCATATTAACAAAAATATTGAAGATGAGGAATGTAAAGAAATTTTAAAAAT
GAACTCATTCAAGGAATTGAAACATTGCATGAAGACATAAAATGGCGGTTTAGAGAATGTGAGAAACGAT
TTGATGGAGAGATAAAAGAAGCTATTAAACAACCTGAATACAGAATTAAAGATTCTCTAGCAATGTTAGA
GCGCATCAGTATTGATAGAGACTTTAATCTTAATTTTGATACTGATAGCGGTATTGATGGAACAAAATTA
GCCACTTCAATAGGAGGTTTGGGTTTGCTTGGGATATTTAACGCTTGGAATCCTATGGGTTGGCTTGCTC
TGACCGCA GGGATTATTGCAGGATTAGTTGGAATAGCCAGCTCAATCTGGAAATTTTTTAGTTCAAGATA
TCAAAGATCCCAACAAAGAAAAGAAGTGGATAAGAATTTGCATCAAATTTGTGAAAAAATTGTGCAGGAT
GTGAAAAGCCGACTTGAAAGTCGCAAAAAGACATATGGGAAAAGATTGAAAACTCAAAGCCAATCTTA
GACCTGTTGACAATTACAAACGCATGAAAAGACAATTGAAAGAAGCCCATGAAAGATTAAGATACATCTC
TCATAGTATCCATCTAATAATATCAAAACAAGGAGCATGCAATGAAGAATGA

The Characterisation of Dynamin-Like Proteins in the Gastric Pathogen *Helicobacter pylori*

Emma Mary Dawson

S2: Conservation screen for Dynamin-Like proteins in *H. pylori*

Listed here are the strains with one, two or three putative dynamins detected in the *H. pylori* screen.

1. Dynamin detected in the following *Helicobacter spp* (not *pylori*) (N=5)

- *Helicobacter acinonychis* strain Sheeba
- *Helicobacter bizzozeronii* CCUG 35545
- *Helicobacter bizzozeronii* strain CIII 1
- *Helicobacter cecorum* strain ATCC BAA 4297128
- *Helicobacter macacae* MIT 99 5501

2. One putative dynamin detected in the following *H. pylori* strains (N=8)

- *Helicobacter pylori* Aklavik117
- *Helicobacter pylori* PZ5086
- *Helicobacter pylori* strain 35A
- *Helicobacter pylori* strain B38
- *Helicobacter pylori* strain Gambia94 24
- *Helicobacter pylori* strain Sat464
- *Helicobacter pylori* UM038
- *Helicobacter pylori* UM298

3. Two putative dynamins detected in the following *H. pylori* strains (N=19)

• <i>Helicobacter pylori</i> UM077	• <i>Helicobacter pylori</i> Shi417
• <i>Helicobacter pylori</i> strain v225d	• <i>Helicobacter pylori</i> PZ5056
• <i>Helicobacter pylori</i> strain Shi470	• <i>Helicobacter pylori</i> Puno135
• <i>Helicobacter pylori</i> strain PeCan4	• <i>Helicobacter pylori</i> PZ5004
• <i>Helicobacter pylori</i> strain India7	• <i>Helicobacter pylori</i> PeCan18
• <i>Helicobacter pylori</i> strain F32	• <i>Helicobacter pylori</i> Hp P 4c
• <i>Helicobacter pylori</i> Shi169	• <i>Helicobacter pylori</i> Hp P 4d
	• <i>Helicobacter pylori</i> Hp P 4

- Helicobacter pylori Hp A 14
- Helicobacter pylori FD577

- Helicobacter pylori CPY1962
- Helicobacter pylori CG IMSS 2012

4. Three putative dynamins (back-to-back) detected in the following *H. pylori* strains (N=136). They are listed corresponding to their ETEC homologs with A = LeoA homolog, B=LeoB homolog and C=LeoC homolog.

*Three putative dynamins detected in an additional 81 *H. pylori* strains but were found to not have consecutive gene names (accessions) and are therefore unlikely to be co expressed which was the focus of this thesis*

Helicobacter pylori GAM80Ai

- A. HMPREF1426_00727
- B. HMPREF1426_00728
- C. HMPREF1426_00729

Helicobacter pylori Hp H-21

- A. HPHPH21_0895
- B. HPHPH21_0896
- C. HPHPH21_0897

Helicobacter pylori HP116Bi

- A. HMPREF1437_01125
- B. HMPREF1437_01126
- C. HMPREF1437_01127

Helicobacter pylori Hp P-26

- A. HPHPP26_1404
- B. HPHPP26_1405
- C. HPHPP26_1406

Helicobacter pylori Hp P-13b

- A. HPHPP13B_1007
- B. HPHPP13B_1008
- C. HPHPP13B_1009

Helicobacter pylori Hp H-44

- A. HPHPH44_0956
- B. HPHPH44_0957
- C. HPHPH44_0958

Helicobacter pylori Hp P-13

- A. HPHPP13_1013
- B. HPHPP13_1014
- C. HPHPP13_1015

Helicobacter pylori GAM119Bi

- A. HMPREF1400_00679
- B. HMPREF1400_00680
- C. HMPREF1400_00681

Helicobacter pylori Hp P-2b

- A. HPHPP2B_0833
- B. HPHPP2B_0834
- C. HPHPP2B_0835

Helicobacter pylori (strain J99)

- A. jhp_0668
- B. jhp_0669
- C. jhp_0670

Helicobacter pylori Hp P-2

- A. HPHPP2_0829
- B. HPHPP2_0830
- C. HPHPP2_0831

Helicobacter pylori GAM71Ai

- A. HMPREF1425_00655
- B. HMPREF1425_00656
- C. HMPREF1425_00657

Helicobacter pylori GAM201Ai

- A. HMPREF1403_01621
- B. HMPREF1403_01622
- C. HMPREF1403_01623

Helicobacter pylori NQ4110

- A. HPNQ4110_0887
- B. HPNQ4110_0888
- C. HPNQ4110_0889

Helicobacter pylori Hp H-11

- A. HPHPH11_0948
- B. HPHPH11_0949
- C. HPHPH11_0950

Helicobacter pylori R055a

- A. OUQ_0960
- B. OUQ_0961
- C. OUQ_0962

Helicobacter pylori Hp A-6

- A. HPHPA6_0801
- B. HPHPA6_0802
- C. HPHPA6_0803

Helicobacter pylori Hp H-18

- A. HPHPH18_0958
- B. HPHPH18_0959
- C. HPHPH18_0960

Helicobacter pylori 26695

- A. HP_0731
- B. HP_0732
- C. HP_0733

Helicobacter pylori GAM112Ai

- A. HMPREF1395_01295
- B. HMPREF1395_01296
- C. HMPREF1395_01297

Helicobacter pylori GAM42Ai

- A. HMPREF1424_01146
- B. HMPREF1424_01147
- C. HMPREF1424_01148

Helicobacter pylori R018c

- A. OUC_1034
- B. OUC_1035
- C. OUC_1036

Helicobacter pylori R056a

- A. OUS_1141
- B. OUS_1142
- C. OUS_1143

Helicobacter pylori Hp A-26

- A. HPHPA26_0736
- B. HPHPA26_0737
- C. HPHPA26_0738

Helicobacter pylori GAM210Bi

- A. HMPREF1404_01695
- B. HMPREF1404_01696
- C. HMPREF1404_01697

Helicobacter pylori Hp P-15b

- A. HPHPP15B_1337
- B. HPHPP15B_1338
- C. HPHPP15B_1339

Helicobacter pylori Hp P-15

- A. HPHPP15_1070
- B. HPHPP15_1071
- C. HPHPP15_1072

Helicobacter pylori Hp P-74

- A. HPHPP74_1386
- B. HPHPP74_1387
- C. HPHPP74_1388

Helicobacter pylori (strain HPAG1)

- A. HPAG1_0715
- B. HPAG1_0716
- C. HPAG1_0717

Helicobacter pylori Hp H-27

- A. HPHPH27_1374
- B. HPHPH27_1375
- C. HPHPH27_1376

Helicobacter pylori Hp P-23

- A. HPHPP23_1113
- B. HPHPP23_1114
- C. HPHPP23_1115

Helicobacter pylori NQ4044

- A. HPNQ4044_0981
- B. HPNQ4044_0982
- C. HPNQ4044_0983

Helicobacter pylori GAM103Bi

- A. HMPREF1393_00149
- B. HMPREF1393_00150
- C. HMPREF1393_00151

Helicobacter pylori CPY6271

- A. HPCPY6271_0257
- B. HPCPY6271_0258
- C. HPCPY6271_0259

Helicobacter pylori GAM249T

- A. HMPREF1410_00998
- B. HMPREF1410_00999
- C. HMPREF1410_01000

Helicobacter pylori NQ4076

- A. HPNQ4076_0812
- B. HPNQ4076_0813
- C. HPNQ4076_0814

Helicobacter pylori (strain F16)

- A. HPF16_0630
- B. HPF16_0631
- C. HPF16_0632

Helicobacter pylori Hp H-1

- A. HPHPH1_0785
- B. HPHPH1_0786
- C. HPHPH1_0787

Helicobacter pylori GAM264Ai

- A. HMPREF1420_00525
- B. HMPREF1420_00526
- C. HMPREF1420_00527

Helicobacter pylori R046Wa

- A. OUO_0704
- B. OUO_0705
- C. OUO_0706

Helicobacter pylori (strain P12)

- A. HPP12_0740
- B. HPP12_0741
- C. HPP12_0742

Helicobacter pylori Hp H-43

- A. HPHPH43_0770
- B. HPHPH43_0771
- C. HPHPH43_0772

Helicobacter pylori GAM244Ai

- A. HMPREF1407_00104
- B. HMPREF1407_00105
- C. HMPREF1407_00106

Helicobacter pylori OK310

- A. HPOK310_0614
- B. HPOK310_0615
- C. HPOK310_0616

Helicobacter pylori GAM254Ai

- A. HMPREF1415_01057
- B. HMPREF1415_01058
- C. HMPREF1415_01059

Helicobacter pylori CPY6311

- A. HPCPY6311_0758
- B. HPCPY6311_0759
- C. HPCPY6311_0760

Helicobacter pylori Hp P-11b

- A. HPHPP11B_0948
- B. HPHPP11B_0949
- C. HPHPP11B_0950

Helicobacter pylori Hp P-11

- A. HPHPP11_1074
- B. HPHPP11_1075
- C. HPHPP11_1076

Helicobacter pylori Hp H-10

- A. HPHPH10_0935
- B. HPHPH10_0936
- C. HPHPH10_0937

Helicobacter pylori GAM93Bi

- A. HMPREF1429_00366
- B. HMPREF1429_00367
- C. HMPREF1429_00368

Helicobacter pylori GAMchJs114i

- A. HMPREF1432_01081
- B. HMPREF1432_01082
- C. HMPREF1432_01083

Helicobacter pylori R037c

- A. OUK_1001
- B. OUK_1002
- C. OUK_1003

Helicobacter pylori BM012A

- A. U063_1035
- B. U063_1036
- C. U063_1037

Helicobacter pylori BM012S

- A. U064_1039
- B. U064_1040
- C. U064_1041

Helicobacter pylori CPY1313

- A. HPCPY1313_0376
- B. HPCPY1313_0377
- C. HPCPY1313_0378

Helicobacter pylori GAM115Ai

- A. HMPREF1397_00439
- B. HMPREF1397_00440
- C. HMPREF1397_00441

Helicobacter pylori Hp H-30

- A. HPHPH30_0952
- B. HPHPH30_0953
- C. HPHPH30_0954

Helicobacter pylori SouthAfrica50

- A. HPSA50_1093
- B. HPSA50_1094
- C. HPSA50_1095

Helicobacter pylori GAM114Ai

- A. HMPREF1396_00345
- B. HMPREF1396_00346
- C. HMPREF1396_00347

Helicobacter pylori NQ4053

- A. HPNQ4053_0693
- B. HPNQ4053_0694
- C. HPNQ4053_0695

Helicobacter pylori Hp A-17

- A. HPHPA17_0823
- B. HPHPA17_0824
- C. HPHPA17_0825

Helicobacter pylori Hp P-25c

- A. HPHPP25C_0968
- B. HPHPP25C_0969
- C. HPHPP25C_0970

Helicobacter pylori Hp P-25

- A. HPHPP25_1094
- B. HPHPP25_1095
- C. HPHPP25_1096

Helicobacter pylori Hp P-25d

- A. HPHPP25D_1156
- B. HPHPP25D_1157
- C. HPHPP25D_1158

Helicobacter pylori GAM121Aii

- A. HMPREF1402_01489
- B. HMPREF1402_01490
- C. HMPREF1402_01491

Helicobacter pylori (strain B8)

- A. HPB8_938
- B. HPB8_939
- C. HPB8_940

Helicobacter pylori GAM260Bi

- A. HMPREF1417_00971
- B. HMPREF1417_00972
- C. HMPREF1417_00973

Helicobacter pylori HP260Bi

- A. HMPREF1452_00708
- B. HMPREF1452_00709
- C. HMPREF1452_00710

Helicobacter pylori GAM120Ai

- A. HMPREF1401_00783
- B. HMPREF1401_00784
- C. HMPREF1401_00785

Helicobacter pylori GAM105Ai

- A. HMPREF1394_01272
- B. HMPREF1394_01273
- C. HMPREF1394_01274

Helicobacter pylori Hp H-9

- A. HPHPH9_0805
- B. HPHPH9_0806
- C. HPHPH9_0807

Helicobacter pylori Hp P-28b

- A. HPHPP28B_0798
- B. HPHPP28B_0799
- C. HPHPP28B_0800

Helicobacter pylori R038b

- A. OUM_0882
- B. OUM_0883
- C. OUM_0884

Helicobacter pylori GAM117Ai

- A. HMPREF1398_01532
- B. HMPREF1398_01533
- C. HMPREF1398_01534

Helicobacter pylori Hp H-42

- A. HPHPH42_0857
- B. HPHPH42_0858
- C. HPHPH42_0859

Helicobacter pylori NQ4099

- A. HPNQ4099_0984
- B. HPNQ4099_0985
- C. HPNQ4099_098

Helicobacter pylori SouthAfrica20

- A. HPSA20_0770
- B. HPSA20_0771
- C. HPSA20_0772

Helicobacter pylori HP260AFi

- A. HMPREF1448_00453
- B. HMPREF1448_00454
- C. HMPREF1448_00455

Helicobacter pylori HP260AFii

- A. HMPREF1449_00848
- B. HMPREF1449_00849
- C. HMPREF1449_00850

Helicobacter pylori GAM268Bii

- A. HMPREF1422_01430
- B. HMPREF1422_01431
- C. HMPREF1422_01432

Helicobacter pylori HP260ASii

- A. HMPREF1450_00962
- B. HMPREF1450_00963
- C. HMPREF1450_00964

Helicobacter pylori GAM260ASI

- A. HMPREF1416_00640
- B. HMPREF1416_00641
- C. HMPREF1416_00642

Helicobacter pylori GAMchJs106B

- A. HMPREF1431_00972
- B. HMPREF1431_00973
- C. HMPREF1431_00974

Helicobacter pylori GAM83T

- A. HMPREF1428_00817
- B. HMPREF1428_00818
- C. HMPREF1428_00819

Helicobacter pylori GAM83Bi

- A. HMPREF1427_01158
- B. HMPREF1427_01159
- C. HMPREF1427_01160

Helicobacter pylori GAM101Biv

- A. HMPREF1392_00657
- B. HMPREF1392_00658
- C. HMPREF1392_00659

Helicobacter pylori Hp H-29

- A. HPHPH29_0819
- B. HPHPH29_0820
- C. HPHPH29_0821

Helicobacter pylori G27

- A. HPG27_687
- B. HPG27_688
- C. HPG27_689

Helicobacter pylori GAM263BFi

- A. HMPREF1419_00369
- B. HMPREF1419_00370
- C. HMPREF1419_00371

Helicobacter pylori GAM100Ai

- A. HMPREF1391_00668
- B. HMPREF1391_00669
- C. HMPREF1391_00670

Helicobacter pylori Hp P-8

- A. HPHPP8_0912
- B. HPHPP8_0913
- C. HPHPP8_0914

Helicobacter pylori Hp P-8b

- A. HPHPP8B_0806
- B. HPHPP8B_0807
- C. HPHPP8B_0808

Helicobacter pylori Hp H-41

- A. HPHPH41_0809
- B. HPHPH41_0810
- C. HPHPH41_0811

Helicobacter pylori Hp H-45

- A. HPHPH45_0792
- B. HPHPH45_0793
- C. HPHPH45_0794

Helicobacter pylori CPY3281

- A. HPCPY3281_0892
- B. HPCPY3281_0893
- C. HPCPY3281_0894

Helicobacter pylori CPY1124

- A. HPCPY1124_0894
- B. HPCPY1124_0895
- C. HPCPY1124_0896

Helicobacter pylori NQ4216

- A. HPNQ4216_1077
- B. HPNQ4216_1078
- C. HPNQ4216_1079

Helicobacter pylori GAM245Ai

- A. HMPREF1408_01330
- B. HMPREF1408_01331
- C. HMPREF1408_01332

Helicobacter pylori HP250ASii

- A. HMPREF1442_01181
- B. HMPREF1442_01182
- C. HMPREF1442_01183

Helicobacter pylori GAM252T

- A. HMPREF1414_01354
- B. HMPREF1414_01355
- C. HMPREF1414_01356

Helicobacter pylori HP250BFii

- A. HMPREF1444_01326
- B. HMPREF1444_01327
- C. HMPREF1444_01328

Helicobacter pylori GAM250AFi

- A. HMPREF1411_01272
- B. HMPREF1411_01273
- C. HMPREF1411_01274

Helicobacter pylori HP250BFi

- A. HMPREF1443_01529
- B. HMPREF1443_01530
- C. HMPREF1443_01531

Helicobacter pylori HP250ASi

- A. HMPREF1441_01126
- B. HMPREF1441_01127
- C. HMPREF1441_01128

Helicobacter pylori HP250BFiii

- A. HMPREF1445_00884
- B. HMPREF1445_00885
- C. HMPREF1445_00886

Helicobacter pylori GAM252Bi

- A. HMPREF1413_00160
- B. HMPREF1413_00161
- C. HMPREF1413_00162

Helicobacter pylori HP250BSi

- A. HMPREF1447_00116
- B. HMPREF1447_00117
- C. HMPREF1447_00118

Helicobacter pylori HP250BFiV

- A. HMPREF1446_01180
- B. HMPREF1446_01181
- C. HMPREF1446_01182

Helicobacter pylori HP250AFiV

- A. HMPREF1440_01195
- B. HMPREF1440_01196
- C. HMPREF1440_01197

Helicobacter pylori HP250AFii

- A. HMPREF1438_01413
- B. HMPREF1438_01414
- C. HMPREF1438_01415

Helicobacter pylori HP250AFiii

- A. HMPREF1439_01373
- B. HMPREF1439_01374
- C. HMPREF1439_01375

Helicobacter pylori GAM250T

- A. HMPREF1412_00155
- B. HMPREF1412_00156
- C. HMPREF1412_00157

Helicobacter pylori CPY6081

- A. HPCPY6081_0707
- B. HPCPY6081_0708
- C. HPCPY6081_0709

Helicobacter pylori Hp A-16

- A. HPHPA16_0927
- B. HPHPA16_0928
- C. HPHPA16_0929

Helicobacter pylori GAM118Bi

- A. HMPREF1399_00375
- B. HMPREF1399_00376
- C. HMPREF1399_00377

Helicobacter pylori Hp P-30

- A. HPHPP30_1352
- B. HPHPP30_1353
- C. HPHPP30_1354

Helicobacter pylori Hp P-16

- A. HPHPP16_0887
- B. HPHPP16_0888
- C. HPHPP16_0889

Helicobacter pylori Hp H-23

- A. HPHPH23_0703
- B. HPHPH23_0704
- C. HPHPH23_0705

Helicobacter pylori CPY6261

- A. HPCPY6261_0948
- B. HPCPY6261_0949
- C. HPCPY6261_0950

Helicobacter pylori Hp H-28

- A. HPHPH28_0855
- B. HPHPH28_0856
- C. HPHPH28_0857

Helicobacter pylori OK113

- A. HPOK113_0739
- B. HPOK113_0740
- C. HPOK113_0741

Helicobacter pylori Hp H-5b

- A. HPHPH5B_1012
- B. HPHPH5B_1013
- C. HPHPH5B_1014

Helicobacter pylori 83

- A. HMPREF0462_0677
- B. HMPREF0462_0678
- C. HMPREF0462_0679

Helicobacter pylori Hp H-19

- A. HPHPH19_1074
- B. HPHPH19_1075
- C. HPHPH19_1076

Helicobacter pylori (strain F30)

- A. HPF30_0602
- B. HPF30_0603
- C. HPF30_0604

Helicobacter pylori R036d

- A. OUI_0847
- B. OUI_0848
- C. OUI_0849

Helicobacter pylori HP260BFii

- A. HMPREF1451_00537
- B. HMPREF1451_00538
- C. HMPREF1451_00539

Helicobacter pylori GAM260BSi

- A. HMPREF1418_00379
- B. HMPREF1418_00380
- C. HMPREF1418_00381

Helicobacter pylori Hp P-62

- A. HPHPP62_0833
- B. HPHPP62_0834
- C. HPHPP62_0835

Helicobacter pylori (strain 52)

- A. HPKB_0615
- B. HPKB_0616
- C. HPKB_0617

Helicobacter pylori Hp A-9

- A. HPHPA9_0577
- B. HPHPA9_0578
- C. HPHPA9_0579

Helicobacter pylori Hp P-3b

- A. HPHPP3B_0833
- B. HPHPP3B_0834
- C. HPHPP3B_0835

Helicobacter pylori (strain 908)

- A. hp908_0740
- B. hp908_0741
- C. hp908_0742

Helicobacter pylori CPY1662

- A. HPCPY1662_0220
- B. HPCPY1662_0221
- C. HPCPY1662_0222

Helicobacter pylori (strain 51)

- A. KHP_0593
- B. KHP_0594
- C. KHP_0595

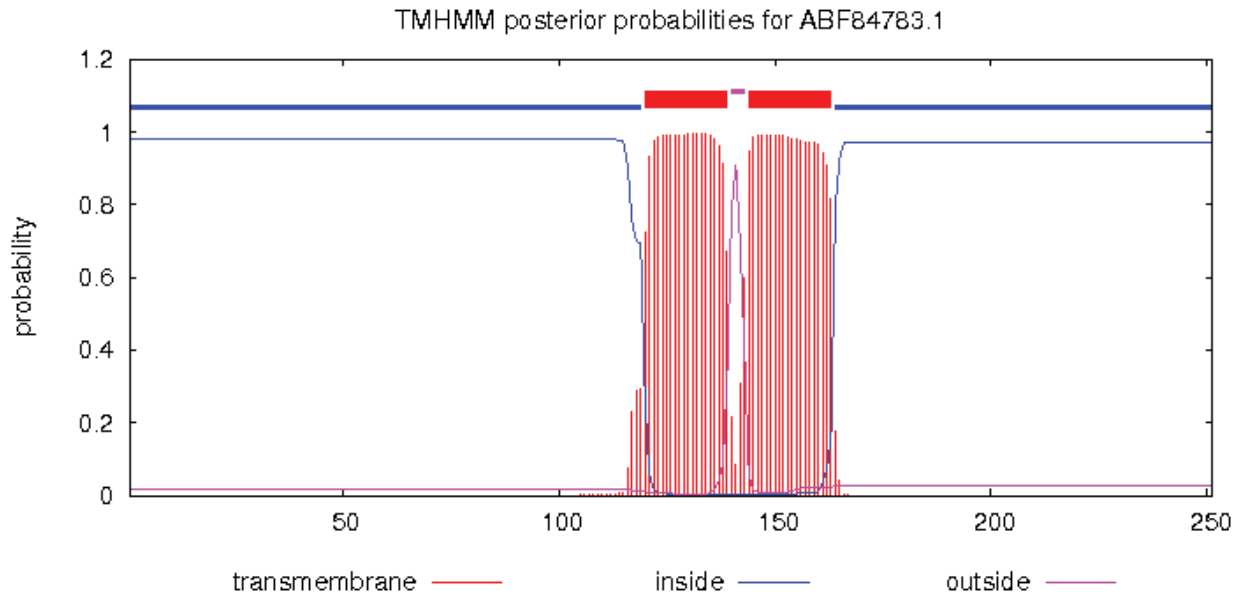
The Characterisation of Dynamin-Like Proteins in the Gastric Pathogen *Helicobacter pylori*

Emma Mary Dawson

S3: Transmembrane predictions of selected *H. pylori* Dlps

Transmembrane predictions carried out on a select 10 *H. pylori* strains (names above corresponding predication) found to contain Dlp1a, Dlp1b and Dlp2 are found below. The protein sequence of Dlp1b was used as the query sequence in the online server, TMHMM. If no transmembrane was detected, Dlp1a was used. It should be noted that the predicted transmembrane is generally 100 amino acids or so from the end of each protein.

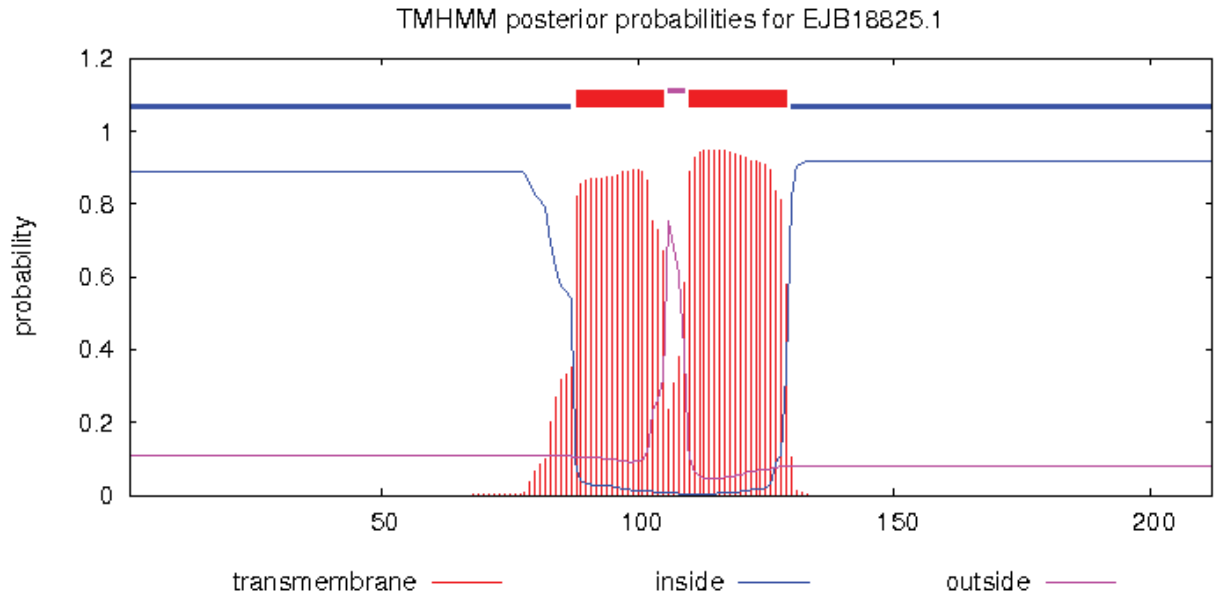
HPAG1



ABF84783.1 Length: 251
ABF84783.1 Number of predicted TMHs: 2
ABF84783.1 Exp number of AAs in TMHs: 40.75362
ABF84783.1 Exp number, first 60 AAs: 0
ABF84783.1 Total prob of N-in: 0.98281

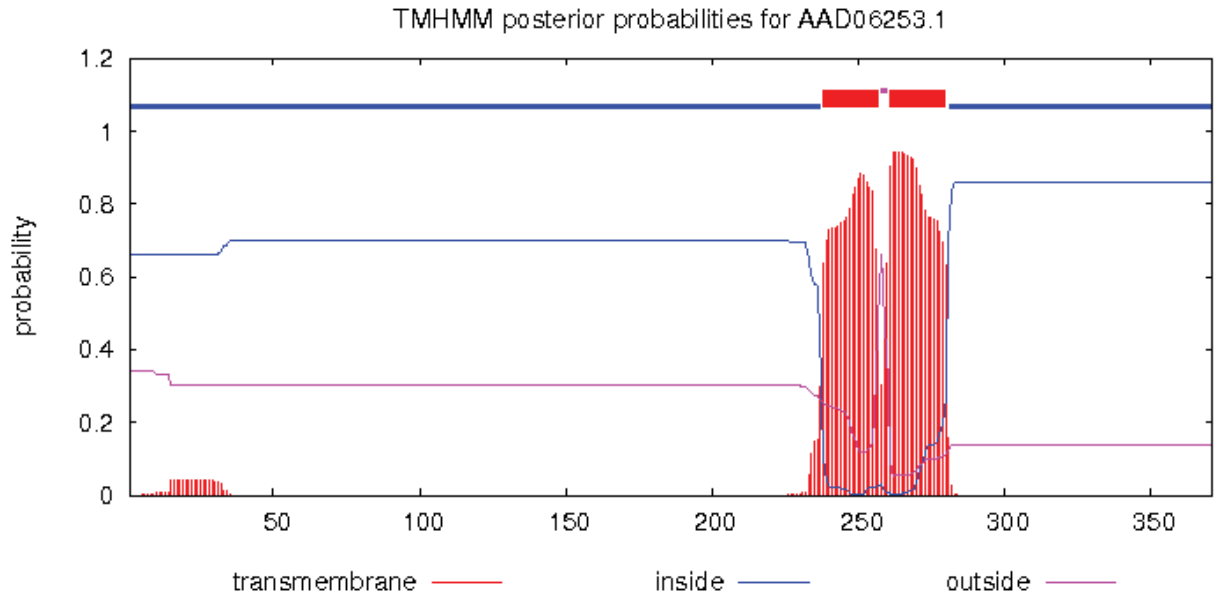
ABF84783.1	TMHMM2.0	inside	1	119
ABF84783.1	TMHMM2.0	TMhelix	120	139
ABF84783.1	TMHMM2.0	outside	140	143
ABF84783.1	TMHMM2.0	TMhelix	144	163
ABF84783.1	TMHMM2.0	inside	164	251

CPY6081



EJB18825.1 Length: 212
EJB18825.1 Number of predicted TMHs: 2
EJB18825.1 Exp number of AAs in TMHs: 36.6646
EJB18825.1 Exp number, first 60 AAs: 0
EJB18825.1 Total prob of N-in: 0.89032

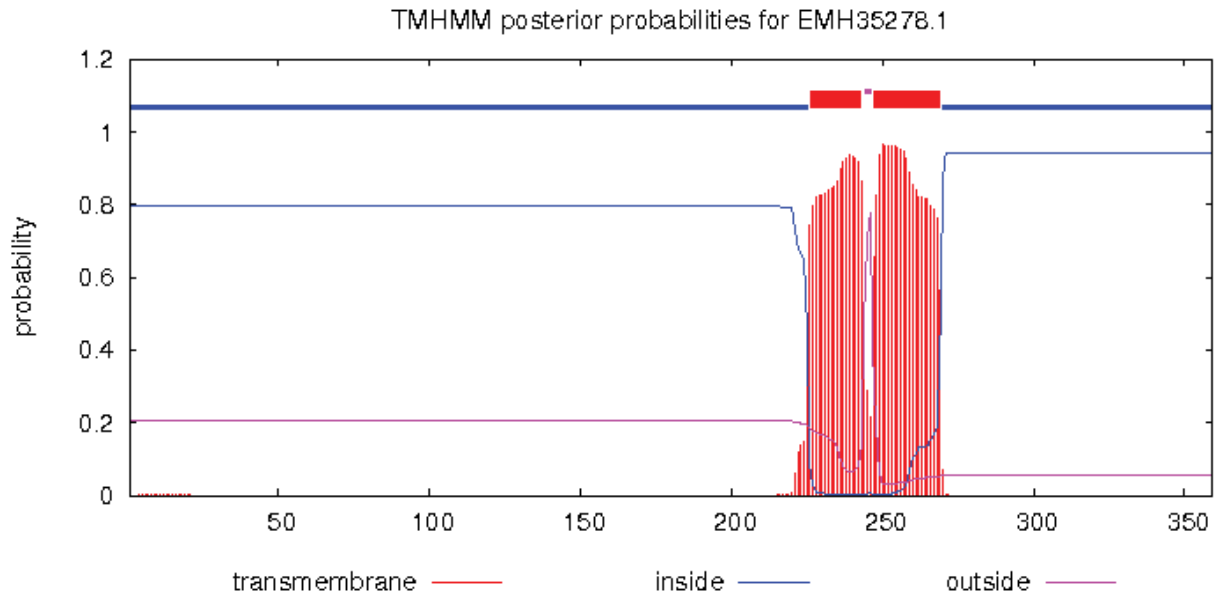
EJB18825.1	TMHMM2.0	inside	1	87
EJB18825.1	TMHMM2.0	TMhelix	88	105
EJB18825.1	TMHMM2.0	outside	106	109
EJB18825.1	TMHMM2.0	TMhelix	110	129
EJB18825.1	TMHMM2.0	inside	130	212



AAD06253.1 Length: 371
AAD06253.1 Number of predicted TMHs: 2
AAD06253.1 Exp number of AAs in TMHs: 35.3251
AAD06253.1 Exp number, first 60 AAs: 0.76967
AAD06253.1 Total prob of N-in: 0.66043

AAD06253.1	TMHMM2.0	inside	1	237
AAD06253.1	TMHMM2.0	TMhelix	238	257
AAD06253.1	TMHMM2.0	outside	258	260
AAD06253.1	TMHMM2.0	TMhelix	261	280
AAD06253.1	TMHMM2.0	inside	281	371

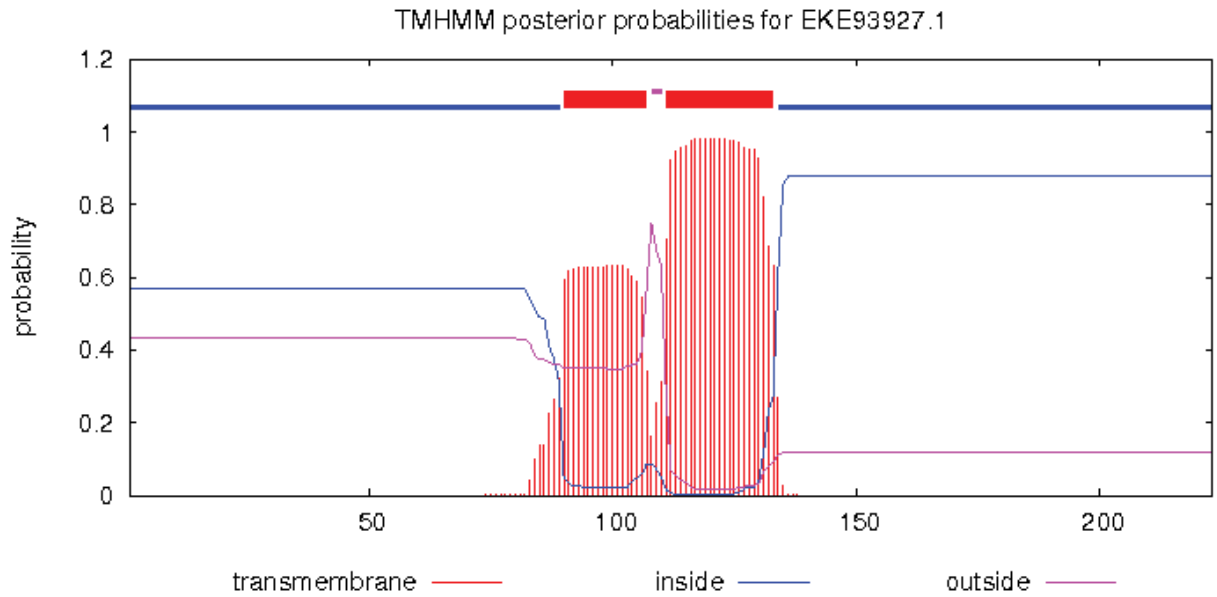
GAM80Ai



EMH35278.1 Length: 359
EMH35278.1 Number of predicted TMHs: 2
EMH35278.1 Exp number of AAs in TMHs: 37.22327
EMH35278.1 Exp number, first 60 AAs: 0.00778
EMH35278.1 Total prob of N-in: 0.79528

EMH35278.1	TMHMM2.0	inside	1	225
EMH35278.1	TMHMM2.0	TMhelix	226	243
EMH35278.1	TMHMM2.0	outside	244	246
EMH35278.1	TMHMM2.0	TMhelix	247	269
EMH35278.1	TMHMM2.0	inside	270	359

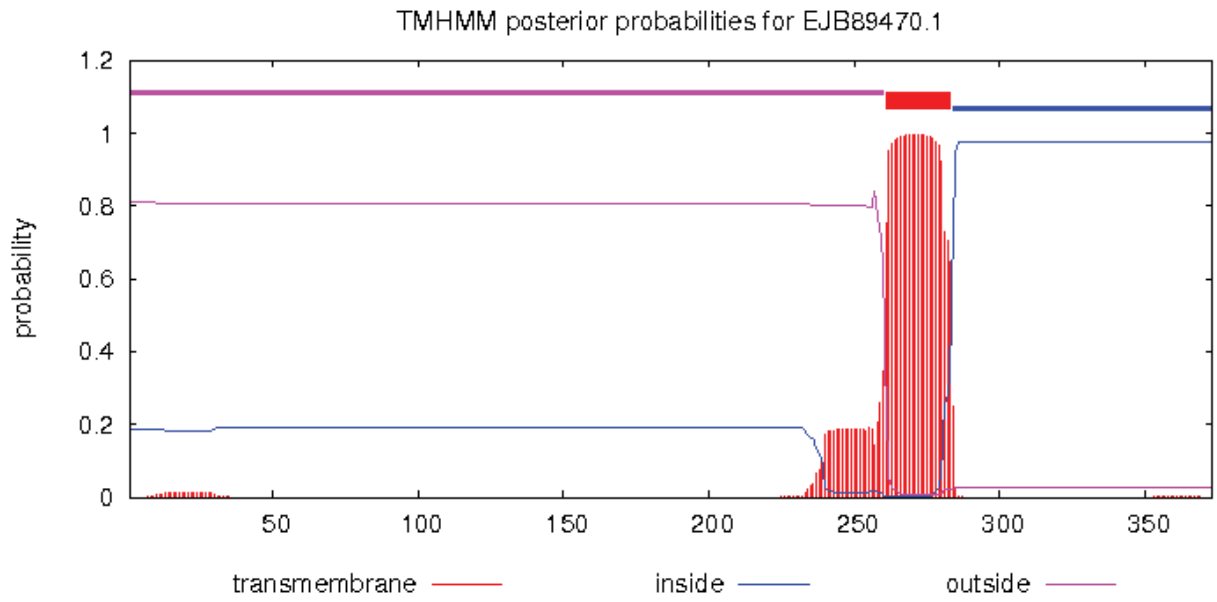
R055a



EKE93927.1 Length: 223
EKE93927.1 Number of predicted TMHs: 2
EKE93927.1 Exp number of AAs in TMHs: 34.2831
EKE93927.1 Exp number, first 60 AAs: 4e-05
EKE93927.1 Total prob of N-in: 0.56840

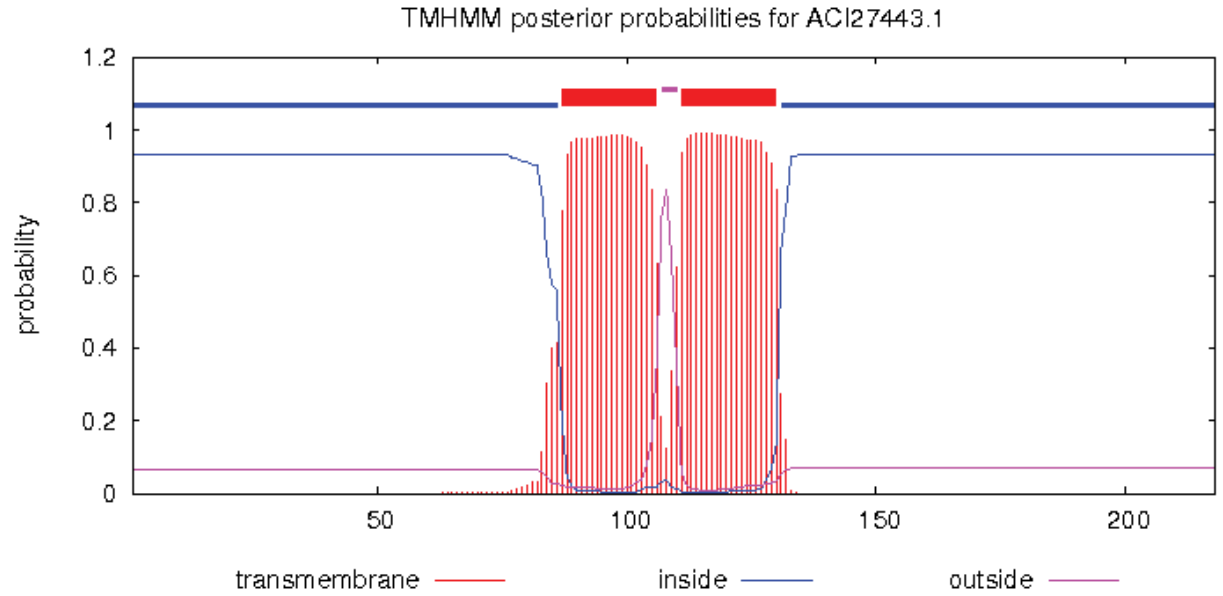
EKE93927.1	TMHMM2.0	inside	1	89
EKE93927.1	TMHMM2.0	TMhelix	90	107
EKE93927.1	TMHMM2.0	outside	108	110
EKE93927.1	TMHMM2.0	TMhelix	111	133
EKE93927.1	TMHMM2.0	inside	134	223

HP H-18



EJB89470.1 Length: 373
EJB89470.1 Number of predicted TMHs: 1
EJB89470.1 Exp number of AAs in TMHs: 26.43746
EJB89470.1 Exp number, first 60 AAs: 0.24032
EJB89470.1 Total prob of N-in: 0.18730

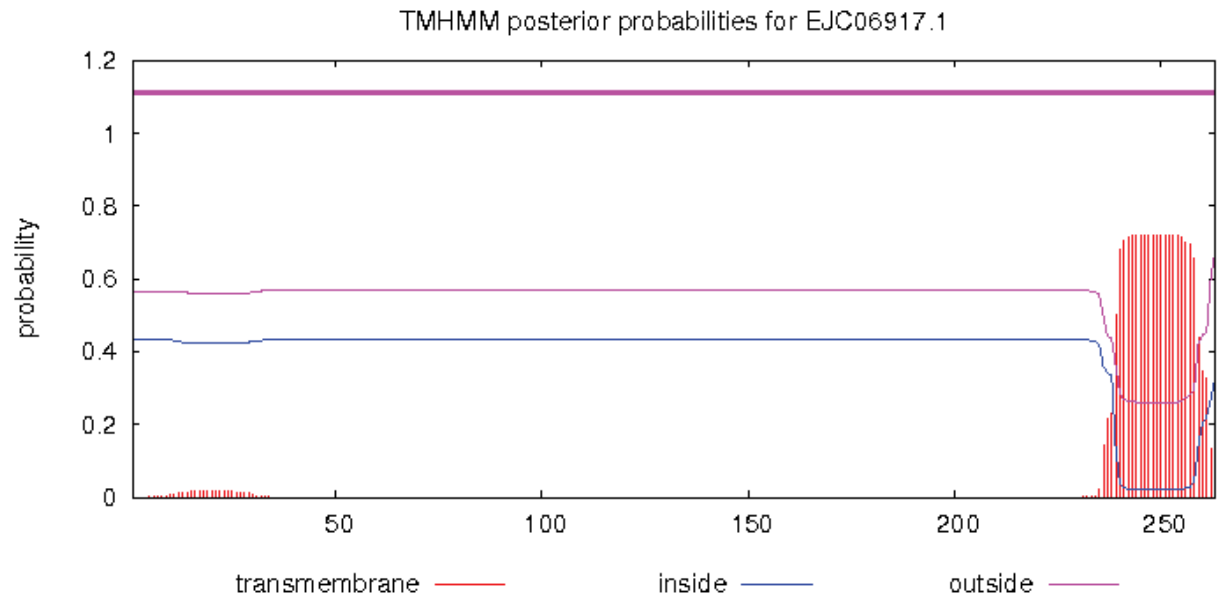
EJB89470.1	TMHMM2.0	outside	1	260
EJB89470.1	TMHMM2.0	TMhelix	261	283
EJB89470.1	TMHMM2.0	inside	284	373



ACI27443.1 Length: 218
ACI27443.1 Number of predicted TMHs: 2
ACI27443.1 Exp number of AAs in TMHs: 41.12137
ACI27443.1 Exp number, first 60 AAs: 2e-05
ACI27443.1 Total prob of N-in: 0.93212

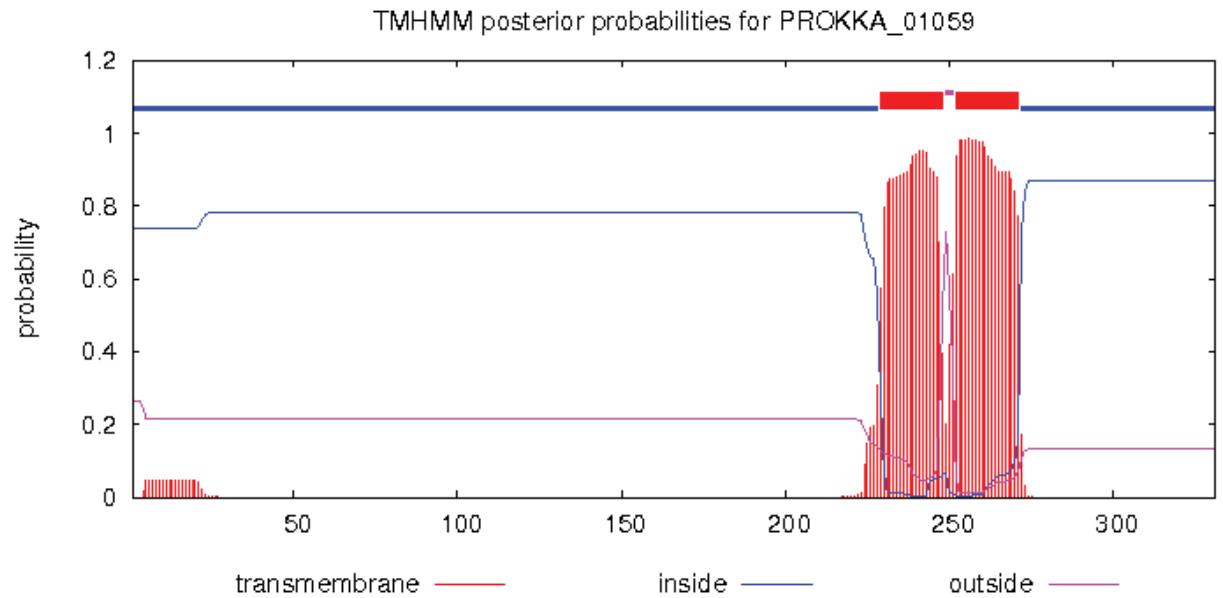
ACI27443.1	TMHMM2.0	inside	1	86
ACI27443.1	TMHMM2.0	TMhelix	87	106
ACI27443.1	TMHMM2.0	outside	107	110
ACI27443.1	TMHMM2.0	TMhelix	111	130
ACI27443.1	TMHMM2.0	inside	131	218

HP P-13



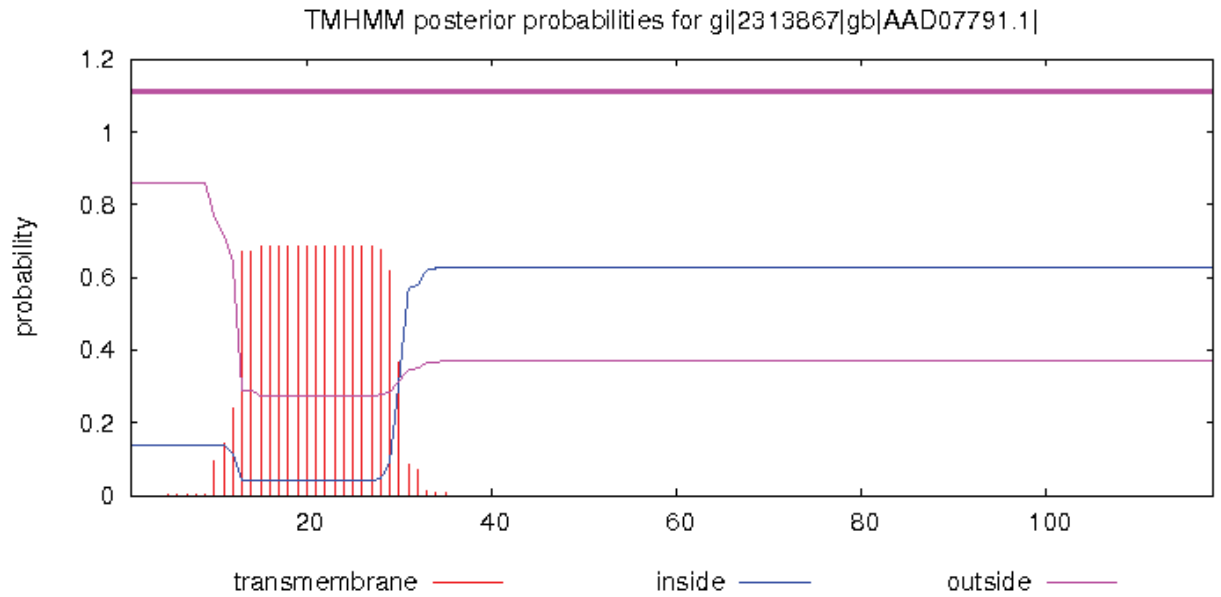
EJC06917.1 Length: 263
EJC06917.1 Number of predicted TMHs: 0
EJC06917.1 Exp number of AAs in TMHs: 16.15156
EJC06917.1 Exp number, first 60 AAs: 0.28648
EJC06917.1 Total prob of N-in: 0.43399
EJC06917.1 TMHMM2.0 outside 1 263

B128.7.13



PROKKA_01059 Length: 331
PROKKA_01059 Number of predicted TMHs: 2
PROKKA_01059 Exp number of AAs in TMHs: 38.77751
PROKKA_01059 Exp number, first 60 AAs: 0.84361
PROKKA_01059 Total prob of N-in: 0.73785
PROKKA_01059 TMHMM2.0 inside 1 228
PROKKA_01059 TMHMM2.0 TMhelix 229 248
PROKKA_01059 TMHMM2.0 outside 249 251
PROKKA_01059 TMHMM2.0 TMhelix 252 271
PROKKA_01059 TMHMM2.0 inside 272 331

26695



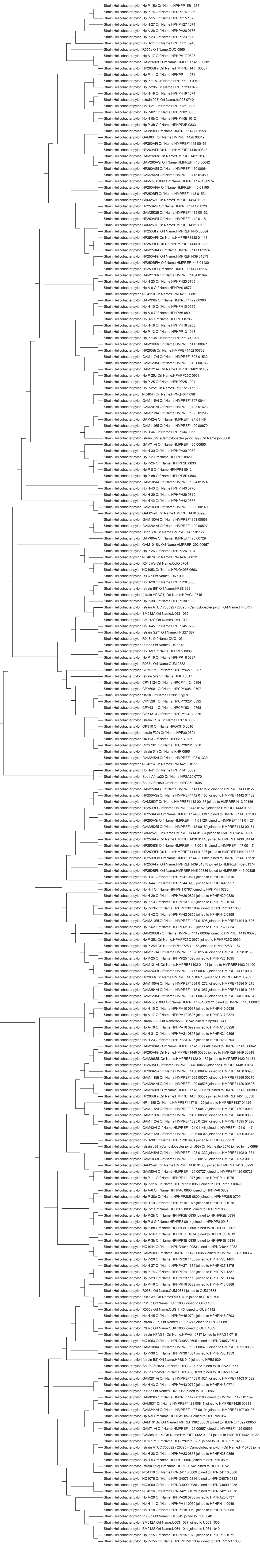
gi|2313867|gb|AAD07791.1| Length: 118
gi|2313867|gb|AAD07791.1| Number of predicted TMHs: 0
gi|2313867|gb|AAD07791.1| Exp number of AAs in TMHs: 12.57369
gi|2313867|gb|AAD07791.1| Exp number, first 60 AAs: 12.57283
gi|2313867|gb|AAD07791.1| Total prob of N-in: 0.13933
gi|2313867|gb|AAD07791.1| POSSIBLE N-term signal sequence
gi|2313867|gb|AAD07791.1| TMHMM2.0 outside 1 118

**The Characterisation of Dynamin-Like Proteins in the Gastric Pathogen
*Helicobacter pylori***

Emma Mary Dawson

**S4: Completely Labelled *H. pylori* dynamin (back-to-back)
phylogenetic tree**

Below is the completely labelled phylogenetic tree of Dlp1 and Dlp2 homologs from all *H. pylori* strains found to have them consecutive in their genome.



The Characterisation of Dynamin-Like Proteins in the Gastric Pathogen *Helicobacter pylori*

Emma Mary Dawson

S5: *H. pylori* B128.7.13 predicted proteome annotated by PROKKA

Below are the predicted ORFs and their corresponding annotation for the mouse adapted *H. pylori* strain B128.7.13 generated from PROKKA. This predicted proteome was used preliminary in the assignment of peptides during the SWATH-MS analysis.

```
>PROKKA_00001 hypothetical protein
MSVGYQIGGTQQFINNKQLENQNIINSVTQSAINVAGPTTGLITLSSQSVIDALGYGV
NTVGNQLEGISNILNQIGKEKTFILAVKSLAFPSKS
>PROKKA_00002 hypothetical protein
MAELLQQIEHSITKTTSTTYAQSLLSNLTDVNASSNNTAYVSALVNALNTLGVGVFPTT
TSTHXXXPTRTNRILSS
>PROKKA_00003 hypothetical protein
MEQFIQNLSPLTATPTSNNQANQQVQAIQKQLQSVAINTLDNNAINNNTTYNLNNLHNA
LNFQAYESTIEQYNNALKQISWISFTEPKNLLKNTSNNYQIGTVTNAQQQONISAYDCT
SATGSLSSNASSGISCSATSSTSTNSFDNSLVATSKVQTINGKEQIGVNSFNLVSQV
SVYNSLKTSEENLQKNANILCTNGSQSGTSPCNSSSGGLSIGNAQLQNILSSTNGT
TTTTQAKNASKKAMVVVNNEEEAKTTNLAQNSGPTTQSSNSTVMGALNTVLQNI
NINFQQSIQSAFQNZENNIQAWANAIYNTNGNQSQDMTLNTNQDLRIQLRANFYQ
LINTINQQAPTDM SALIAQSQTQQTSGATSNNNACASGMNGSSSNWCYQQWSD
SKAYYSGLQ SALGIKHKQQLKAGAVGTASPTMSNKSRS LAVGYLITSFHLLRVL
VIAEVMGEAVEAQMALAKSTQPTRCSQTPVMGN
>PROKKA_00004 Helicobacter outer membrane protein
MLTDASDGKLGTYSSNSGSGNNGYTPCNSGANGNSSNCYTPSKTQNTTTATTTT
DSNLQKVYNDAQKIANIIASSGNNKGVENGLKQFFKALNNSSSLNLCNGSSNC
SGGLINLLGAIPIGVSDTNNLVNLLTEFIKTAGFIQNNDSSVSTSLTNAFQAIT
NAISQGFQALQNDISPNAILTLLQEITSNTTTIQSFSQTLRQLLGDKTF FMV
QQKLIDAMINARNQVQNAQNQANNYGSQPVLSQYAAAKSTQHGM SNGLGVSL
GYKYFFGKARKLGLRHYFFFDYGFSEIGLANQSVKANIFAYGVGTDFLWNLF
RRTYNTKALNFGLFAGVQLGGATWLSLRQQIIDNWGSANDIHSTNFGVALNFG
VRTNFAEFKRFAKKFHNQGVISQKSVEFGIKVPLINQAYLNSAGADVNYRR
LYTFYINYIMGF
>PROKKA_00005 Pyrroline-5-carboxylate reductase
MEILQFIGYGNMAQAILEGAHEILSKRFILEITGRNPEKIAPFLQEKNIQAQIV
PYKDAIDIHQKFVFLLFKPYNLKDFNYQGQAKSVLSALAGVGFKALSDAINSL
HYLKMPNIA SKFALSSTAVCEKSPMPLISQKALSIIESFGNCVRVGNEEQVD
SSVATNGSALAFSLVASSLKDAGIREGLNARDSLELVKMSFKGFAKLLERPE
MIIIEQICTPKGATIEGLSVLEKKGVRGAFIEACQKSVKKMRL
>PROKKA_00006 Adenosine monophosphate-protein transferase
NmFicMHLDRQSLEKAKHLIQSGLIDTIEVGTIKGLQEIHRLFEGLYEFAGKIR
DKNISNFRFANCLYLDLILPRIESMPQNNFNQIVEKYVEMNIAHPFLKG
>PROKKA_00007 Adenosine monophosphate-protein transferase
NmFicMLKKELKKIVLWDRIDKAAYLSAIKRSPVNDLEIKTLLKKHLSSNTND
PLTLIKGITQSYYYEGL
```

>PROKKA_00008 Endoribonuclease YbeY
MLEIDNQTPLESDFLLLEKIADVLAAPTQIIELVLVSDETMREINRDLRNC DYATDVL SFP
LEAIPHTPLGSGVINAPLAQENALKLGHSLENEIALLFIHGVLHLLGYDHEKDKGEQRQK
ESELIEAFNLPLSLIERAQD
>PROKKA_00009 Flavodoxin
MGKIGIFFGTDSGNAEAI AEKISKAIGNAEVVDVAKASKEQFNSFTKVILVAPTAGAGDL
QTDWEDFLGTLESSDFANKTIGLVGLGDQD TYSETFAEGIFHIYEKAKAGKVVGHTSTDG
YHFEASKAVEGGKFVGLVIDEDNQDDLTD ERISKWVEQVKSSFA
>PROKKA_00010 Inner membrane protein YohD
MQEALLRFQEGFKEWGYLILFLYSLGGGYVGIVIASILSATTHALDIKITILVAFLGNMV
GSGALVIFARYQKRDFLQYFQKHRRKLALASLWVKRYAFLMIFINKYLYGIKSVVPLAIG
FSKYPLKRFLWLNVLSSFLWALIVGSVSFQASDWVKTLYERLSHYTSFFVISFVLIALLI
WFLLKRYSRKMGF
>PROKKA_00011 Cytochrome oxidase maturation protein cbb3-type
MNTEILTIMLVVSVLMGLVGLIAFLWG VKSGQFDDEKRMLESVLYDSASDLNEAILQEKR
QDPKN
>PROKKA_00012 Ferredoxin--NADP reductase
MDQEILDVLIVGAGPGGIATAVECEIAGVKVLLCEKTESHSGMLEKFYKAGKRIDKDYK
KQVVELKGHIPFKDSFKEETLENFTNLLKEYRITPSYKTDIESVKKEGELFKITTTSTNT
YHAKFVVVAIGKMGQPNRPTYKIPVALSKQV VFSINDCKENEKTLVIGGGNSAVEYAIAL
CKTTPPTTLNRYKKEFSRINEDNAKNLQEV LNNNTLKS KLGV DIESLEEDNTQIKVNFTDN
TSESFDRLLYAIGGSTPLEFFKRC SLELDPSTNIPVVKENLESNNIPNLFIVGDILFKSG
ASIATALNHGYDVAIEIAKRLRS
>PROKKA_00013 Major Facilitator Superfamily protein
MLRKNILAYYGANFLLIIAQSLPHAILTPLL LSKGLSLSEILLVQTFFSFCVLVAEYPSG
VLADLMSRKNLFLVSNVFLIASFSFVLFFDSFIFMLLAWGLYGLYSACSSGTIEASLITD
IKENKKDL SRFLAKNNQITYLGMII GSSLGSFLYLKVHAMLYIVGIFLIMLCMLTIVIYF
KEKEGDFKSQKSLKLLKEQVKGSLKELKDNPKLKILLVGHLITPIFFMSHFQMWQAYFLK
QGIKEQYLFMFYIAFQVISILIHFLKAKNYSQKIALSSLLVLLGVSPLLLSNIPYCFIGV
YALMVAFFAYMSYCLGYQFSKFVSKNNIS SLLSSCVRVVSVLILSLSSLELRYFSPL
TIITMHFALTIIILFFFLYKAKPFDE
>PROKKA_00014 Glucose-6-phosphate isomerase
MLTQLKTYPKLLKH YEEIKEMHMRDWF SKDKERASRYFVQLESLSLDYSKNRLNDTTLKL
LFELAKDCSLKEKIEAMFKGEKINTTEKRAVLHTALRSLNDAEILLDNMEVLKS VRSVLK
RMRAFSDSVRS GKRLGYTNQVITDIVNIGIGGSDLGALMVCTALKRYAHPRLKMHFVSNV
DGTQILDVLEKINPASTL FIVASKTFSTQETLTNALTARKW FVERS GDEKHI AKHFVAVS
TNKEAVQQFGIDEHNMFEFWD FVGGRYSLWSAIGLSIMIYLGKKNFNALLKGAYLMDEHF
RNAPFESNLPVLMGLIGVWYINFFQSKSHLIAPYDQYLRHF PKFIQQLDMESNGKRISKK
GEIIPYDTCPVVWGD MGINAQHAFFQLLHQGTHLIPIDFIASLDKKPNAKGHHEILFSNV
LAQAQAFMKGKSYEEAFGELL SKGLDKDEAKDLAHRVFFGNRPSNILLLEKISPSNIGA
LVALYEHKVVFVQGV IWDINSFDQWGV ELGKELAVPILQELEGHKS NAYFDSSTKHLIELY
KNYNQ
>PROKKA_00015 Putative outer membrane protein
MKKASQVLF FGAFLSSSLQGFEAKLNGFVDQSSTIGFNQHKINKERGIYPMQQFATIAGY
LGLGFSLLPKKVSDHVLKGKIGGMVGSIFYDGTKKFEDG SVAYNLFGYYDGMGGYTNIL
QTD SLETQNMKHKNVRNYVFS DAYLEYAYKNYFEIKAGRYLSTMPYKSGQTQGFQVSGQ
YKKARLTWFSSFGRAFAYGSFLMDWFAARTTYSGGFTKNDKGGYDSHGRKVLYGTHAVQL
TYKPHRFLIEGFYYLSPQIFNAPGVKIGWDSNP NFSGTGFRSDTAVMGFFPIYYPWMIVK
SNGSPVYRYDTPATQNGQNLIIRQRFDINN NVSIAFYKV FQNGWIGNMGNPSGVIMG
SNSVYAGFTGTALKRDAATILLSCGGTHFAKKFTWK FATQYSNSVVSWEARAMISLGYKF
NEYLSGSVDL AYYGVHTNKGFKAGENGVPVKNFPALYSDRSALYTALVASF
>PROKKA_00016 Carbon starvation protein A
MQKSLVSLAWVFVAILGAICLGV LALHKGESINTLWL VVASACIYSIGYRFYSHFIA YRV
LKLD DS RATPACIRNDGRDFVPTDKAITFGHHFAA IAGAGPLVGPILAAQMGYLPSILWI
LIGSVLGGCVHDFVVL FASIRRDGKSLGEMIKLEMGGFVGMIASLGILGIMLII IAILAM
VVVKALAHSPWGFFT IAMTIPIAILMGLYMRFFRPHKILEVSVIGFILLIIAIYAGKYVS
LDPKLASIFTFDASSLAWMIMGYGFVASILPVWFL LAPRDYLS TFLKIGVIGVLVVAIVF
VAPPLQIPKITPFVDGSGPVFAGSVFPFLFITVACGTISGFHALISSGTPKMLAKESDA

RLVGYGSMVMESVVALMALVCAGILHPGLYFAINSPEVSIGKDIADAASVISSWGFSISA
EEISEMTKNIGESSILSRTGGAPTFAIGLAMIVYHILGDPSVMAFWYHFAILFEALFILT
AVDAGTRTARFMIQDLLGNVYKPLGNLSSYKAGIFATLLCVAGWGYFLYQGTIDPKGGIY
TLWPLFGVSNQMLAGMALLLVTVVLFKMGRFGAMVSALPAVLILFITFYSGILKVVPKS
NDSVLNNVSHVAQMQUIKEKIALTTDEKALKTLQKSFFNHAIIDAILCVFFMLVALLVLIV
SVRICSNAYFKNQIYPPLAETPYIKAA

>PROKKA_00017 Glutamine transport system permease protein GlnP
MALDWFDMFHSIPAFKLGLELTLYISFFGILLSLLVGFLCAIVLYFKMRFISPVVYVYGE
IARNTPLLIQLFFLYYGLNEIGLSALECAILALGFLGGGYMSQSFLLGFKSLASIQRESA
LSLGFSPKMMYYIILPQSLSVSMPSIGANVIFLLKETS SVVGAIALTDIMFVAKDLIGIY
YKTTESLLMLSLTLYLIALPLSVLFVILERSFKKKVA

>PROKKA_00018 glutamine ABC transporter permease protein
MGVLELDELNLKRLLEGFEITLLIALSSAVISIIVGTLLGSLMAFGSQIVVLACRVYLESI
RIIPLLAWLFIYVYFGLASWFDLHISAVLASVIVFSLWGVLK

>PROKKA_00019 Inner membrane amino-acid ABC transporter permease protein YhdY
MMDLTRGVLTSSVKHQVESALALGMDSSKKVIFNIIFPQSFLSLLPSSNLNLFTRMIKTAL
VSLIGAIDLKLVGQQIIELNLLRMPNASFVVYGVILMFYFSLCYSLSLYSSYLEKKFQHI
RG

>PROKKA_00020 Arginine transport ATP-binding protein ArtM
MSVILETKGLKKTYQNHLVLDGINFTLNKGEVAVILGPSGCGKSTFLKCLNGLEKIDGGE
ILFENTNLNNKTTNWNQMRQKIGMVFNQYELFPHLNVLDNILLAPLKVQKRSKDEVISQA
VELLKRVGLEHKQQAYPKELSGGQKQ RVAIVRSLCMRPKIMLFDEV TASLDPEMVKEVLE
VILELAT TGMSMIVTHEMKAQKIASKIVFFD SGKIAEENSAKEFFNHPKSQRAQKFLE
TFHFLGSC

>PROKKA_00021 ABC transporter glutamine-binding protein GlnH precursor
MKTNGLLKMWGLFLVLIALVFSACSDSHKEKKDALEVIKQRGVLKVGVFSDKPPFGSVDS
KGKYQGYDVVIXKRMALDGLDENKIEFIPVEASARVEFLKANKVDIIMANFTRTKEREK
VVDFAKPYMKVALGVISKDGVIKNIEELKD KELIVNKGTTADFYFTKNYPNIKLLKFEQN
TETFLALLNNKATALAHDNTLLLAWIKQHPEFKLGITSLGDKDVIAPA IKKGNPKLLEWL
NNEIDSLISSDFLKEAYQETLAPVYGDEIKPEEIIIFE

>PROKKA_00022 hypothetical protein
MSYFYKRCLKFSLIGLLGLLSVQLDARSFVDGDLDIQKFSYEDSLLKKGDPNGVHKVQVR
DYKGMQEAIEHSEIRIALKPGVKKEVKKGKIYSAQINDGMCYAFRMLQTGDNTTGLDSK
EFPKQSREKKGRVITLIGKDEVPLYILETDCQVGDIAKISLVGNFDGTGFLTEHKFKDAK
PIY

>PROKKA_00023 L-fucose-proton symporter
MQKTSNTLALGSLTALFFLMGFITVLNDILIPHLKPIFDLTYFEASLIQFCFFGAYFIMG
GVFGNVISKIGYPFGVVLGFVITATGCALFYPAAHFGSYGFFLGALFILASGIVCLQTAG
NPFVTLSSKGKEARNLVLVQAFNSLGTTLGPIFGSLLIFSTTKMGDNASLIDKLADAKSV
QMPYLGLAVFSLLLALIMYLLKLPDVEKEMPKE TTQKSLF SHKHVVF GALG IFFYVGGEV
AIGSFLVLSFEKLLNDPQSSAHYLVYYWGGAMVGRFLGSVLMNKIAPNKYLAFNALSSI
VLIALAIIIGGKIALFALT FVGFFNSIMFPTIFSLATLNLGHLTSKASGVISMAIVGGAL
IPPIQGA VTDMLTATESNLLYAYGVPLLCYFYILFFALKGYKQEENS

>PROKKA_00024 Putative permease YicO
MGFFKLKEHNTNIATEFRAGLTTFITMIYIVPLNALILSHANMPYEALLSATAIITILSS
VFNGLWANTPIAMSVGLGLSAYFSFGLVQGLKLPWQSALGIVALSGAIFVILSFTKFRSW
VMRSIPSDLRRRAVSAGIGAFIAFIGLKEMHIVVTHKATLVTLGDFGDPHVLLGVVGIIILT
FALYTLKIRGSFIIAVLITSILAWVLKLAPYPSEFFSMPASIGPIAFQLDFKGIFFDASG
AFTLALVPV IITFFVTDLFD SLGTLAGIGHKTDFFNDEEKNKELEKTLEADAVASLGS AV
VGVSTTTAFIESASGVEEGRTGLTAVFTGLFFVLTFLCLPLLKAIPSNAIYPVLVVVGV
LMFSVLEGVNFKDMAISVSTFLTVMMLPLTFSIADGLAFGFLSYSIIKLVQKDFKALNSG
IIILCIISVSVFIFR

>PROKKA_00025 Helicobacter outer membrane protein
MKKTKKTILLSLTLAASLLHAEDNGVFLSVGYQIGEAVQKVKNADKVQKLSDAYEQLSRL
LTNDNGTNSKTS AQAINQAVNNLNERAKTLAGGTTNSPAYQATLLALRSVLGLWNSMGYA
VICGGYTKSPGENNQKNFHYTDENGNGTTINCGGSTKSDGTHSSNGTNTLKADKNVLSI
EQYEKI HESYQILSKALKQAGLAPLNSKGEKLEAHVTTSKYQQDNQTKTTTSVIDTTNDA
QNLLTQAQTIVNTLKDYCPMLIAKSSSGSSGQATTNTPSWQTAGGGKNSCATFGAEFSAA

SDMINNAQKIVQETQQLSANQPKNITQPHNLNLNTPSSLTALAQKMLKNAQSQAIEILKLA
NQVESDFNKLSSGHLKDYIGKCDASAISSANMTMQNQKNNWGNCAGVEETLASLEKSNA
SFSSQTPQINQAETLANAIVQELGHNPFKRVGI ISSQTNNGAMNGLGVQVGYKQFFGEKK
RWGLRYYGFFDYNHAYIKSNFFNSASDVWTVYGVGSDDLNFNFINDKNTNFLGKNNQISFGL
FGGIALAGTSWLNSQFVNLKTI SNVYSAKVNTANFQFLFNLGLRTNLARPKKKDSHHAQ
HGMELGVKIPTINTNYYSFLDTKLEYRRLYSVYLNIVFAY
>PROKKA_00026 Modification methylase TaqI
MLSLPLIEERPFLNHERIKLSFSQIKAYFDTLNFDTSHFVNSNDICTPLDCVKEMIDAI
PSDFFKQEHKLILDCCCGNGNFFAYLETKTSLNLYFNEINPKRIEHVKYFGNNIHLSC
KDFLKFD RATLYDLIVANPPFAKFNALGRTSKNHNLARDFIKKALELTKNGGYILFIVPN
HWMFSFSDRNVLPNLLSQYQFIHLNIGGAKKYFKKVGSSFTWFLQKVPNQKSFSVENHYV
LKDRQRVSLKTHLNFIPLYLNLVCSILDKTINNAHLPTYKIETTSFLHRTTKKEFLSPI
QNKDYPYKIIHTPSQVVYSKIPHKYQEGYKVFLSLTNQYGTIDNCGMTQSIAFVRCKNY
KEALKIKTEL DNEIYLFNLNLTRYGNFNINIRVLQHLPLLESIVLNKQELEFIQKFNGAYY
GKKKERYNFKEC
>PROKKA_00027 hypothetical protein
MGRIMAKRKSDIILKSVDDLKDEIDYKDFEYKEYFNLLCELVPNNSLEKLEINAIDEKNM
KNEGLVYVFVIQGKIFKIGHSITPITKRVQSYNCGKVEYRKNGTCSTTNYFVLQSLKIN
KIVQVYAFFPEQPTYTLFGKTYQDSFSTSKRAENVILENFIKNHNKKPIGCTQT
>PROKKA_00028 Purine nucleoside phosphorylase DeoD-type
MTPHINAKIGDFYPQCLLCGDPLRVSYIAKKFLQDAKEITNVRNMLGFSGKYKGKGISLM
GHGMGIASCTIYVTELIKTYQVKELLRVGTCGAISPKVGLKDIIMATGASTDSKTNRVRF
LNHDL SATPDFELSLRAYQTAKRLGIDLKGVNVFSSDFFYSFETHAFDLMAYNHLAIEM
EAAGLYATAMELSAKALCLCSVSDHLITKEVLNPKERVESFDNMII LALEMMTQ
>PROKKA_00029 Phosphopentomutase
MQKRVVILLLLDSFGIGASEDAKDFGDLGANTLGNIACAFNNLADSNDRHGALKLPYLES
LGLGLSALKATNELPLGFELKPNLIGAYAYAQELSSAKDTISGHWEMMGVPVLFEWGYFK
DKNDSFPKEILDEIMHKTKIKGYLGNC HASGTEIIKDLGEKHLETLYPIFYTSADSVFQI
AVHEEKFGLDNLYALCEEAFQILEPLKIARVIARPFIGANREDFKRTAKRKDYAIKPHKK
LLFEKFIEEKQGEVISIGKIADIYAHVGITQKFKAGSLMELCDVTLEQVKNKNNSLIFT
NFVHFDS DYGHRRDISGYANALEYFDARLKEVLENLRENDLLILCADHGC DPSFKGTDHT
REYIPVLFYHKDLQPAFLGKSESFADIGQSIAYFLGLSPLDYGKNLLNFKGQS
>PROKKA_00030 Nucleoside permease NupX
MISSSLFSVVGMAVFLIAWVFSSNKRAINYRTIISAFVIQVALGALALYVPLGREMLQG
LASGIQSVIGYGYEGVRFLFGSLAPNAKGDQGIGDFIFAINVLAIIFFASLISLLYYLK
IMPLFINLIGGALQKCLGTSKAESMSAAANIFVAHTEAPLVIKPYLKSMDSSEIFAVMCV
GMASVAGPVLAGYASMGIPLYLIAASFMSAPGGLLFAKIIYPQNEAISSHADVSAEKHV
NAIEA IANGASTGLNLALHVGAMLLAFVGM LALINGLLGVVGGFLGMEHLSLGLILGTLL
KPLAFMLGIPWSQAGIAGEIIGIKIALNEFVGVMQFLPYLGDNPPLILSEKTKAIITFAL
CGFANLSSVAMLIGGLGSLVPKKDLIVRLALKAVLVGTLSNFMSATIAGLFIGLNAH
>PROKKA_00032 Inner membrane transport protein YajR
MFKKIFPLALVSSRLFLGLFIVLPVISLYADSFHSSPLLVLAVGGAYLTQIVFQTPMG
ILSDKIGRKVVVMVCLLLFLAGSLVCFIANDIVWLVIGRFIQGMGALGGVVSAMVADEVK
EEERTKAMAIMGAFIFISFTISMAIGPGVVAFFGGAKWFLLLTAILTLLSLLMLLKVKDA
PKISYQIKNIKAYQPNSKALYLLYLSSFFEKAFTLIFVLIPLALVNEFHKDESFLILVY
VPGALLGVLSMGVASVMAEKYNKPKGVMLSGALLFIMS YLCLFLADSSFLGKYLWLFIVG
VAFFFIGFATLEPIMQSLASKFAKANEKGKVLGQFTTFGYLGSFVGGVSGGLSYHHLGVS
NTSLIVVALGLIWGLSLFLLNNPSKQKNIFYPLDAYNEEQFETLGDKIIEWYVNISEEII
IVKYNSDHISEEEIIRLAQSFRK
>PROKKA_00033 tRNA 2-thiocytidine biosynthesis protein TtcA
MAYEISKVVLHIVGKTNATYKLIIEGDKILLGLSGGKDSIMLACILARMQKHAPFKFDFK
AVTVHYGLGEDLKWLSDLCEQEQGIEHEIITYTQIAATINEKRREKSSFCFSRLRRGTLY
SKALEEGYNKVAIAHHLDDAVESFFMNFTYNGSLRSMPIIYRAENGLLVIRPLIKVREAS
SIHFVTSQNIPVAPDCNCPAKQPTS DKPPIARLATKNFLKEMQNLHPRFFDSLENAFNNV
QANSFSDAKYLD A
>PROKKA_00034 Sodium/hydrogen exchanger family protein
MHAEFFTFALIMLLIVMAPYISRISRLPITVVEILFGSVGAYVGFIETPKGFEIMSEIGF
LFLMFLCGLEVEIYLFKKLGVSLLRKIFAYLLILYTLTSFILTFSLNLEPIFMVIFPIISL

GMIMTLVKDYRKEILWLDLVLKVGIGELLISIFGLVVVDGVYSHGLGMELIKDLGILIVF
LILIIIVAFQIFKTLFWWFPHLKL FVMPKSSQFNQDVRFSMLFFSLVAIVVWLKIEMVLG
AFLAGLVVSTFFPHKSELIHKLNDVGGFFVPLFFI HVGSTLDLKLVLNPHLILQGILI
VIAMLSLHLITSTLLWRKYFKEAKHLFSFALGASMP LTFVLVTTAAVGLKAQAISQNTYYA
LLMAAIFEGVLFTIAIKILNKKA

>PROKKA_00035 Multidrug export protein MepA

MLKKKIDLHKDSIRKLFFYYFIPLAFSMISLSTYSMIDGMFVGKKLGKEAIAAVNIAWPI
FPGLIAYELLFGFGAVSIVGYFLGQNKTHRARLVFSSVFYFVAISAFILSMALLPFSETI
ARFFGSNDALLSMSKRYIEIILMGAVFMVLHPLADV FVVNDKRPILAMVAMLIGSLANIF
FNYLFIFVLEVGVQGSATVIGHAIGVLVLMQHFWFKKGQLYFIKRFSLSSVISSAKSG
VPQSTAEFSASVMILLFNTAIMHTAGERFVS VYGIVMYNAIIFFTTLFAISQGIQPIASF
SYGARNLERVKEVFVFLKAAFCIGVVFYGVYYFLDEF LIKLYLQPSEQDPLFMQETKRA
MNIYYVGIVFLGMTLLCAVFFQSIQRTKSSFIITLSHTLGFMVILLPILSHFYGINGVWV
TYPIAQFLAFLVALGVTYYEIKKGVFTPYKEKNPIALKT

>PROKKA_00036 Sugar efflux transporter B

MMITKQSYQKFALMRVVFVLSAFIFNTTEFVPVALLSDIAKSFEMESATVGLMITAYAW
VVSLGSLPLMLLSAKMERKRLLLFLFALFIASHILSALAWNFWVLLLSRMGIAFAHSIFW
SITASLVIRVAPRNKKQQALGLLLALGSSLAMILGLPLGRIIGQILDWRSTFGVIGGVATL
IMLLMWKLLPHLPSKNAGTLASVPVLMKRPLLMGIYLLVIMVISGHFTTYSYIEPFIIQI
SQFSPDITTLMLFVFGLAGVVG SFLFGRLYAKNSRKFI AFAMILVICPQLLLFVFNLEW
VVFLQIFLWIGIGITS LGISLQMRVLQLAPDATDVAS AIFSGSYNVGIGSGALFGSIVIHQ
LGLGYIGFVGALGLLALFWLRFITIKFKKT

>PROKKA_00037 Carbonic anhydrase precursor

MKKTFLIALALTASLIGAENTKWDYKKNKENGPHRWDKLHKDFEVCKSGKSQSPINIEHY
HTQDKTDLQFKYAASKPKAVFFTHHTLKASFEPTNHIN YRGHNYVLDNVHFHAPMEFLIN
NKTRPLSVHFVHKDAKRLLVLAIGFEEGKENPNLDPI LEGIQKKQNFKEVALDAFLPKS
INYYHFNGSLTAPPCTEGVAWFVIEEPLEVS AKQLAEIKKRMKNSPNQRPVQPDYNTVII
KSSAEIR

>PROKKA_00038 hypothetical protein

MESVKTGKTNKVGKNTETADTKANKETHFKQANAITNIIRSVGGFFTKIMKRVRELVKKH
PEKSNAALVVLTHIACKKAKELDDKVQDKSKQAEKENQINWWKYSGLTIATSLLLAACSV
GDIDKQIELEQERQKANNTQKDLVKEQKDLVKEQKDLVKEQKDFIKYAEQNCQENHSQFF
IKKLGIKGGIAIEIEAECKTPKPTKTNTQTPIQPKHLPNSKQPHSQRGSKVQELIAYLQKE
LESLPYSQKAIKQVDFYRPSSIAYLELDPDRDFKVTEEWQKENLKIRSKAQAKMLEMRNL
KPDPPQAHLASQSLLFVQKIFADVNKEIKVVANTEKKVEKAGYGYSKRM

>PROKKA_00039 Aspartate-semialdehyde dehydrogenase

MKTYNVAIVGASGAVGQELIKLENS SFPIKKFVPLASARSAGKKIRAFNKDYEILETTH
EVFEKEKIDIAFFSAGGSVSEEFATSAAKTVLVIDNTSFFRLHKDVPLVPEINAKEIFN
APLNIIANPNCSTIQMTQILNPLHLHFKIKSVIVSTYQAVSGAGNKGIESLKNELKTALE
HLEKDPAIDLNLQILQAGAFAYPIAFNAIAHIDTFKENG YTKEELKMVHETHKIMGVDFPI
SATCVRVPVLRSHSESLIAFEKEFDLKEVYEV LKNAPSVVVCDDPSRNLYPTPLKASHT
DSVFIGRLRKDLFDKKTLLHGFCVADQLRVGAATNALKIALHYIKNA

>PROKKA_00040 Histidine--tRNA ligase

MITPKVLSGFKDRLPKDAIQKAQLLAKVS VVFQSFSGFVPIETPHLEYAQTL LDPDASSDIQ
KEIYRFKDHGDRDVALRFDLTVPLARFVSLHHQILGMPFKRYAIGNVFRGERAQKG RYRE
FTQCDFDFIGSESLVCD AEIIQVIIASLKALDLEDFCVSINHRKILNGICEYFGISQVNE
ALRIVDKLEKIGLNGVEEELKKECDLNSNTIKELLE MVQIQQNDLSHAEFFEKIAYLKDY
NENLKKGIQDLERLYQLLGD LQISONLYKIDFSIARGLGYYTGIVYETTLNDMKSLGSVC
SGGRYDHLTKNFSKENLQGV GASIGIDRLIVALSEMQLLDERSTQAKVLIACMH EEFYSY
ANRLAESLRQSGIFSEVYPEAQKIKKPF SYANHKGHEFVVIIGEEEFKSETLSLKNMHSG
MQLNCLSFLEKALEIIGENDEDL

>PROKKA_00041 ADP-heptose--LPS heptosyltransferase 2

MSVNAPKRMRI LRLPNWLG DGMASSLFYTLKHHYPNAHFILVGPQITCELFKKDEKIE
AVFIDDTKKSFFRL LATHKLAQKIGRCDIAIALNNHFYSAFLLYATKTPIRIGFAQFFRS
LFLSHAIAPAPKEYHQVEKYCFLFSQFLEKELDKKSVLPLKLA FNLPHTPTNTPKKIGFN
PSASYGSAKRWPASYAEVAAALLEKGHEIYFFGAKEDTIVFEEILKLIKGLLKNPLL SH
NAYNLCGKTSIEELIQRIAILDLFITNDSGPMHVAASTQTPLIALFGPTDEKETRPYKAQ
KTIVLNHHLSCAPCKKRVCP LKNEKNHLCMKSI TPLEVLEAAHTLLEKP

>PROKKA_00042 hypothetical protein

MKKQILTGVLLSVLAVSSAYAHKDKKDAKKPELSSQLVAHKDKKDAKKPKNSVA

>PROKKA_00043 General stress protein 69

MTYGYGEVHDKKQMVKLIHKALELGINFDDTAEAYGEDNEKLLGEAIKPFDRDKVVVASKF
GIYYADPNDKYATMFLDSSPNRIKSAIEGSLKRLKVECIDLYYQHRMDTNTPIEEVAEVM
QALIKEGKIKAWGMSEAGLSSIQKAHQICPLSALQSEYSLWWREPEKEILGFLEKEKIGF
VAFSPLGKGLGAKFEKNATFASDFRSVSPRFNQENLAKNYVLVELIQDHAHAKGVTPA
QLALSWILHTQKIIVPLFGTTKESRLIENIGALQVSWSQKELEIFQKELTAIKIEGTRYP
ERINEMVNR

>PROKKA_00044 Elongation factor G

MARKTPLNRIRNIGIAAHIDAGKTTTTSERILFYTGVSCHKIGEVHDGAATMDWMEQEKERG
ITITSAATTCFWKDHQINLIDTPGHVDFTIEVERSMRVLDGAVSVFCSVGGVQPQSETVW
RQANKYGVPRIVFVNKMDRIGANFYSVENQIKLRLKANPVPINIPIGAEDTFIGVIDLVQ
MKAIVWNNETMGAKYDVEEIPSDLLEKAKQYREKLVEAVAEQDEALMEKYLGGEELSVEE
IKKGIKTGCLNMSLVPLCGSSFKNGVQTLDAVIDYLPAPTEVVDIKGIDPKSEEEVF
VKSSDDGEFAGLAFKIMTDPFVGQTLFVRVYRGKLESGSYVYNSTKDKKERVGRLLKMHS
NKREDIKEVYAGEICAFVGLKDTLTGDTLCEKNVAVLERMEFPEPVIHIAVEPKTKADQ
EKMVALGKLAEEDEPSFRVMTQEETGQTLIGMGELHLEIIVDRLKREFKVEAEIGQPQV
AFRETIRSSVSKEHKYAKQSGGRGQYGHVFIKLEPKEPGSGYEFVNEISGGVIPKEYIPA
VDKGIQEAMQNGVLGYPVVDKVTLYDGSYHDVDSSEMAFKIAGSMAFKEASRAANPVL
LEPMMKVEVEVPEEYMGDVIGDLNRRRGQINSMDRLGLKIVNAFVPLVEMFGYSTDLRS
ATQGRGTYSMEFDHYGEVPSNIAKEIVEKRKG

>PROKKA_00045 30S ribosomal protein S7

MRRRKAPVREVLGDPVYGNKVVTKFINKMMFNGKKSVAEKIIYKAFNKIEEKSGEKGIEV
FEKALERVRLPEVRSRRVGGATYQVPVEVRASRQQSLSIRWILEATRKRNERMMVDRLA
NELMDAASDKGAAFKKKEDVHKMAEANKAFAHYRW

>PROKKA_00046 30S ribosomal protein S12

MPTINQLIRKERKKVVKTKSPALVECPQRRGVCTRVTTPPKPNLSALRKVAKVRLTSK
FEVISYIPGEGHNLQEHSLVLRGGRVKDLPGVKYHIVRGALDTAGVNRKTVSRSKYGTK
KAKATDKKATDNKKK

>PROKKA_00047 Bifunctional DNA-directed RNA polymerase subunit beta-beta'

MSKKIPLKNRLRADFTKTPTDLEVPNLLLLQRDSYDSFLYSKEGKESGIEKVFKSIFPIQ
DEHNRITLEYAGCEFGKSKYTREAMERGITYSISLKIKVRLILWEKDTKSGEKNGIKDI
KEQSIFIREIPLMTERTSFIINGVERVVVNQLHRSPGVIFKEEESSTSLNKLICTGQIIP
DRGSWLIFYEYDSKDVLYARINKRRKVPVTILFRAMDYQKQDIKMFYPLVKVRYENDKYL
IPFASLDANQRMEFDLKDPPQGVILLAGKKLTSRKIKELKENHLEWVEYPMDILLNRHLA
EPVMVGKEVLLDMLTQLDKNKLEKIHNLGVEFVIINDLALGHDAIIHSFLADYESLKL
LKQTEKIDDENALAAIRIHKVMKPGDPVTTEVAKQFVKKLFFDPERYDLTMVGRMKMNHK
LGLHVPDYITTLTHEDIITTVKYLMIKKNQKIDDRDHLGNRRIRAVGELLANELHSGL
VKMQKTIKDKLTTMSGAFDSLMPHDLVNSKMITSTIMEFFMGQLSQFMDQTNPLSEVTH
KRRLSALGEGGLVKDRVGFEARDVHPHTHYGRICPIETPEGQNI GLINTLSTFTRVNDLGF
IEAPYKKVVDGKVVGETIYLTAIQEDSHIIAPASTPIDEEGNILGDLIETRVEGEIVLNE
KSKVTLMDLSSSMLVGVAASLIPFLEHDDANRALMGTNMQRQAVPLLRSDAPIVGTGIEK
IIARDSWGAIKANRAGVVEKIDSKNIYILGEGKEEAYIDAYSLQKNLRTNQNTSFNQVPI
VKMGDKVEAGQIIADGPSMDRGELALGKNVRVAFMPWNGYNFEDAIVVSEKITKDDVFTS
THIYEKEVDARELKHGVEEFTADIPDVKEEALAHLDSESGIVKVGTYSAGMILVGKTS
GEIKSTPEERLLRAIFGDKAGHVVNKSLYCPPSLEGTVIDVKVFTKKGYEKDARVLSAYE
EEKAKLDMEHFDRLTMLNREELLRVSSLLSQAILLEPFSSHNGKDYKEGDQIPKEEIASIN
RFTLASLVKKYSKEVQNHYEITKNNFLEQKKVLGEEHEEKLKSIKDDILPNGVIKKVKL
YIATKRKLKVGDKMAGRHNKGIVSNIVPVADMPYTADGEPVDIVLNPLGVPSRMNIGQI
LEMHLGLVGKEFGKQIANMLENKTKDFAKELRAKMLEIANAINKDP LTIHALENCSDDEE
LLEYAKDWSKGVKMAIPVFEGISQEKFYKLFELAKIAMDGKMDLYDGRGTGEKMRERNVVG
YMYMIKLHLHVLDEKVHARSTGPYSLVTHQPVGGKALFGGQRFGEVWALEAYGAAHTLK
EMLTIKSDDIRGRENAYRAIAKGEQVGESEIPETFYVLTKEQLQSLALDINIFGDDVDE
VDPKPIVIEDDRPKDFSSFLTALASPEKIHWSYGEVKKPETINRYRTLKPERDGLFCMKI
FGPTKDYECGKYYKPRFKDIGTCEKCGVAITHSKVRRFRMGHIELATPVAHIWYVNSL
PSRIGTLLGVKMKDLERVLYYEAYIVKEPGEAAVDNEGTKLVMKYDILNEEQYQNISRRY
EDRGFVAQMGGEAIKDLLEEIDLITLLQSLKEEVKDTNSDAKKKKLIKRLKVVESFLNSG

NRPEWMLTLVLPVLPDLRPLVALDGGKFAVSDVNELYRRVINRNQRLKRLMELGAPEII
VRNEKRMLQEAVDVLFDNGRSTNAVKGANKRPLKSLSEIIKGKQGRFRQNLGKRVDFSG
RSVIVVGNPKMDECGLPKNMALELFKPHLLSKLEERGYATTCLKQAKRMIEQKSNEVWEC
LQEITEGYPVLLNRAPTLHKQSIQAFHPKLIIDGKAIQLHPLVCSAFNADFDGDQMAVHVP
LSQEAIACEKVLMLSSMNIILLPASGKAVAIPSQDMVLGLYYLSLEKSGVKGEHKLFSSVN
EIIITAIDTKELDIHAKIRVLDQGNIIATSAGRMIIKSILPDFIPTDLWNRPMKKKDIGVL
VDYVHKVGGIGITATFLDNLKTLGFRYATKAGISISMEDIITPKDKQKMVEKAKVEVKKI
QQQYDQGLLTDQERYNKIIDTWTEVNDRMSKEMMSAIAKDKEGFNSIYMMADSGARGSAA
QIRQLSAMRGLMTKPDGSIETPIISNFKEGLNVLEYFNSTHGARKGLADTALKTANAGY
LTRKLIDVSQNVKVVSDDCGTHEGIEITDIAVGSELIEPLEERIFGRVLLEDVIDPITNE
ILLYADTLIDEEGAKKVEAGIKSITIRTPVTCKAPKGVCAKCYGLNLGEGKMSYPGEAV
GVVAAQSIGEPGTQLTLRTFHVGGTASRSQDEREIVASKEGFVRFYNLRTYTNKEGNII
ANRRNASILVVEPKIKAPFDGELRIETVYEEVVSVKNGDQEAQFVLRSDIVKPSELG
VGGKIEGKVYLPYASGHKVHKGGSADI IQEGWNVNRI PYASELLVKDNDPIAQDVYAK
EKGAIKYYVLEANHLERTHGIIKKGDIVSEKGLFAVIADDNGREAARHYIARGSEILIDDN
SEVSANSVISKPTTNTFTKIATWDPYNTPIIADFKGKVSFVDVIAGVTVAEKEDENTGIT
SLVVDYIPSGYKPSLFLEGANGEEMRYFLEPKTSIAISDGSSVEQAEVLAKIPKATVKS
RDITGGLPRVSELFEARPKPKDVAILEVDGIVSFGKPIRNKEHIIVTSKDGRSMDYFV
DKGKQILVHADEFVHAGEAMTDGVVSSHDLIRISGEKELYKYIVSEVQQVYRRQGVSIAD
KHIEIIVSQMLRQVRILDSGDSKFIIEGDLVSKKLFKEENARVIALKGEPAlAEVLLGIT
RAAIGSDSIISAASFQETTKVLTEASIAMKKDFLEDLKENVVLGRMIPVGTGMVKNKKIV
LRTLEDNSKF

>PROKKA_00048 50S ribosomal protein L7/L12

MAISKEEVLEYIGSLSVLELAELVKMFEEKFGVSATPTVVAGAAVAGGAAAESEEKTEFN
VILADSGAEKIKVIKVVREITGLGLKEAKDATEKTPHVLKEGVNKEEAETIKKKLEEVGA
KVEVK

>PROKKA_00049 50S ribosomal protein L10

MQKQHQRQHKVELVANLKSQFADAKALLICDYKGLSVKKLEALRNKARAQGIKVQVIKNT
LAHIAMKEAGYSDLDLKETNVFLWGGDQIALSKLVFDFQKEHKDHFVLKTGLFDKESVSV
AHVEAVSKLPSKEELMGMLLSVWTAPVRYFVTGLDNLKAKEEN

>PROKKA_00050 50S ribosomal protein L1

MAKKVFKRLEKLFSKIQNDKAYGVEQGVEVVKSLASAKFDETVLRLGVDPRHADQMV
RGAVVLPHGTGKKVRVAVFAKDIKQDEAKNAGADVGGDDLAEIKNGRIDFDMVIATPD
MMAVVGKVGRILGPKGLMPNPKTGTVTMDIAKAVSNAKSGQVNFVRDKKGNVHAPIGKAS
FPEEKIKENMLELVKTINRLKPSSAKGKYIRNAALSLTMSPSVSLDAQELMDIK

>PROKKA_00051 50S ribosomal protein L11

MAKKVVGEIKLQIPAGKANPSPVGPALGQRGVNIMEFCKAFNERTKDMGSFNIPVIITV
YQDKSFTFITKKPPVTDLIKASGVEKGSNDPLKNKIAKLTHKQVEEIAQLKMEDLNTST
MEAAKKIVMGARSMSGVEVVD

>PROKKA_00052 hypothetical protein

MDWYAIQTYSGSEQSVKKAIEENLANDHNIRDRIQEIIVPTEDIIEVSKKSKTKVTERS
PGYVFIKVDLDTVLWHKIQSLPRVSRFIGENKKPTPLSEADIGHILEKMNNRAAPKPKIF
FEQGEVVRVVEGPFANFTATVEEYDVEHRKLKLNVSIFGRNTPIEILHSQVEKII

>PROKKA_00053 preprotein translocase subunit SecE

MQYKLAREELSKVIFPIKEQIRNALISVLVVSAITLFLALLDFSLGAFVSSVL

>PROKKA_00055 50S ribosomal protein L33 2

MKVKIGLKCSDCEDINYSTTKNAKTNTEKLELKKFCPRENKHTLHKEIKLKS

>PROKKA_00056 Elongation factor Tu

MAKEKFNRTKPHVNIGTIGHVDHGKTTLSAAISAVLSLKGLAEMKDYDNIDNAPEEKER
G
ITIATSHIEYETENRHYAHVDCPGHADYVKNMITGAAQMDGAILVVSAADGMPQTREHI
LLSRQVGVPPIVVLNKGDMVDDQELLELVEMEVRELLSAYEFPGDDTPIIAGSALRALE
EAKAGNVGEWGEKVLKMAEVDAYIPTPERDTEKTFMLPVEDVFSIAGRGTVVTGRIERG
VVKVGDEVEIVGIRPTQKTTVTGVEMFRKELEKGEAGDNVGVLLRGTKKEEVERGMVLCK
PGSITPHKKFEGEIIYVLSKEEGGRHTPFFTNYRPPQFYVRTTDDVTGSITLPEGVEMVMPGD
NVKITVELISPVALELGTKFAIREGGRTVGAGVVSNIIE

>PROKKA_00061 Heterocyst differentiation ATP-binding protein HepA

MAKKKKHISTLKYFLRSLKQIYMLITFKEKMFVFLVLMVAVSSFVEVMSLTLLMPFITL
ASDPSRALDDKDWMVYDFFHFSSPVRLMYFFSFCLVGIYLFMRFYGVSTYLYLGRFSNK

KAYQIKQQLFLQHIKSNYLSHLNHNLDLSLRDIINNKA DGMFMSFNAFLNLLTELT VIVFF
YSTLLITNWKLTLVFTLIISIQIFIITKKITVLIQKKGEIAAKSRAQTLKVFSKFFSNFK
ITKLKDNHEEAHKLFGENSRKAHDTEIIYATLQVVP RYSLETVGFSLLILAVAYILFKYG
EAKMVLPTISMYALALYRTLPSVTGVLNQYNEIAYNQLATNMVFKSLSKTIVEEDLVPLD
FNEKITLQNISFAYKSKHPVLKNFNLTIQKGQKVALIGHSGCGKSTLADIIMGLTYPKSG
EIFIDNTLLTNENRRSWRKIGIYIPQNIYLF DGTVDGNIAFGSAID EKRLIKVCKMAHIY
DFLCEHEGLKTQVGEGGAKLSGGQKQ RIGIARALYDNPEILVLDEATSALDNETESKIMD
EIIYQIAKNKTLIVIAHRLSTIERCEVIIDMSQH KDNLG
>PROKKA_00062 Phosphoglycolate phosphatase
MALEVVLWDFDGVIFDSMHLKNEGFKALFQKHGNKNQEDLKQFEVYHYQSGGISRNEKIQ
YFYNEILKTPIAQEEMDALALEFGAIEQKLFDR EHLNSEVMAFIDKHYQNYIFHIASAA
LHSELQVLCEFLGIIKYFKSVEGSPPNKP KPIIANIIQKYAYNPSRMLMIGDSVNDYESAK
ANKVAFLGYNSKVLKNLVGQDGYQGKYLESF KGFDLQNFIE
>PROKKA_00063 Modification methylase FokI
MNYIGSKYKLIPFIKENIHAVAGHDLSGAIFCDLFA GTGIVGRAFKKAVNKIISNDLEY Y
SFVLNQNYIGNIQEIPNKEELIDRLNSVASKKGF IYSHYSLGGSSRQYFSETNAQKIDAM
RLKIEELKLSQNIDNCAYYFLLASLLESADKVANTASVYGAF LKRLKKSQAQKELILKGAH
FDLSLNANEVYQQDASELIGKISGDILYLDPPYNARQYGANYHLLNTIAAYTPFTP KGKT
GLPSYQKSSFCRSQILNAFEDLIKKVRFKYIFLSYNN EGLMSEIEIKNILKKYGVYSLI
TKTYMRFKADNKRTHKAVHTKECLHILIK
>PROKKA_00064 ICEA Protein
MIKKRQICKECKETGYRFDATKIPGNHYPFYEGEA EYDGCVGCYQYDPIQYRKT CNDGIY
NEGYQNGYHQKTTL
>PROKKA_00065 ICEA Protein
MAKEFNLEFDKGQTLGNSIDRIRLNGYNT ECVFNQSI RQDIKNHYKQQCCTMCGAHGNSE
NTQIEVDHKDGRKNDSRVSDSNAQTFFD FQVLCKACND
>PROKKA_00066 Serine acetyltransferase
MLDLSYSLERVLQEDPAARNKWEVLLLLYPGIHALLCYRLAHALHKRRFYFIARALSQ LAR
FITGIEIHGPAKIGRGLFIDHGMGVVIGETTEIGDDVTIYHGVT LGGTGKFKGKRHPTLG
NRVVVGAGAKVLGAICVGDDVRIGANAVVLSDLPTG STAVGAKAKTITKDR
>PROKKA_00069 ATP synthase subunit c
MKFLALFFLILAGVAFADHGGMGMDMIKSYSILGAMIGLGIAAFGGAIGMGNA AATIT
GTARNPGVGGKLLTTFVAMAMIEAQVIYTLVF AIIAIYSNPFLS
>PROKKA_00070 Polyribonucleotide nucleotidyltransferase
MDFITINSSNKTEEFALKQVAKQATSSLMYRLGKT IILASVCVEREPVSEDFLPLVVQFL
EKSYAAGKIPGGFVKREGRAQDFEILTSRLIDRTL RPLFPKDYRYPTQITLMVLSHDIEN
DLQVSALNAASAALFLAHIAPIKSVSACRIARVDNEFI INPNTSLLNQSSLDLFVSGTKE
SLNMIEMRSLGQQNLAALEEPMLLKALELAQKSLKETCTLYEEAFTPHQNE LLFKESQGI
FNERLLDLLKNQYFDEIIKGIESSALSERENVFSEI AKKISEAHSEFSLEEIEWSLEKVK
KTEIRRMIIQDKIRPDKRALEEVRPILIESDLLPMAHSSILFTRGQTQSLVGV LGTDND
AQTHESLEHKAPIKERFMFHYNFPFPCVGEASSIGATSRRELGHGNLAKRALETSIKNKE
QVIRLVSEILESNGSSMASVCAGSLALYASGVEIYDLVAGVAMGMVSEGQDHAILSDIS
GLEDAEGDMDFKIAGNLEGITAMQMDTKMSGIRLEVLYQALLQAKEARKHILKIMHEAKE
KIVINFSHLPTEIFNVVPDKIVEIIGQGGRVIKEIVEKF EVKIDLNKPSGEVKIMGNKE
HVLKTKEFILNYLHSLDQELEYAIDEVLEAQVKRIVDFGAFLSLPKGGEGLLRKQNM DR
CQVVLKEGDSIRCRVISFNKGKIALDLA
>PROKKA_00071 hypothetical protein
MNTDFSHITDIEGMRVFNEEDALNKLIN EIHTRHIDLKDSIMLALSFNALYLAHALAQKF
GATYDILFLEPILAPLNSKCEIALVSESMDIVMNE SLINSFDITLDYVYGEAKRAYEEDI
LSHIYQYRKGNAIKSLKDNIFIVDRGIETGFRAGLGVQTCLKKECQDIYILTPILA QNV
AQGLES LCDGVISVYRPECFVSVEH HYKELKRLSNEEIEKYL GANNAPNLKKEH
>PROKKA_00072 LPS-assembly protein LptD precursor
MIYWLYLAVFFLLSALEAKEIAMQRFDKQNHKIF EILADKVS AKDNVITASGNAILLNYD
VYILADKVRYDTKTKEALLEGNIKVYRGEGLLVKTDYVKLSLNEKYEIIFFPYVQDSVSG
IWVSADIASGKDQKYKVNMSASGCSIDNPIW HVNATSGSFNMQKSHLSMWNPKIYV GDI
PVLYLPYIFMSTSNKRITGFLYPEFGTSNLDGFIY LQPFYLAPKNSWDMTFTPQIRYKRG
FGLNFEARYINSKDDRFLFNARYFRNYTQYVKRYDLRNQNIYGF EFLSSSRDTLQKYFHL
KSNIDNGHYIDFLYMNDLDYVRFEKVNKRITDATHMSRANYYLQTENNY YGLNIKYFLNL

NKINNNRTFQSVPNLQYHKYLNLSLYFRNLLYSVDYQFRNTAREIGYGYVQNALNVPVGLQ
FSLFKKYLSSLGLWNDLQLSNVALMQSKNSFVPTIPNESREFGNFVSSNFSMYVNTDLARE
YNKLFHTIQLEAIFNIPYYTFKNGLFSQNMYSALSAQALNSYTSPLLRDIDYQGRLYDSVW
NPSSILPSNASNKTVDLTLTQYLYGLGGQELLYFKISQLINLDDKVSPFRMPLESKIGFS
PLTGLNIFGNVFYSFYQNRLEEISVNANYQRKFLSFNLSYFLRNNSGINSIVENPADY
LKAGFSNDFGYFSMSADVGDIRNNVVLNWNVGIYKKIRCFGIGFQFVNQRRPILTDGPN
QPIRVFENNYVKLELDFSPITKTNVTYRSLQRK

>PROKKA_00073 RDD family protein

MRSFNLEKEETEIIETLLMREKMRCLPLYWRILAFLLTDSLLVAFLLSDLLDACDFLHSLY
WLANPIYHSVFVIVSFIILYGVYEIFFVCLCKMSLAKLVFRIKIIDIYLADCPSRAILLK
RLGLKIVVFLCPFLWFVVFKNPNYHRAWHEKSKSLLVLF

>PROKKA_00074 Phosphoribosylamine--glycine ligase

MKDNNNNYNVLIVGNKGREYALAQRLQQDERVNALYFCLGNGGTQDLGENLECEHYEHIVE
LALKKQIHLAIISEEELLVLGLTEMLEKAGILVFGASKEAAKLEASKSYMKAFFVKECGIK
SASYFETNDLKEALSYIQNASFPLVIKALNKNTSIVHQEEEAIKILEDADFQKSNEPVIIE
PFLEGFELSVTALIANDDFILLPFCQNYKRLLEGDNGVNTGGMGAIPANFFSNELEEKI
KNHIFKPTLEKLQADNTPFKGVLLAEIVVIEEKGVLPEYLLDFSVRFKDIECQTILPLLE
SSLLDLCLATAKELHSLVFSKEFVMSVALVSRNYPTSSSPKQTLIDPVDEKKGHIL
LGEVEQDNGVFESSGGRVIFAIGRKSLEARNHAYEIAQKVHFEGMFYRKDIGFKVLDL
KEYS

>PROKKA_00075 hypothetical protein

MTEETYEAYLDTNIKQLEEIRNQKLNALELCKQSGVLVRAFDGKNFSFKCDKTNRSNNP
NEKIDP

>PROKKA_00076 hypothetical protein

MDNIWFQYKIGKQLDELEIEDSLCLSLFKSLENCIKFKKLSNIGLKEAEIKRDAEISRWH
TELLREKNLKEQRRQKEEQEKKSQENFQKFLEFCKQQNRVLKKFDNINFSFECDEPATKK
PLENPTNNILNSNHQNTQPQNNTKMFYDFDNPMFRIY

>PROKKA_00077 ABC transporter ATP-binding protein YtrB

MLVEIENLTTKTYGSLKALDNINLKLPKQQFIGLLGPNGAGKTTLKILAGLNLNYQGEVK
ILNQKIGIETKKSVAFLSDGDFLDPKLTPLKAIKAFYKDFFSDFDESKALDLLKRFVPLK
REFKALSKGMREKLQILTLNRNASLYLFDEPVAGIDPIAREEIFELIAKEFSQNASLLV
STHLVVDVEKYLDLSAIFLKEAKVVAFGDVGELKKGYSSLEAAYKERLK

>PROKKA_00078 Isoprenyl transferase

MDNALKHLAIIMDGNGRWAKLKNKARAYGHKKGVKTLKDITIWCTNHKLECLTLYAFSTE
NWKRPKSEVDFLMKMLKKYLKDERSTYLNNNIRFRAIGDLEGFSKELRDTILQLES DTRH
FKDFTQVLALNYGSKNELSRAFKSLLNPPNNIDPLESLENEISNRDTRDLPEVDLLLR
TGGEMRLSNFLLWQSSYAEFFTPILWPDFTPKDLENIISDFYKRVKFGELKC

>PROKKA_00079 putative FAD-linked oxidoreductase

MEENYHAFFTEASGFLNDRIFKDYLRRLAYGIDASCYRYIPKIVAWVKDEEEVQKLCVLA
KKHGVSLTFRAAGSSLSGQAICDGVLMATHFFKDAKILNNATSILSCGVIGSNANALL
KPYHKKIGPDPSTINTAMIGGIVANNASGMCCGVEQNSYKTLKSLRVILADGTLLDTANQ
ESVEDFKNARKDLIEGVNLNRKEILEDKELHALIKKKYEIKNTTGYSNLALIDFEDPIEI
ISHLFIGSEGTLGFISSVELECVKDYAYKTCALLFYENLERCAKAAQILAALKAKQPEMI
SSAELMDYACLSVKDLEGMPSPVLEIKEPNACLLIQSESDDPLILENNMQAILNALSTI
PVVLDQSQISSDPNVYQSWWKIRKGIFPIAASKRKSQSSVIEDVCFSQENFVEGAKAIEG
FLKKHGFKDNSIIFGHALSGNLHFVVTPILENETERKAFENLVSEMFLMVSESSGSIKAE
HGTGRMVAPFVEMEWGEKAYKIHKQIKELFDPNGLLPDVIITNDKEIHTKNLKSIIYPIE
EHLDMCMECGFCERICPSKDLSTPRQRIVIHREVERLKERVSHGHNEDQVLLDELLRES
EYLAHATCAVCHMCSTLCPLKIDTGSIALNYYQKNPKGEKIASKILNHMQTTTSAARFSL
KSARLVQNLI GSHNLVSLTKGIKKFIKFPKAFHYMPKNNAYPLENTHKSEEKVIYFST
CINRSFAPSTKMADKRCIQEVFESLCQKAKVSVMYPNGLDALCCGKAFINYDTLTKQNN
KNHAIFLQLSDEGKIPIVLDHSACSTHFFKQMKAYKDLKVYDLSVYIEEVLSSKLKFNPI
NEDIGLYTMCALKLENKEELLFNLAKKCTLGEIVIHKETGCCGFAGNKGFFTPELNESAL
NGFQEFYQSYELKGGFSTSSTCEIGLSEKTQFAWQHIAIYLVDACTL

>PROKKA_00080 molybdopterin biosynthesis protein MoeB

MLEDYAIISLEEVSFDFIVVDVRELDEYEELHLPNATLISVNDPEKLADFLSKHKDEKVL
LHCRAGRRALDAKSMHELGYTPYLEGNVYDFEKYGFRMVYDDTCDKKN

>PROKKA_00081 uroporphyrinogen-III synthase

MGEIVVHVSQRIAPYKTLILNEFCYYPLELDLTPFNALIFTSKNAVFSLLETLNKNSPKLK
 MLQNIPAYALSEPTAKTLQDHHFKVAFIGEKAHGKEFAKEIIPLEKKSVLYLRAKEIAS
 SLDAILLEHGINFQQAVVYENKLKHLTLSEQNALKPKEKSILIFTAISHAKAFLHYFEFL
 ENYTAISIGNTTALYLQEQGIPSYIAQKPSLEACLELALSRLRIKEC
 >PROKKA_00082 Putative fluoride ion transporter CrcB
 MNFVFLWAALGGAIGSSRLRYFVGKMMPSKFLMFESFPLGTFSVNVIGCFVIGFMGHAAK
 KVFGDGDFGIFFVTGVLGGFTTFSSYGLDTLKLQKSQYLEAISYVLGTNLLGLIGVAIGW
 FLAKNFIGVN
 >PROKKA_00083 Oxygen-independent coproporphyrinogen-III oxidase 1
 MILYIHIPFCENKCGYCAFNSYENKHGLKEEYIQALCLDLKHALSQTDEPIESVFIGGGT
 PNTLSVKAFERIFESIHQHARLSLDCEITTEANPELISKAWCQGLKDLGINRLSLGVQSF
 REDKLLFLERQHSKNIAPIETILKSGIENVSIDLIYNTPLDNENSLKEELKLAKELPIN
 HLSAYALSIEKNTNLEKNAKPSCTNFDNVIKEALEGFSFKQYEVSNYARGYQVKHNLAY
 WGAKDYLGCAGAVGCVANERFFAKRLIENYIKDPLKRQVETLNKQDKRLEKFLGLRVC
 LGVELSLLDENKVKFLIEENKAFIKNNRLVASDFFMADEMALWLL
 >PROKKA_00084 Cytochrome c-553 precursor
 MKKVIMTLGVLAFAFANALMATDVKALAKGCAACHGVKFEKKALGKSKIVNMMSEAEIEKDL
 MDFKSGANKNPVMTAQAKKLSDEDIKALAKYIPTLK
 >PROKKA_00085 RNA pyrophosphohydrolase
 MLHKKYRPNVAAIIMSPDYPNACEVFIAERIDIEGAWQFPQGGIDEGETPLEALYRELLE
 EIGTNEIEILAQYPRWIAYDFPSNMEHKFYAFDQKQRYFLVRLKHTNNIDLNKHTPEFR
 AYQFIHLKDLLKKIVPFKRQVYRQVIAYFRKEGYLVC
 >PROKKA_00086 Aspartokinase
 MLIVQKYGGTSMGSIERIHNVQRVLESVKLGHQVVVVVSAMSGETDRLLEFGKNFSHNP
 NKREMDRIVSTGEWISSAALSMALERYGHRAISLSGKEAGILTSSHFNQNAVIQSIDAKRI
 TELLEKNYIVVIAGFQGADVQGETTTTLGRGGSDDLAVAGALKAHLCETDVGDVYTT
 DPRIEKAQKIAQISYDEMELASMGAKVLLNRSVELAKKLSVKLVTRNSFNHSEGTIV
 AEKDFKGERMETPIVSGIALDKNQARVSMEGVEDRPGIAAEIFGALAEYRINVDMIVQTI
 GRDGKTDLDFTIVKTQIEETKQALKPFLAQTDSDIDYDENIAKVSIVGVGMKSHSGVASIA
 FKALAKDNINIMMISTSEIKISVLIDIKYAEALAVRTLHAVYQLDR
 >PROKKA_00087 DNA replication regulator
 MKNFYDWIKEFVRDQGEFIAQQSGWLELERSSYAKLIAQTISHVLNNGSSLLVSADSSRH
 FLNYILSNLNPDKLKERPLLSVIDFNASSFYPKNDANLSLATIELTYQNPMFWHIGKIEN
 EGLKTLSSKIPSFLWLFEELKEDCLLLKEHDSLLDYKLLQLFKLFENALFSVLYNKVTL
 >PROKKA_00088 DNA polymerase III subunit delta'
 MKNLNLRIYTDNLEESLEEAASLFEHHIKFYTEILEKDKKVIKTFNKDFKIEHAKKEVISK
 AHLKHSELNAFLIAAPSYGIEAQNALLKILEEPPNNVCFIMFAKSPNHVLATIKSRLIKE
 DKRQKNPLKPLDLDLRLDLKDIYAFLKNLDKENFDSRENQREKIESLVNRHKIPLN
 EQELQAFDLAIKANSSYYKLSYNLLPLLLSLLSKKTP
 >PROKKA_00089 Dihydropteroate synthase
 MIVKRLNPDALKNALQKIGPEEIAQNMMHMQGVSVFVEIQHLPLSATLILKQEASVGGD
 FATPRDCILAKEPFYDGVLVVSASQLERLIVKCHSQPFGKHLAQELKSHLKAKKPNAPQ
 IMAVNLNTPDSFYEKSRFDSKKALEEIQYQLEKGITLIDIGAASSRPESEIIDPKTEQDR
 LKEILLEIKSQKLYQYVQFSIDTYHAQTAQMALEHYFSILNDVSGFNSVEMLEVAKYK
 TCVLMHTQKTPKDMQENVFYHNLDFEVDVRFKEKLEVLEKYALQDVILDIGFGFAKLKEH
 NLALIKHLSHFLKFKKPLLVGASRKNTIGLITGREVQNRLAGTSLSLHLMALQNGASILRA
 HDIDEHIDLIKVFKSLEETD
 >PROKKA_00090 hypothetical protein
 MAVSSINQFDSNLYGLNNAKSAPKEDLAPVENAEKIEREKKDAPTENLPFNPDKRETYGF
 LVLELMSDREYEAFLRATAGMDESQKRLAAQSLYSLTDFYNGKFSKEVENAPKQPINGLH
 KKALQTFNATNHNAFLQRYQNAYNPNSSMDVV
 >PROKKA_00091 putative DMT superfamily transporter inner membrane protein
 MRNTILFGVSMILLANLCFGIMSAFVKITADYFSPMENVFYRSITMTLLLLLVYPFKPYR
 LKSYKQGGFKKLAFRVVVGGGLAMLAFFYNIEKISLATATAFSQCAPIYTVLLSPLLLKEK
 LKRSALISACIGLVGVVLISDPSVENVGPVEILMGILSGIFVSLAYITLRLDREYYDKQA
 VILAFAFGMSLLGLIGMFIDIPFLSTGIHIPRKEDILWISLIGISGTLGQYFLTYAYMNA
 PAGIIAPIEYTRIVWGLLFGLYLGDFTDLKSSLGVALILCSGLLIALPALLKELKKI
 >PROKKA_00092 hypothetical protein

MQLSPLQSAALLYFRYFIYPEKKTRSFDLSDLI FIVMVFLVLALGLLMSEEISISYNEAKD
FFYSDVWFVKIAQKSTAILGQNDLALRLPFLIAHIINMFLFYLI GRKILKKPKDALYVVL
TYALLPGVNLFAILLAKSVLVLSLGLLISYLIKTQKIPYLTL SACAFLDGAFIPLLLGV
FAYALRRRYFKSAIFALVVLIVNIALFSGDFNKGLPSGYFIDTCELMMLYSPLLFLYYP
YTLYKALLDKKPSLLAFMSASGWLFPLLLSMRQEIDLKTFAPLALIGLPLFVKSVLNSLR
VRLKEFRGQYYLRVFSLYLLMLTETLFLWGSKISGAN EKLLNRHFLAKEVATALQLRGIH
QIRTNDKQLALRLQFYGIKEGGRLRLINTKISKKRPDITIIYADKILQSYSLVRH
>PROKKA_00093 hypothetical protein
MRLKLTHITHISHKIANDFIHSLLELKAPRELLCELIEGILEKSVKKENAIDEQARELL
EENTDEIEFMRMDERQLFWMIKRQIAQKEGFHLFWEERCSDLSHQILNKILDEDLIMFSV
SENLRNLIIYKSIDTYSKAYESIENEVHEKIKHYKRKL PVGSDEYELVFERLYEEELRRK
GFL
>PROKKA_00094 Carbamoyl-phosphate synthase small chain
MVSLYLENGLFQLQAQSFAGSTQVGELVFNTSM SGYQEVISDPSYKGQFVFSMPEIGVV
GANPKDDESFFSCAGILARHYNEFFSNSRADSSLSLYLKKRGVLGISGVDTRSLIKTLRH
YGCLMMVASTIEHDKNKLEKILKNAPKISHSPLVSSVSTPKIITHQRATFDNTLDYKPF
DEKTPHKIIAVLDFGAKGNILNELQNVGLKALIYPHHTKASELIKAYEKKRN
>PROKKA_00095 Carbamoyl-phosphate synthase small chain
MKKKEISGIFLSNGPGDPLSLQQEIGEIKRLIDAKIPMFGICLGHQLLSIAQGYPTYKLG
FGHHGSNHPVKNLKTNAVEITAQNHNYCVPEAIEEIAIITHRNLFDNTIEGVRYKNAPII
SVQHHPESSPGPKESHYIFKEFVELLKDF
>PROKKA_00096 Formamidase
MGSIGSMGKPIEGFLVA AIQFPVPIVNSRKDIDHNIESIIRTLHATKAGYPGVELIIFPE
YSTQGLNTAKWLSEEFLLDVPGETELYAKACKEAEVYGVFSIMERNPDSNKNPYNTAII
IDPQGGKIILKYRKLFPWNPIEPWYPGDLGMPVCEGPGGSKLAVCICH DGMIPELAREAAY
KGCNVYIRISGYSTQVNDQWILTNRSAWHNLMYTVSVNLAGYDNV FYYFGEGQICNFDG
TTLVQGHRNPWEIVTGEIYPKMADNARLSWGLENNIYNLGH RGYVAKPGGEHDAGLTYIK
DLAAGKYKLPWEDHMKIKDGSIIYGYPTTGGRFGK
>PROKKA_00097 Maf-like protein
MELILGSQSSARANLLKEHGIKFTQKALDFDEESLKT TDPKEFVYLACKGKLEKAKKLLT
NHCTIVVADSVVSVGNRMQRKAQNKQEALEFLKRQNGHEIEVLTCSALISPKLEWLDLSV
FRARLKAFDPSEIEKYLESGLWQESAGCVRLDFHRPYIKSSSKNLSVGLGLNVESLLGA
LQLGAKLSLL
>PROKKA_00098 Alanine--tRNA ligase
MDIRNEFLQFFHNKGHAVYPSMPLVPNDATLLFTNAGMVQFKDI FTGIVPRPSVPRAASS
QLCMRAGGKHNDLENVGYTARHHTLFEMLGNF SFGDYFKEEAILFAWEFVTKNLGFKPKD
LYISVHEKDDEAVKLWEKFVPVDRIKKMGDKDNFWQM GDSGPCGPCSEIYIDQGEKHFKG
SEDYFGGEGDRFLEIWNLVFMQYERSNDGVLSP LPKPSIDTGMGLERVQALLEHKLNNFD
SSLFAPLMEEISELTSLDYASEFQPSFRVVADHARAVAFLLAQGVHFNKEGRGYVLRIL
RRALRHGYLMGLKEAFLYKVVGVCQFSNTHAYLKE SKEIVMKECFEEEEERFLETLESG
MELFNLSLMHLNENKIFDGKIAFKLYDTFGFPLDLTNDMLRSHGARVDMQGFENCMQEQV
KRSKASWKGKQNNADFSAILNAYAPNVFVG YETTECPAKVLGFFDSDFKEITEANPNQEV
WVLEKTPFYAEGGGAIGDRGTLFKDNEEAALVLDTKNFFGLNFS LLEIKKALKKGDQVI
AQVSDERLEIAKHHSATHLLQSALREVLGSHVSQAGSLVESKRLRFDFSHAKALNDEELE
KVEDLVNAQIFKHLTSQVEHMLPNQAKDKGALALFSEKYAENVRVVSFKEVSIELCGGIH
VENTGLIGGFRIVKESGVSSGVRRIEAVCGKAFYQLAKEENKELKNAKTLLKNNDVIAGI
NKLKESVKNSQKAPVSMDLPIEKIHGVSLVVG VVEQGDIKEMIDRLKSKHERLLAMVFKK
ENERITLACGVKNAPIKANAWANEVAQILGGKGGGRDDFASAGGKDTEKLQAALNLAKNT
ALKALEG
>PROKKA_00099 hypothetical protein
MFHEFRDEISVLKANNPHFDKIFEKHNQLDDDIKTAEQQNASDAEVSHMKKQKLKDKDEI
HSMII EYREKQKSER
>PROKKA_00100 Helicobacter outer membrane protein
MSAEDDGFYMSAGYQIGEEAQMVKNTKGIQELSDNYEKLNNLLTRYSTLNTLIQLSADPS
AINSARENLGASAKNLIGDTKTS PAYQAVLLAINAAVGFWNVLGYATQCGGNANGTKSTS
STTIFNNEPGYRSTSITCSLNGYTPGYYGPM SIDNMKKLNEAYQILQTALKQGLPALKEN
NGKVNVTYTYTCSGEGNDNCKAPVSNNGGTKTETQTIDGKSVTTTIS SKVVD SXAXGNTRG
VSYTEITNKLDGVPDSAQALLAQASTLINTINTACPYFSVTNKS GGPQMEPTRGKLCGFT

DEISAIQKMITDAQDLVNQTSAINANSQSTSVGESNKPFPNFTDASFAQGMLANASQAQAK
MLDLANQVGQAINPSNLSGTFKNFVTGFLATCNNPSTSGANTTQGSAPGTVTTQTFASGC
AYVQETIANLTNSIAHFGTEAGQIQQAADIADTLVHFHFKSKYSELGNTYNSITTALSKVFN
AQSLQNVVSKKNNPYPSPQGIETNYLNNQSYNQIQTINQELGRNPFRKVGIVSSQTNNGA
MNGIGIQVGYKQFFGQKRKWGARYYGFDDYNHAFIKSSFFNSASDVWTYGFGADALYNFI
NDKATNFLGKNNKLSVGLFGGIALAGTSWLNSEYVNLATVNNVYNAMNVANFQFLFNMG
VRMNLARPKKKSDHAAQHGIELGLKIPTINTNYYSFMGAELKYRRLYSVYLNIVFAY

>PROKKA_00101 30S ribosomal protein S18

MERKRYSKRYCKYTEAKISFYIDYKDLDMKLKHTLSERYKIMPRRLTGNSKKWQERVEVAIK
RARHMALIPYIVDRKKVVDSPFKQH

>PROKKA_00102 Single-stranded DNA-binding protein

MFNKMVIMVGRRLTRNVELKYLPSGSAAATIGLATSRRFKKQDGTLGEEVCFIDARLFGRTA
EIANQYLSKGSSVLIEGRLTYESWMDQTGKKNSRHTITADSLQFMDKKSDNPQANAMQDS
IMHENFNNAYPANHNAPSQDPFNQAPSYAQNAYAKENLQAQPSKYQNSVPEINIDEEEIF

>PROKKA_00103 30S ribosomal protein S6

MRHYETMFILKPTLVEEEIKSKIEFYKEVITKHHGVIETSLDMGMRNLAYEIKKHKRGYY
YVAYFKAEPMSILELERLYRINEDVLRFRIVIKYESKKEVEAWHALVDRANKKPSHAKEKH
EKTEHTSHHTEEAESVGSHE

>PROKKA_00104 DNA polymerase III subunit delta

MYRKDLDNLYLKQRLPKAVFLYGEFDFFIHYIQTISTLFGKNNPDTETSLFYASDYEKSQ
IATLLEQDSLFGGSSLVLKLDLAFHKKFKENDINLFLKALERPSHNKLIIGLYNAKSDT
TKYKYTSEIIVKSFQKSPKDEAICARFFIPKAWESLKFLQERANFLHLDISGHLLNALF
EINNEDLSISFNDLKLAVLNAPITLEDIQELSSNAGDMDLQKLILGLFLKKSVDIYDY
LLKEGKKDADILRGLERYFYQLFLFFAHIKTTGLMDAKEVLGYTPPKEIAENYAKNALRL
KEAGYKRVFEIFRLWHIQSMQGGQKELGFLYLTPIQKIINP

>PROKKA_00105 Ribonuclease R

MQGFLRSLFFGVKKIPKPFAPLIEKGVLEKEALELKKDRYFLKEGFDIGRIEWAENKAFFI
SLAKNYSKDPLIKNLPPSFKTDALILCQIECSKKRPIAFFKAALLSTDHTMIAYLAKEKN
QIVAIPFKEPFKKPILLKHSQRSLELPRHCVVKIDLKKREISEILGALEDPLIDENLSL
SLFDRVKDFSKDCLNLAQHYAQLKASDFKDRINYSHIPFITIDPKDAKDFDDAIFYDQEK
RVLFVAVADVSEFVPHKSSLDKEARLRGFSVYFPNSVYPMLPLSLSQGACSLKAFEKRLA
LVYEIPLDDLKNARLSQGVIEVRANCTYEEINYNFLNANQSSLDKDLQQSLLGFLEMALKL
KKERLKKGFNFNSFENKLYLNKEGRIEKIETEKESDAHTLIEEAMLLANQSSARLLDEHF
HNKGIYRTHKEPSLEQQKRLYAKLFDYEVIRPKNMSFFPFLEHALKIAKEKKLEREVSRL
IISQNALALYSPMQESHFGLGFASYTHFTSPIRRYSDLALHRLKELLFHQAAGCSYLLE
ETPELCAELNALQKKTALIERNFIKRKFARLALFLEKEFLGVVLEVVDWVVGLEKEFIG
LKVLVKTNKVFKPLEKVRVTITHADLILGQVRGEIIDKD

>PROKKA_00106 Shikimate dehydrogenase

MKLKSFGVFGNPIKHSKSPLIHNACFLTQKELGFLGHYHPILLPLESRIKNEFLHLGLS
GANVTLFPKERAFAQICDKIKGIALECAVNTLVLENDELVGYNLDALGFWLSLGDSEYQS
ALILGSGGSAKALACELKKQGLEVSVLNRSSRGLDFFQRLGDCDFIEPPKSAFDLIINAT
SASLNNELPLNKEVLKGYFKESKLVYDLAYGFLTFFLSLAKELKTPFQDGKDMLIYQASL
SFEKFSTSQIPYSKAFVMRSVF

>PROKKA_00107 Bacterial SH3 domain protein

MKTEMKSSKLKFMRLVVLAFMMLLYALVHAALGFYVKDSTPINPNIEKTETERQNSAL
SPKQEEANTTTTATEENPTKDTVPLDAAQKQETKQETKQEQEKENEKQDSVSPIQNN
QKTPTTPLMGKKPLEYKVAVSGVNVRAFPSTKGKIIGSLTKDKSVKVLEIQNDWAKIEFS
NKTGYVFLKLLKKA

>PROKKA_00108 Inner membrane ABC transporter permease protein YejB

MIAYILKRLLLIIPITLLAIMTINFFLIQSAPGGPIEQMMAKINNTQSKEIQGVVKDRAYR
GSQGLESDDLLENLKKLYGFDKPIGERYLLMLKKYLQDFGESFYRQIKVVDLIKEKLPVS
ISLGLFSTLLIYLIISIPLGIFAKRNNELDLVSSVVIIVANAIPAFLFAVVLIVFFAGG
NYWHWFPLKGLVSDNFESLSTLGKIKDYLWHITLPVLCISLGGFASLTLLVKNSFLDEM
KLYVLSAKANGCSVGRIFYAHVFRNAILLVAGFPQAFLGMFFSSSLIEIVFSLDGLGL
LGYESIVSRDYPVVFSGSLYIFTLLGLVASLISDLLCVVIDPRIDFEKR

>PROKKA_00109 Oligopeptide-binding protein AppA precursor

MVFKILGLWLGVFCFLKATPYLYLGEEPKYKDNFTHFEYANPNARKGGVLRNDAIGTFDS

LNPFAKLGTKAEGLDLIYDTLMVQSLDEPFPAEYPLIAKDAEVAKDNSYVIFTIDKRARFS
 NNAPILASDVKFSFDTIMKLGSPYRQYYQDVKKAVILDKHHVKFIFKTENKELPLILG
 QLQIFSKKAFQKDYFEKNPLLI PVSSGPYVIASFDVGKKITYQRNPNYWARNLPSRKGQF
 NFDEIKFEYYKDETIALQAFLSGAYDWRIESTAKVWARGYVGKAIDNKKITKYLIAHKMP
 SGMQGGFFNTREIFKDKRVREALFYAFDFEWANKNLFFS QYKRTTSFFSNSIYASPLP
 SPEEKALLAPYEKSLDERVFKEPYIVPRTDGPDLGYNLRENLYAQKLLESAGFSYKNM
 RLVDKNNKPFSTLLNSPAFERLALAFANKLRVLGIEMKIQRVDSLQYVNRVKS YDFDM
 IVGVIGQSSFPGNEQRFYFGSLSAKEKGARNYAGISSKAVDDLIEKI INAKDYKEQLAAI
 QAMDRVLLWGFYVIPHFYLPNYRIAAYNYIGMPEISPSYGFSPYLWWVKEKDLQ
 >PROKKA_00110 Tryptophan--tRNA ligase
 MHKKRVFSGIQPTGQIHLGNYLGAIKHWVEMQDEYENLFCVVNSHAITLPIDPAFLKSQS
 YELVKLLLACGISPKQSGFLFIQSEVDEHAALAWLLNCQVSMGEMQRM TQFKDKSLKNPKS
 VNVGLFNYPILMASDILLYQSDLVPVGEDQKQHLELTRNIAEKFN RDFGSCFKVPEPLIA
 KVGARVMGLDDPKVKMSKSYQGANAIFLLDEPDIIVRKIKKAATDSMGVIAFDEKREGI
 FNLLNIYMLLSNESPENIEERFKNKG YGDFKKELAEVVIQALKPIQERYKEISDDEVKAI
 LNHGVKKARPLAQMTYQKAKELMGLI
 >PROKKA_00111 biotin biosynthesis protein BioC
 MDSFYSFNQALAFNQHSFNKHAKTYHLFAHIQQQIAICLVQFLKQKH YTKVLDLGS GS GAV
 FNALERQHIVIEDFIALDNSINMLKLHPTQSLNIQRISLEHADFE EHVFCDYDLVSSSS
 LQWARDLKS VLEKIALSSKEVALAIHTDFSLHEVHEFLGTPSPLRDLKTLKSLIKNAFKH
 FQIELENKRFALYFN RKQDCLNYLKKCGLLGGSTLSFKQKKHFFQNMAFEKLSYEVLLFS
 GIKRS
 >PROKKA_00112 Protein-export membrane protein SecG
 MTSALLGLQIVLAVLIVVVLLQKSSSIGLGAYSGSNDSLFGAKGPAS FMAKLTMFLGLL
 FVINTIALGYFYNKEYGKSILDETKTNKDLSPVPATGTNLNPTLNPTLNPTLNPLEQAPT
 NPLMPTQTPNELPK EPAKAPSVESPKQNEKNEKNDAKENG IKGVEKTKENAKTPPTTHQK
 PKTHATTNAHTNQKKDEK
 >PROKKA_00113 Ribosome-recycling factor
 MLQAIYNETKDLMQKSVQALNRDFSTLRS AKVSVNILDHIKVDYYGTPTALNQVGSVMSL
 DATTLQISPWEKNLLKEIERSIQEANIGVNPNDGETIKLFFPPMTTEQRKLI AKDAKAM
 GEKAKVAVRNTRQDANNKIKKLEKDK EISEDESKKAQEIQKITDEAIKKIDESVKNKED
 AILKV
 >PROKKA_00114 Orotate phosphoribosyltransferase
 MDIKACYQNAQALLEGHFLSSGFHSNYYLQSAKVLEDPKLAEQLAKELAKQIQEAHLNI
 ECVCSIPAIGGILAGYELARALGVR FIFTERVDNIMTLRRGF EVKKNEKILVCEDIITTGK
 SAMECAKVLEEKGAQIVAFGALANRGICKRVHSHLKAQEGACLP SHLPLFALEDVFVDMH
 KPSSCPLCTTSVAIKPGSRGN
 >PROKKA_00115 RDD family protein
 MATKKTTRNKTPEKKRVLESPLKGLYLSRLKAFITDIFMIYTPMLYIMTYAILGSAKDF
 RENQSAIFLCLLFYALTHSFFIAFKSQSPGMRYAQFKLVKNNGEEVGF FLALWRFVLWVL
 SMGLLIGFVTPFIFKFFLHDKLSGTRIEIIEAT
 >PROKKA_00116 NAD-dependent protein deacylase
 MKNLVILSGAGISAESGIKTFRDADGLWEGHDIMEVASPYGWKKNPQKVLD FYNQRRRQL
 FEVYPNKAHKALAELEKHYQVHIITQNVDDLHERAGSSRILHLHGELL SVRSEKDPNLVY
 RWEKDLNLGDLAKDDSQLRPDIVWFGEAVPLLKEAISLVKQAHLLIIIGTSLQVYPAASL
 YTHASKDTLIYYIDPKAKNAHLPQNVQCINESAVHAMQDLMPKLIEMAS
 >PROKKA_00117 NAD(P)H-quinone oxidoreductase subunit 3
 MQQATEALNHPYFGVFVLLVFTFVFNLT LRIQRFLSRKMAQKKGEK LKLAPYECGPVAL
 KQPNRVSHHFYIMAMLFILFDVEIVFMFPWAIDFKKLGLFGLVEMLGFVFFLTIGFIYAL
 KRNALSWQKLEVK
 >PROKKA_00118 NADH-quinone oxidoreductase subunit 6
 MQQAPVVLSTLDKLLNWGRNSNLWPLTYGLACCAIEMMATGGS RFDFDRFGTIFRASPRQ
 SDVMIIAGTLTKKHAEFMRRLYDQMP EPKWVISMGSCANTGGMFN TYATVQGADRIVPVD
 IYLPGCAPRPETLQYALMVLQDKIRRSKA IKQDAPKRLV
 >PROKKA_00119 NAD(P)H-quinone oxidoreductase subunit J
 MVRKQSPYEDVQKQSRQHDPYKII EPTPKKYLEGSAYEVIYNHLSYKHEILD KYIETNTA
 VFWIKKDDIFSVATILRHLGYECLSEMSAIDLCAKKGHFELFYQFVGFSDNCKNRRRV
 KCVLLPNESVDSL SFLYRSANWSEREAYDLLGIVFDKHPY LKRLIMPHDWVGHPLLRSY

LKGDEFAQWYEVDKIFGKEYREVVGKEQRDSARVDEKDTFNFAKIGYEQGKGEELKEIEE
KHAFKKIPFVKDLHKIAPTILKKRL

>PROKKA_00120 NADH-quinone oxidoreductase subunit D

MAQNFTKLNPOFENIIFEHDDNQMILNFGPQHPSHGLRLILELEGEKIIKATPEIGYL
HRGCEKLGENTYNEYMPTTDRLDYTSSTSNYAYAYAVETLLNLEIPRAQVIRTILLE
LNMISHIFFISVHALDVGAMSVFLYAFKTRGYGLDLMEDYCGARLTHNAIRIGGVPLDL
PPNWLEGLKKFLGEMRECKLIQGLLDKNRIWRMRENVGVVTPKMAQSWGMSGIMLRGT
GIAYDIRKEEPELYKELDFDVPVGNYGDSYDRYCLYMLEIDESIRIEQLIPMYAKTDT
PIMAQNPHYISAPKEDIMTQNYALMQHFVLAQGMRPPIGEVYAPTESPKGELGFFIHSE
GEPYPHRLKIRAPSFYHIGALSDILVGQYLADAVTVIGSTNAVFGVDR

>PROKKA_00121 hypothetical protein

MKRFDLRPLKAGIFERLEELIEKEMQPNEVAIFMFEVGDGFSNIPKSAEFIQSKGHELLNS
LRFNQADWTIVVRKKA

>PROKKA_00122 hypothetical protein

MILNGFNPLNSPLIASSSLSLKEAYYLEKLSLQKGFKIHYKMTKDSLLEKSDLCVLF
GFSNACLNENERLILENINQLKRPYALLRPLQDTRDLQENCLFASYEIHTEAAILALIL
GILEKTSQKLGHVLEKVDVGYLSSEANMSEELQELIALIIKAKKRALVLNREITKHADS
AFLYTLLSELQNYLEILHIPCNDSNATTAFFYDFKDQEWLLETAFKEGILPFESQLKSKDL
ELLERMGEANGSFVYVSYKSLKTPKLSFSKQFKIANKIEHSAKAGFQISNKTLECELEESP
HLKGLIAILEGAFFDAYPIPIILSHSQGIS

>PROKKA_00123 NADP-reducing hydrogenase subunit HndC

MITMNINGKTIECQEGQSVLEVARSAIYIPTICYLSGCSPTVACKMCMVEMDGKRIYSC
NTKAKNNATILTNPTLMDERKSIMQTYDVNHPLECGVCDKSGECELDQDMTHLTGVEHQ
YAVADDFKALDSWAKALYDPNLCIMCERCVTTCNDVGENNLKATKADLHAPDKFKDSMS
KDAFSVWSRKQKGIISFVGSVPCYDCGECIAVCPVGALSYKDFAYTANAWELKKIHSTCS
HCSAGCLISYDVRHFDTLGEESKIFRVLNDFYHNPICGAGRFAFDVSSSPKGSTNLKEAQ
NALKECEAVRIGGDITNEEAFLIERLRKELDFKIYNQEVYHFQQFLKVLGEVSRPNIEEI
KTSNLVVITIGSSIKTENPLVRYAINNALKLNKASLIAMHPIKDNALANLCRSSFCITHEV
GAEEILLGMLLKMLNIESTALKSLEDSKQSIQVDEAALKALEEERKKALEQANQGCSEES
KAENQEEEDKIEAAAPKEENPEENKTEVKEEKIEVPTKTTYLLLEEAGINLETYEKILALL
QKSNNTLLVVGEEIYSHKQAHNIAKMLRLLAQSAVKLILIPPSTNALGSIQCLSEEV
FEHEKIVGIRAQGDFTINSDDRFGKDALSKVDFILPSLNQLEGTITNVEGRVLPKLPAL
RFEGYDLSDIMQGGFVEENLTECTHKLPTAGFKAIEFDYLTNYFTNDRVNRHGYLLGT
SHFEKSAKECETIECEPIKPLKEKIAFNAYLKYPETQFNATNKSLENLQKAGVYVSKAF
LKKLNKEVGQNTITLFKEEEELTGVLVYLDSESLDQEVFVISPSLLTNHNSNFFREGVFDSDV
KEQA

>PROKKA_00124 NADH-quinone oxidoreductase subunit 8

MSAYIIETLIKILILVAVFSALGGFATYIERKVLAYFQRRLGPCYVGPFGLLQVAADGIK
LFTKEDIIPQGANKFIFTLAPIIAMVSAFVSMAPIPFFPNFTLFGYEIKPLISDINIGFL
FFLAVGSAGIYAPILAGLASNNKYSLIGSARATIQLLSFEVVSTLTILAPLMVVGSLSLV
EINHYQSGGFLDWLVFKQPLAFVFLFIASAEINRTPFDLLEHEAEIVAGYCTEYSGLKW
GMFFLAEYAHLFASFVISIVFFGGFNAGWFIPGGIAILIKAGFFVFLSMWVRATYPHVR
PDQLMNMCKIMLPLALLNIVLTGIIILI

>PROKKA_00125 NADH-quinone oxidoreductase subunit I

MAKQYEQQLPKRAEAHSATEQFKDTIKTSLGLDLFKGLGLTIKEFFSPSVTIHYPMEQLP
LSPRYRAVHNLRLLDSGSERCIGCGLCEKICTSNCIRIITHKGEDNRKKIDSYTINLGR
CIYCGLCAEVCPELAIVMGNRFENASTQRSQYGSKEFLTSEQDAKNCSHAEFLGFGAVS
PNYNERMQATPLDYVQEPSKEESQEETPTNPESHKGDENV

>PROKKA_00126 NADH-quinone oxidoreductase subunit J

MFETIAFYFFAILTSMALVVITTTNIIYAITALASSMVFISAFFFLDAAEFLGVVQITV
YVGAVIVMYAFGMMFFNSAAEVIERKQSPKILCILSFGVAVLLTLILSAPSIGENLSNQV
NSNAIDAQIPNIKAIGYVLFNTYILPFEEAALMLLVAMVGGIATGIQKIHGKNHTQFIKE
SL

>PROKKA_00127 NADH-quinone oxidoreductase subunit K

MIGLNHYLIVSGLLFCIGLAGMLKRKNILLFFSTEIMLNAINIGFVAISKYTHNLDGQM
FALFIIAIAASEVAIGLGLVILWFKKFKSLDIDSLNAMKG

>PROKKA_00128 NADH-quinone oxidoreductase subunit L

MQYSSLLSVVLFPLIGAVYAGLFGAKAKALHVGVFNSLCVLVSFIGAVVLFIQAWHHQS

YEKYLFDWIVVGNFKVGFSMLMDNINAVMIVVTLVSVFLVHVYSIGYMEHDKGFNRYFSY
LSGFVFSMLVLVLSDNFLGLFIGWEGVGLCSYLLIGFWYHKTSANNASIEAFVMNRITDL
GMLMGIILIFWNFGTLQYKEVFSMLNNADYPMLFYISVFLFIGAMGKSAQFPMHTWLANA
MEGPTPVSAIHAATMVTAGVYLVRANPLYSAVFEVGYFIACLGAFVALFGASMALVNK
DLKRIVAYSTLSQLGYMFVAAGLGAYAIALFHLFTHAFFKSLFLGSGNVMHAMEDNLDI
TKMGALYKPMRITAVFMIIGSVALCGIYPFAGYFSKDKILEVAFGMHHHILWFVLLIGAI
FTAFYSFRLIMLVFFAPKQHEINHPHEAKNFMLLSMLPLGLVLAIVAGFFEEPFFHFVSQV
IPNVGEYLVPLALLISITTIVVLLSIAYAIFKYKNGITSKKEGGFLYKLLNQQYYIPQLY
QGIKVFSAIASFLHQVVELKIIDAIVDAIGRSVVFVIGRVFRISQDGNLTSMLRFMVAGV
LILLAFVAFFGR

>PROKKA_00129 NAD(P)H-quinone oxidoreductase chain 4 1

MQFLHAHLLSVVIFFPMLSALLAFFMSDQASRAYAIVIALIELLLILLWHGFDIQTAGM
QFEEMKELVYQIGVNYHVGVDGIAFFLLLLNAIVVLLSVIYVKEHRKDFAIACLLLEGIL
MGVFSSLNMIFFYAFWEISLLPVLYLIGRFRNNKIYSGMKFFLYTFLASLCMLLGILYI
GYYYASNYGMMSFDILDWYQLNFSSGVKTWLFVAFLLIGIAVKIPLFPLHTWLPYAYSNA
TLGSVMLSALLSKMGTYALLRFLPLFPELSEIYLTPIAIVALCMIYGGFLAYAQKDLK
TLIAYSSFSHMGVVVLGVFSFNVEGISGAVFMMFAHGIIVMGLFLLAGILEERASSLEIA
RFGSIAKSAPVFAAFMIVLMANVGMPLSIGFVGEFLSLLGFFATYPLLAIIAGTSIILS
AVYMLTSYKDVFFGNLKAGNNQISVFEDLNAREIGVLSVILALILILGIYPKALLKPIEQ
GSKQLLEVIEIRSLPFLGSLDTKIKEVSYVNR

>PROKKA_00130 NADH-quinone oxidoreductase subunit N

MLIDSLHISFDSFNESVLPMLVLVCGGIFTLLINAFTSRFSRNLNVFLCMLFLVLDLFLV
VLGLEEQENAFFGFLSLDTLSLISQSIVLISAFLLIFLALSKERFNEFQTAEFYSLYLF
AAGFQFMVSSNHLILLIGLETASPLCVLMALSDKRYGLEAGIKYFTMGAMASAFFAMG
AMAFYLLTGSLNLEVITLYLHTEGVTNPMLFAMGAIFLIGAIGFKVSLVPFHTWMPDVYE
GNNPVFASYISIVPKIAGFVVATRLFGAFIDTRIAWVEDIFYVLILMTITIPNFIALWQE
DVKRMLAYSSISHSGFALACVFIHTEDSQQAMFVYWFMAFTYIGAFGLLWLLKSREKTW
DERYDHPYSKFNGLIKTHPLVAILGAIFVFGLAGIPFVSFVGKFLAVESALESNHILLA
VVMLVNSAVAAFYFRWLVAFFNKPLQSYAQNDIYTQATMPIYAVIIAMALACLSV
MMRGLLEFVA

>PROKKA_00131 hypothetical protein

MGLLSHSLNALSLTLTQGKEGGEDFSVLTLRHNKAFSCFYTNEKPPSGIEASLSIIRAKR
PIECVIDSIPKEGFTPLENAFFNITYSMRQQQFILHIKPKVMRRLTLFSFDRDYKKAIP
FVENDLKAKMWQIIGYDQNIPLSGKDNAQKGLNFPIVIKDAQTPIIQELDVNNKPLLT
KGYDLNAYLEAKQIDSQAYFDALRTISRAFKNYPQTIFKKDLYLLEIIALGQLGIKKS
LIDIGTKWIKNYPTDPNIPEALYVAKALDENNNYKQAMRYYKRILLEYKNSRYAPLAQM
RLAIEAAEGSDLSNANMLFKEAFSNAKDKESEIALNWAEAEINYQNFNNAKYLIDKVV
QSNPDYISTHSESALDLLKLLKKNQMNASAEIAHLLFNQDDDLKAKEQALYDLGALYAR
IKDFKNAHLYNLQYLQDHAELDKVSVMRDEKALFSMEGNTQEKIAHYDKIIQNFPSN
EALKALELKAQLLFENKRYAEVLGMQKNLPKDSPLIQKTLNIIAKTPLEDNRCEEALKYL
SQITAFEFTPKEEIQAFDCLYFASLKEKAQIIALNALKAAKTPSEKLVWLYRLGRNYYRL
GDFKNSTLASKDALILAQSLNKKEFYDIAFVLFVLDYMQSNEKGLALNLYAFLEKHFKDDK
RMALVYFKLLENEKDPKSVKIYATSLKLQDAYKDYSYTPFSEFALIDAYRTTKDYKAL
ETLDKLLDRRLSLEDHQKALYLQSSLLDLTNQAKSKSSLEKCVQLKQKQDNQNLCE
QGLNLFKNKES

>PROKKA_00132 Phosphomannomutase/phosphoglucomutase

MDISIFREYDIRGIYPTTLDEKSAFSIGVELGKIMREYDKSVFVGHDARVHGRSLFEALS
AGLQSSGLKVYDLGLIPTPVAYFAAFNGINGVQCPNSIMITGSHNPKEYNGFKITLNQNP
FYGKDIQALKDTLLNAKHEIKPLKETPEKVNALAEAYQRYLIKDFKHLKLNLYKIALDFGN
GVGALGLEPILKALNIDFSSLYSDPDGNFPNHHDPSEAKNLKDLEKHMQENASIGFAF
DGDADRIAMLSSHIIYAGDELAIFAKYLHAQGITPFVIGEVKCSQVMYNTINTFGKTL
YKTGHSNLKIKLKETNAHFAAEMSGHIFFKERYFGYDDALYACLRALELLEQSPSDLEN
IIKNLPYSYTTPEEKIYVSEEEKFEIIHNLQKALKNPPSHFPKIKEIISIDGVRVVFHEH
FGLIRASNTTPYLVSRFEGKDETTALEYKRALLNLEK

>PROKKA_00133 hypothetical protein

MDNNKNNNNKNASHLTHEMKNHGDLOTASMKALITQKVDELNDIRKKQLAEKEIQAEIEE
ESRNKERESFKRERKSLEEEKLMEKGFQVEREILEICYNNGYLPQKNPKIEHDDTIDI
EPED

>PROKKA_00134 Tryptophan synthase alpha chain
MRYQNMFETLKKQDKMAFIPFVTLGDPNYEWSFEIIKTLIDSGVNALELGFAFSDPVADG
VIIQASHLRALKHASMAKNFQLLKKIRSYNHDIPIGLLTYANLIFS YGVDNFYAQIKECG
VDSVLIADMPLEKELVIKSAQKHHIKQIFIASPNASDKDLEQAAIHSQGYIYTLARSGV
TGISHALENNASAIIKTLKAFSPTPALLGFGISQKEHIKNAKMGADGVICGSALVKIIE
ENLNNENAMLEKIKGFIGEMIF
>PROKKA_00135 Tryptophan synthase beta chain
MNKKAYFGEFGGSFVSELLVPALRELEQAFDACLKDEKFQKEYFRLLKDFVGRPSPLTLC
QNIVSNPKVKLYLKREDLIHGGAHKTNQALGQALLAKKMGKTRIIAETGAGQHGVATAIA
CALLGLKCVIFMGSKDIKRQEMNVFRMRLLGAEVREVN SSGSATLKDAVNEALRDWASSYK
DTHYLLGTAAGPHPYPTMVKTFQKMIGDEVKSQILEKENRLPDYVIACVGGGSNAIGIFS
AFLNDKEVKLIGVEPAGLGLETNKHGATLNKGRVGILHGNKTYLLQDDEGQIAESHSISA
GLDYPGVGPEHSYLKEIGRAVYESASDAEAEAFRLLCQKEGIIPALESSHALAYALKLA
QKCEESIIVVNLSGRGDKDLSTVYNALKGGLK
>PROKKA_00136 Tryptophan biosynthesis protein TrpCF
MPSVLENILKDKLLEVSMKKNHTLPINITPSDRDFKKALLEKKT SFILECKKASPSKGL
IRKDFDLLKIAKTYEKFASCISVLADSKHFLGSYENIKIVSQHSTKPILCKDFIIDAFQI
KLARVMGANAVLLMLSVLDDKNYLELFNLAKSLNMSVLTEVS NQQEIERLLKLQYDIIGI
NNRDLHTLTDDINNTLKLRSLLPKDALVVSESGIYSHVQIKALAPYVNGFLVGSSLMKEK
DLKKACAKLILGENKVCGLTRIKDAKAVYKNHFIYGG LIFEKSSPRYIKPKEALKITKAV
KKLDFVGVFVKDSIKKIAKIAKKLDLKA VQLYGYSPKKIAQLKKS LPKTCTIWQVVNAMD
ANDLAPKIKEASLILYDTKGDKTGGNGVSFDWGI LENVKTPFMLAGGLNLDNIQKALKIK
ALGLDFNSGLEISPGIKNKDKIKRLARILREY
>PROKKA_00137 Bifunctional protein TrpGD
MKDILNALYHQKDLNDEEVKKLFTLIINEKVSPAQLGAILCAL KIKGESFKEISVAATTL
LEHAPKPFNSGLDLIDNCGTGGDGLKTINISTIAALIASSM GLPMAKHGSRVSSHSGSA
DLLENLGVNIEMNPMQLENCFKQSHFGFLFAPLYHQSF KKSAPLRKELFTKTIFNCLGPL
INPLRPKIQLLGVDKSLCKTMALALKALGVKRAMVVNGG GTDEIVLHDITHACELKNE
ILEYDLSAKDFDLPPYDLKELQIENAKESAQA CLDILENKGKNSHTMVVAANVASLLYLS
HRAKDLKEGVGMTLEHLKTKAPYAHLQKIIWL SHA
>PROKKA_00138 Anthranilate synthase component 2
MKIFFIDNFD SFSYNLVYELECLGYEVAVYQNDIEPDY LMDLMKKESKTPLLFISPGPGN
PNSSGNLLKIIEMAKKKFPILGVCLGLQALAQSYGAKI IRSKEIVHGKATTIALKKHAVF
KGLGESMVVGRYHSLMASGLPKNLEVIAEHDNIPMAIVNEEDKILAYQFHPESIMTLQGR
ALLEQSVGFLKEL
>PROKKA_00139 Anthranilate synthase component 1
MVSLIEKAPYIPYPLALYEQLEQPHTL LFESAEIESKAHTKSLLMAKACLK LICNHNIVT
ITSLTPNGGAFLQKLSVFFKTPIQDNALTLYTKNKTQDE FLKLFEPSPF DALRGLFKS
VKTKPKHPFTLLSAGVFSFEMLNFFEDLPNLKAQDN TAHD FIFYVAQNLI IIDHKEKSAE
ILGACFDERFKTEIAQELQDLKELAKNIKSDFIPK KAKQSREVS VSCDDSEFEKRVLSLQ
EEIKKGEIFQAVLSRSFYMECLEGLSAYYHLKLSNP SPYMFYIKDSDFILFGAS PESALK
YNALTNTAEIYPIAGTRLRGKDKQGNIDYDLDSKMEFD LQHDYKERA EHI MLVDLARNDM
ARVSKKRYCDKLLKVDKYSNMHLVSRVVGELKKGCD SLHAYRSFMNAGT LSGAPKISAI
KLIYQLEKQRRGSYGGSVGYLNSESSMDCITIRSCFVKNNR ALIQAGAGIVLDSVPQNE
ANETRAKAQALIDAIRKTSL
>PROKKA_00140 hypothetical protein
MRKLEILIRNAKNHFNQIQLLSVVKVVC AVLVVVIFLVCKKD WVR
>PROKKA_00141 Glycosyltransferase family 9 (heptosyltransferase)
MDFVGFEDLKCKDKENSQKV FVIRNDKLGDFILAIPALIALKQAFLEKGKEVYLG VVVPS
YTTPIALEFPFIDEV IIEDNHL SATLSKSIDALIFLFSNFKNARLAFSLRKFIPIYILAP
KTKIYSWL YQKSVRQSRSLCKTEYEYNLDLIHV FCKDHNLPNASIKKI AWWKLKDKSKER
SIIASKLNADVGLLWIGVHMHSGSSPVLPA SHFIKLIDCLHNNLSCEIILICGP GERKA
TEELLKKVPFARLYDTSHSLVDLAKLCANLSVYIGNASG PLHVNALFDNQSIGFYPNELS
ASIVRWRPFNERFLGITPPNGSNDMGLIDIKKESENIIEFIAPNL S CTQER
>PROKKA_00142 Lipoprotein E precursor
MSVLNAKECVSPLTRS VKYHQQSAEIRALQLQSYKMAKMVLDNN LKLVKD KPAVILDLD
ETVLNTFDYAGYLT KHC IKYTPETWDKFEKEGSLSLIPGALDFLEYANSKGVKIFYISNR
TQKNKAFTLTKLSFKLPQVSEESVLLKEKGKPKAVRREL VAKDYEIVLQVGD TLHDFDA

LFAKDAKNSQEQRAKVLQNAQKFGTEWIIIPNSLYGTWEDEPIKAWQNKK
 >PROKKA_00143 hypothetical protein
 MKKALISTLFGVSLVFAKPYTIDKANSSVWFVVKHFKFNETRGAFDNFDGKIDADPNTKA
 LNVFEGKIDIKSLNTRNKKRDDHLKTAEFFDVVKYPKGSFKMTKYEDGKIHGDLTLHGVT
 KPVVLEAKIQAPLQNPMMKKEFMVLQAEGKINRKDFGIGKTFNDAVVGDEVKIELKLEAY
 AQ
 >PROKKA_00144 Thiaminase-2
 MQVSQYLYQNAQSIWGDCISHPFVQIGRGTLERDKFRFYIIQDYLFLEAYAKVFALGVV
 KACDEAVMREFSNAIQDILNNEMSIHNHYVRELQITPKELQNARPTLANKSYTSYMLAEG
 IKGSIKEVAAAVLACAWSYLVIAQNLSQIPNALEHAFYGHWIKGYSSKEFQACVNNINL
 LDSLTASSKQEIIEKLKEIFITTSEYEYLFWDMAQYQS
 >PROKKA_00145 hypothetical protein
 MSGRLTFSCVVVLCGAMANVAIASPKIEARGELGRFVGIGFGAMGGAMGGAIGAPGGPV
 GSAIGRSFGHEIGREFGREIGDRVEDYIRGVDRPQAPREPTYDRHFVYDR
 >PROKKA_00146 hypothetical protein
 MNVKNRLSDWEYQSAVALVYTICISINARIFYDIDGSASDLIFDPKNSYYMWLVGLIAAL
 LSNLLFDPRGRDCYKSFQVRYPRFLKAIFKARFFGMFYNAVLSRLRDFYVMLLTMPFIA
 AIHEISAYCGHPSTFLIEGLVMLGFLVCFGICSRCAKLGW
 >PROKKA_00147 Nicotinamide riboside transporter PnuC
 MLITTQLSKRFYATLILSCVFLTITNIIKVGSGFINLLAGLSGVLYAFFAGERQTICFVFG
 LVYNLSYAYVAYQWKLNADVILCLFLYMPVTIYGLFAWKKTEQHEGAIIKAQKLPKNWRFI
 LILGIGVLTYVSALFFKEIKTNFLWAESFNFVIFIIAFILQVLRYYENYALVTLGNIVSI
 IVWFCIFQISTESLVQLFTTILYLFIGLYYFNRWNKSCKQ
 >PROKKA_00148 Thiamin pyrophosphokinase, catalytic domain
 MQAVILANGEFPKSKQCLDLLKNAPFLIACDGAVISLHALQFKPSVVIGDLDSIDSHLKA
 LYNPIRVSEQNSNDLSKAFFYALNKGCDDEFIFLGLNGKREDHALANTFLLLEYFKFCQKI
 QSISDYGLFRVLETPFTLPSFKGEQISLFSLDLKAQFTSKNLKYPLKNLRLKTLFSGSLN
 EATDSYFSLSSSTPKSVVLVYQKFS
 >PROKKA_00149 50S ribosomal protein L17
 MRHKHGYRKLGRSTSSHRKALLKNLAIALIEHNKIETGIYKAKELRSYIEKLTTAARVGDF
 NAHRHVFAYLQNKETHKLVTETAPKYAQRNGGYTRIQRRTTFRRGDASTLATIEFV
 >PROKKA_00150 DNA-directed RNA polymerase subunit alpha
 MKVIKTAFLIPSEIKVLEKEGNRVKISLAPFEFGYAVTLAHPIRRLLLLSSVGYAPIGLK
 IEGVHHEFDSLGRGVTEDEVSLFIMNLKNIRFIKALVGQDSSLENQSVVVDYSFKGPMELR
 ARDLNSEQIEIVNPEMPLATINEDAQLNFSLIYYKGMGYVPSENTRELMSEGYMPLDGSF
 TPIKKVVYEIENVLVEGDPNYEKIIFDIETDGQIDPYKAFLSAVKVMSKQLGVFGERPIA
 NTEYSGDYAQRDDAKDLSAKIESMNLARSFCNCLDKIGIKYVGELVLMSEELKGVKNMG
 KKSIDEIAEKLNLDLGYPVGTLSHEQRESLKKRLEKLEDKGGND
 >PROKKA_00151 30S ribosomal protein S4
 MARYRGAVERLERFRFVSLALKGERRLSGKSALDKRAYGPGQHGORRAKTSYGLQLKEK
 QKAKMMYGISEKQFRSIFVEANRLDNGTGENLIRLIERRLDNVVYRMGFATTRSSARQLV
 THGHVLVDGKRLDIPSYFVRSGQKIEIKEKTKSNSQVVRAMELTAQTGIVPWIDVEKDKK
 YGIFTRYPEREEVVPIEERLIVELYSK
 >PROKKA_00152 30S ribosomal protein S11
 MAKRNVTAKKKVVKKNIARGVYISATFNNTNITITDEMGNVICWSTAGGLGFKGSKKSX
 PYAAQQAVESALSKAKEHGVKEVGIKVQGPGRSGRETAIKSVGATEGVKVLWIKDITPLSH
 NGCRPPKRRRV
 >PROKKA_00153 30S ribosomal protein S13
 MARIAGVDLPKKRVEYALTYIYGIGLKSSREILEAVGISFDKRVHELSEDEVSSIAKKI
 QQSYLVEGDLRKKVQMDIKSLMDLGNRYRGIRHRKGLPVGRQTTKNNARTRKGKKKTGVSX
 >PROKKA_00154 50S ribosomal protein L36
 MKVRPSVKKMCDKCKIIKRRGVIRVICATPKHKQRQG
 >PROKKA_00155 Translation initiation factor IF-1
 MARDDVIEVDGKVIKALPNATFKVELDNKHVVLCRISGKMRMHYIRIALGDRVKLELTPY
 SLDKGRITFRYK
 >PROKKA_00156 Methionine aminopeptidase 1
 MAISIKSPKEIKALRKAGELTAQALALLEREVRPGVSLLLELDKMAEDFIKSSHARPAFKG
 LYGFNPSVCMNLNEVVIHGIPTDYVLQEGDIIGLDLGEVDGYYGDSALTLPIGAISPQD

EKLLACSQESLMHAIN SIRVGMHFKEL SQILEGAIVERGFVPLKGF CGHGIGKKPHEEPE
IPNYLEKGVKANS GP KIKEGMVFCLEPMVCQKQGE PKILADKWSVVSVDGLNTSHHEHTI
AIVGNKAVILTER

>PROKKA_00157 preprotein translocase subunit SecY
MNKAIAS KILITL GFLFLYRVLAYIPIPGVDLAAIKAFFDSNSNNALGLFNMFSGNAVSR
LSIISLGIMPYITSSIIMELLSATFPNLAKMKKERDGMQKYMQIVRYLTILITLIQAVSV
SVGLRSISGGANGAIMIDMQVFMIVSAFSMLTGTMLLMWIGE QITQRGVNGISLIIFAG
IVSGIPSAISGTFNLVNTGVINILMLIGIVLIVLATIFAI IYVELAERRIPISYARKVVM
QNQNKRIMNYIPIKLNLSGVIPPIFASALLVFPSTILQQATS NKT LQAVADFLSPQGYAY
NILMFLLIIF FAYFYSSIVFNSKDIADNLRRNGGYIPGLRPGE GTSSFLNSVASKLTLWG
SLYLALISTVPWILVKAMGVPFYFGGTAVLIVVQVAIDTMKKIEAQIYMSKYKTL SAVGF
>PROKKA_00158 50S ribosomal protein L15
MGLENLKPAGKSVKKIKRVGRGQSGMGKTATRGKGQTARTGYKAKRGFEGGQQPLQRR
LPKIGFRTKDSHIYSINVEKNEAIKNLEEITFSSLRALHHPFLYIEGVKLIGKDAKNLAS
KIKDERIKTSGQK

>PROKKA_00159 30S ribosomal protein S5
MEEINREEFQEVVNIGRVTKVVKGGRRFRFNALVVGNKNGLVGFG LGKAKEVPDAIKK
AVDDAFKNLIHVTIKGTTIAHDIEHKYNASRILLKPASEGTGVIAGGSTRPIVELAGIKD
ILTKSLGSNNPYNVHATFDALAKIKA

>PROKKA_00160 50S ribosomal protein L18
MNAKALYKKKALRRRKLRIKSKLLGDALRPRVSVFRSNRYFYAQAI DDVKQNTITHIDG
RKMFGKNTQEDAKKLGVLF AEELKKAGIERAVYDRNGYLYHG VVA AFAESLRENGIAL

>PROKKA_00161 50S ribosomal protein L6
MSRIGKR IIEIPSSVQASVEGSKLLFKNSKEKHELETHNRVKVTLENNQLSFQPVGEDAQ
SRAYWGTYGALANNIVIGLSTGFSKTLEVNGVGYKVALGNKTL DLSLGFSHPVKYPI PAG
IEMVVEKNTITIKGSDKQKVGQVAAEIRSFRPPEPYKGKGVKYSNEVIIRKAGKTAKK

>PROKKA_00162 30S ribosomal protein S8
MVNDI IADSLTRLRNASMRRL EFTQLY YAKIVVSILEIFKEKGFIKDFNVKDKDKKQSVY
VQLAYDEKGHSKISEVKRLSKPGRRVYKQKNELKRFKNGYGVIVVSTSKGVITNEEAYRQ
NVGGEVLC SIW

>PROKKA_00163 50S ribosomal protein L5
MFG LKQFYQNEVRTKLAQELDIKNPMLLPKLEKIVISVGAGAYAKDMKIMQNI AQTISLI
AGQKAVITKAKKSVAGFKIREGMAVGAKVTLRNKRMYNFLEKLIVISLPRVKDFRGISR N
GFDGRGNYTFGINEQLIFPEVVYDDIMVSHGMNITMVTSTNNDKEAFKLELLGLPFAKV
R

>PROKKA_00164 50S ribosomal protein L24
MKSEIKKNDMVKVIAGDDKGKVAKVLAVLPKTSQVVVEGCKVVKKA I KPTDDNP KGGFIH
KEKPMHISNVKKA

>PROKKA_00165 50S ribosomal protein L14
MIQSFTRLNVADNSGAKEIMCIKVLGGSHKRYASVGSVIVASV KKAIPNGKV KRGQVVKA
VVVRTKKEIQRKNGSLVRFDNAAVILDAKKDPVGTRIFGPVSREVR YANFMKIISLAPE
VV

>PROKKA_00166 30S ribosomal protein S17
MNTKEPHKRLVQGKVISKFAEKSAVILVERKV VHEKYRKIVKKFKKYTIHDENNQVKVGD
FVSAIECRPLSKTKSFTLKEILVVG V

>PROKKA_00167 50S ribosomal protein L29
MKYTELKDKS I KELEELLHAKKAELFELRVKLKAMQLSNPNEIKKARRNIARINTAINAH
YSSSVE

>PROKKA_00168 50S ribosomal protein L16
MLMPKRTKYRKQMKGRNRGKAHRGNSIAFGDIAIKAIEHGRIDSRQIESARVAMTRHIKR
AGKVWIRVFPDKPLTAKPLETRMGKGKGSVEKWVMNIKPGRIVYEMLGIEEGLAREALAL
AQSKLPFKTKIVTCESENEIY

>PROKKA_00169 30S ribosomal protein S3
MGQKVN PVGLRLGINRNWTSRWFPSARTAPSNIDEDNKIRKFLKKELY YAGVSEIVIERA
AKKL RVTVVAARPGLIIGKKGV DIEKVKEGLKALIKKEVSINIKEVKRPQADAQLAAENV
ATQLEKRVAFRRAMKKVMQAALKSGAGIKVRVSGRLAGAEIARTEWYMEGRVPLHTLRA
KIDYGFAEAMTVYGIIGVKVWIFKGEVLQKGIQFEKKEEAKEEREPRRSRRGRQ

>PROKKA_00170 50S ribosomal protein L22

MSKALLKFVRLSPTKARLIARQIQGMNAELAIASLEFTPNKAARVLSKVVASAVANGSLD
AKSALIVSCRVDAGPVLRRSIPRAKGRATAIRKPTSHVFVEVVEGKEMKSSKSHKKNQAE
GK
>PROKKA_00171 30S ribosomal protein S19
MSRSIKKGPFIDDHLMKKTLKAKEGKDNRPKIKTWSRRSTILPEMIGFTYNVHNGRVFVPV
YITENHVGKYLGEFAPTRTFKGHGKGSVQKKIGK
>PROKKA_00172 50S ribosomal protein L2
MAIKTYKPYTPSRRFMSVLDSKDITAKSSVKGLLTKLKATAGRNNNGRITSRHKERGAKK
LYRIIDFKRNKYNIEGKVAIEYDPYRNARIALVVYPDGDKRYILQPSGLKVGDSVIAAE
GGLDIKVGFMKLKNPIPIGTVVHNIEMHPGAGGQLARSAGMSAQIMGRENKYTILRMPSS
EMRYILSECMASVGVGNEDFINVSIKAGRNRRHRGIRPQTRGSAMNPVDHPPHGGGEGKT
GTSGHPVSPWGTPAKGYKTRKKKASDKLIISRKHK
>PROKKA_00173 50S ribosomal protein L23
MADIMDIKSILYTEKSLGLQEGVLVVQTAQNVTKNQLKEVFKEYTFGFEPLKINSLKQEG
KVKRFRGKLGQRKSFKKFYVKVPEGASIAALGA
>PROKKA_00174 50S ribosomal protein L4
MSKAIVLDSHLKEKGSVELPKRYEGINSHNLVLYVKHYLSSVRANTAKSKNRAEVSGGGR
KPWAQKGGGRARAGSITSPVFVGGGVSHGATNNRNYNLKINKKQKRLALEYALEEKAQAN
KLFVVEKIAIKGVVEDNKRKHLTKEANQMFQALEQRDTLFCMNMDEYTELAFSNLKKCL
IVDVSELNAYLLAAFSSVVMEEAAAFQHVVQDKTEE
>PROKKA_00175 50S ribosomal protein L3
MEFLVQKIGMSRTIDANSTPVTLKVLQAKVCQLENGKALVAYAMHKKHNKAIEGQQKKY
QLSKEFNHFATLKASQQKELGDLDLSALETCLKRVKASFCTKGRGFAGVMKRWNFQGGPAA
HGSRFHRRPGSIGNREWPGRVQKGRKMAGHYGNELVTCQNEVLVSFDKESMVLVLKGSVAG
FSGAYGRIRAV
>PROKKA_00176 30S ribosomal protein S10
MEKIRLKLKAYDHRVLDRSVVAIVEAVKRSRGSEIRGPIPLPTKNKRYTVLRSPHVNKDSR
EQFEIRVYSRLIDIISATPETVDSLMLDLAPEVDVEVTSMETK
>PROKKA_00177 hypothetical protein
MPLSCLHPFGPFETPKPALNNPLLKAHSSDKICLLGPMKSGKTTFALKLAKDFKNPVYI
NYNDMRLNQNILSSWLLKWHLEKKMDLLILDRIERLDFSLPKLPKIVLIPNYLSPITALD
FSLCYALGLNFKEYISFFKPNTPKNTLFRNRLRDNALDSLFTENEQEKILKKQENIQLI
FQAYAPLMAKICSYQSKFVSAFYLYTQLKKELKISKDTLYKLLHALEKQRILFLAPSFEN
NKTKLYLCDFALPYSLTSPSLLSVFENMVFLELYKQFPNHELISHDNGIFILQNPTNKL
ALIAHAFTPHFLEKQLLWCHKHGGFFNIVVVSINAPISATNPYPYKHLNFIDFSLDIQSIL
I
>PROKKA_00178 hypothetical protein
MNALKKLSFCALLSLGLFAQTAHAKHLKGTINYPDWLEINLFDEKNPPNQYVGSASISGK
RNDFYANYIPYDDQLPPEKNAEEIALLRTRMNAYSSLESVLLTKMHKSIVKVLQVKNNVT
NHLFGLVDFLTSSKILAKRFVDTTNRHVYVMVQFPFIQPEDLIVYFKKKRIDLSLTSATH
LSAVLNKALFHL
>PROKKA_00179 Ribonuclease HII
MVLGIDEAGRGCLAGSLFVAGVACDEKTALEFLKMGLKDSKKLSPKKRFFLEDKIKTHGE
VGFFVVKKSANEIDSLGLGACLRLSIEEILENGCSLANEIKIDGNTAFGLNKRYPHIQTI
IKGDEKIAQIAMASVLAKASKDREMRQLHALFKEYGWDKNCGYGTKQHIEAIIKLGATPF
HRSFTLKNRIFNPKLLEVEQRLI
>PROKKA_00180 hypothetical protein
MNPEKATHSKGIRKILKEALHDNKDLNLYLESGSKHAKLTSGANHLTIPSSPSDRKSAKN
FEKELTEFIKKIRGNRITHGACE
>PROKKA_00181 Fumarate hydratase class II
MQFRIEHDTMGEIQVDDSQYWGAQTQSRLENFKIGTEKMPKELIGAFAKLKRSLAVVNHK
LGKLSPEKSQAIKACDCILKGELCGEFLAIWQTGSGTQTNMNLNEVIANKATEILGGN
FREKKLIHPNDVDNMSQSSNDTFPTAMHIVSVLEITHKLLPSLENLLKTFKDKSQQFKEI
VKIGRTHLQDATPLTLGQEFSGYASMLEHKSQQILESLEHLRELAIGGTAVGTGLNAHKE
LSQKVAEELSQFSGVKFISAPNKFHALTSHDAIAYAHGAFKALAANLMKIANDIRWLASG
PRCGLGELNIPENEPGSSIMPGKVNPTQCEAMTMVAVQVMGNDAIGIAASQGNFELNVF
KPVIIYNFLQSLRLSDSMESFNHCASGIEPNREKIDYYLHHSMLMLVTALNPHVGYENA
AKIAKNAHKKGISLKESALELKLKLSAEDFDKFVVPEKMIGPKA

>PROKKA_00182 hypothetical protein

MKKLAALFLVSVLGVMSLNAWEQTLKANDLEVKIKSVGNPIKGDNTFILSPTLKGKALEK
AIVRVQFMMPEMPGMPAMKEMAQVSEKNGLYEAKTNLSMNGTWQVRVDIKSKEGQVYRAK
TSLDL

>PROKKA_00183 Outer membrane efflux protein

MLSFVSAFYKRGVSIRLLSAFLLLFSLGLAKDLEIQSFVAKYLSKNQKIQALQEIDALS
SQEKVVSKWDNPVLYLGYNNANVSDFFRLDSTLMQNMSLGLSQKVDLNGKKLTQSQMIHL
EKQKKILELKKTKQQLVINLMINGIENYKNQQEIELLNTAIKNLENTLYQANHSSSPDLI
AIAKLEILKSQLEIKNNLEEALSSSHYSMGELTFKENELLSIAPKNFEFNHEQELYNIS
ATNYDIAIAKLDEEKAQKDIALAKKSFLEDVNVTVGVYFRSKQYYNYDMFSVALSIPLPL
YGKQAKLVEQKKKESLVFKSEVENTKNKTHHLAKLLKKLETLOKNLESINKIIKQNEKI
AQIYALDLKSNGDYNAYNAFNDKITTTQITQLETLSALNSAYLSLQNLKGLE

>PROKKA_00184 Cation efflux system protein CusB precursor

MKRLWLALTLFFSLSCANAQEIKETQETKKTKEAKSQTRFNISTTKVIEKEFSQSRCYY
ALLEPNEALIFSQTLRFDDGYVEKLYANKTYTPIKKGDRLLSVYSPELAGVQSELLSSLK
NQQVGAIKEKLLKLLGLENSSIEKIISSHKVQNEMTIYSHFNGVIFKKSPDLNEGSFIKKG
QELFQIIDLSQLWALVKVNQEDLEFLKNTHKAILFVEGVKGEQEITLENINPIINVQDKM
LEARFNVPNLKQLYYPNMFQAQVEIFQKPQKMKILPKEAVLIKGGKAIVFKKDDFGLSPLE
IKAVRLSDGSYEILEGLKAGEEVANNALFVLADADAQNNGDY

>PROKKA_00185 Cation efflux system protein CusA

MIEKIIDLSVKNKLLTTLVTLIFLASLWAIKSVRLDALPDLSPAQVVVQITYPNQSPKI
VQEQVTYPLVSTFMSIANIDTVRGISSYESGLIYIIFKDGVNLYWARDRVLEQLNRVNNL
PKDAKVEIGSDSTSIGWAYQYALSSDSKNLSDLKVLQDFYRYALLGVDGVSEVASVGGF
VKDYEVTLOQNDSLIRYNLSLEQVANAIAKNSNNDTGGGVILENGFEKIIIRSHGYIQSLKDL
EEIVVKKEGAIPKIKDIASVRLTPKPRGAANLNGDKEVVGGIVMVRYHADTYKVLKAI
KEKIATLQASNPDKITSVYDRSELIEKGIDNLIHTLIEESVIVLVIIAIFLLHFRSALV
VIITLPLSVCISFLLMRYFNIEASIMSLGGIAIAIGAMVDAAIMVMENAHKHLQHIDTKD
NAQRVNGIIEGVKHVGGAIFFALMIIVVSFLPIFALTGQEEKLFAPLAYTKTFAMLVGAL
LSITMVPILMVWLKGRILEESKNPINAFFIKIYGVSLNVVLKFRYAFLIASVLGLGGGLY
VAYKKLNWEFIPQINEGVIMYPVTINGVGIDTALEYLKKSNSAIKRLDFVKQVFGKVGR
ANTSTDAAGLSMIETYIELKPQNEWKEKLSYKEVRDKLEKTLQLKGLTNSWTYPPIRGRD
MLLTGIRTPGLGIKLYGNDTDKLQELAILMEQQKLTLESLSVFAERSNNGYYITLDLND
NLARYGINKSAVLDAIKFALGGATLTMTIKGVESYPISLRLEDTERNTIEKLNLYIKTA
YNYMPLRELVRIYYDNPAVLKSEKGLNVNFIYIVPQNGISSDTYRQLAQKALEKIQLP
GYYYEFSGESQYLEEAFKTLQYIVPVSVFIIIFILIVFALKNFTNSLLCFFTLPPFAFLGGL
IFMNIMGFNMSVAALVGFLALLGVASETAIVMIIYLEDAFQKFIKTPLEQTTTALKEAI
MHGAVLRVRPKLMTFFSILASLIPIMYSHGTGSEIMKSIAAPMLGGMISVVLTLFIIPT
AYFVIKSARIRKHEY

>PROKKA_00186 Branched-chain amino acid transport protein (AzlD)

MHSILIIILVIIITTYFTRIWPFMVFNSKNPPNDFVRYLGRALSCSVIGMLVVYCFKDIQI
LKPPYGINEITAFLSVILLHRIFKVFLSITLPTILYMLVLVQSHALEKAFFNIHVS

>PROKKA_00187 Inner membrane protein YgaZ

MHEFLKALKNAFPHTISIFLGYLLMGMTFGMLLVQHGYDYKVALFMSLFIYAGAVQFVAI
TLLSAQASLMNVVIVSLLVNARQTCYALSMLDRFKNTKWRLPYLAHALTDETFALLNLYA
PKEGVSEKDFIFSISLLNHSYWIFGSLVGSVLSHFSFDAQGMFVMTAIFIVLFMEQYK
RTTNHKNAWLGIFIAVVCLALFGTEYFLLIALVLMVLALMLFRKQLEC

>PROKKA_00188 Chaperone protein DnaJ

MELSYYEILEVEKHSNQETIKKSYRKLALKYHPDRNAGDKEAEEKFKLINEAYGVLSDEK
KRALYDRYGKKGLNQAGASQGDFFSDFEDLGSFFEDAFGFGARGSKRQKSSIAPDYLQTL
ELSFKEAVFGCKKTIKVQYQSVCESCDGTGAKDKALENCKQCSGQGVFMRQGFMSFAQT
CGACQGGKGIKVTCPQACKGKTYILKDEEIDAIIEGIDDQNRMLVKNKGNEYEKGRGD
LYLEARVKEDEHFKREGCDLFIKAPVFFTTIALGHTIKVPSLKGDELELKIIPRNARDRQT
FAFRNEGKHPPESSYRGLSIVELQVIYPKSLNKEQQELLEKLHASFGYEGEPHKSIVLETC
ISKIKDWFK

>PROKKA_00189 hypothetical protein

MMTFDSQTNAKLSRSNEQLSDMLYKLNESLRIYQSVLSNNQDQLKEIKKANSTLNSQRRF
FNASQIRLMDTDALLKQSALELEKLQALEKRLKESMEQERLIKESQTLFLQEHCPLYSGV
KNLEEASNALEVQEQNNALFLLKEPKLAHLLSQLDLMSALNALCDQVLESQAHNQQSHNK

TLEYNALKNHDFQAYKAMRLKKFKNKIQSQIQAKEDALKTFLEKRLLETLKTHFLCDKE
 NLKSCAKELHQRYQNALIERDKELKNAKDNKEKHALILANYEHTLKTNLIEFLSELNKQM
 AFLNETMALNAQVLALLAKQHAKTPKPSNLNGDLSDKKASIKNIRLDPHGFPSPFENFKQE
 >PROKKA_00190 tRNA-specific 2-thiouridylase MnmA
 MNGFCARLRAITHNERLKMKIAVLLSGVDSSYSAYSLKEQGHELVGIYKLHASEKKHD
 LYIKNAQKACEFLGIPLEVLDFQKDFKNAVYDEFISAYEEGQTPNPCALCNPLMKFGLAL
 DHALQLGCEKIATGHYARVKEIDKVSIIQEALDKTKDQSYFLYALEHEVIAKLVFPLGDL
 LKKDIKPLALSAMPFLGTLETYKESQEICFVEKSYIDTLKKHVEVEKEGVVKNLKGEEIG
 THKGYMQYTIGKRKGFNIGALEPHFVVGIDAKKNELVVGKKEDLAMRFLKAKNKSMLMKD
 FKSGEYFIKARYRSVPAKAFVSLKDEVIEVEFKEPFYGVAKGQALVVYKDDILLGGGVIV
 >PROKKA_00191 DnaJ domain protein
 MAKIELLAKFTQIALPNSHPLLKVLNYAKKHFSQCHMLSSSLILNDTECFKKNYLLNW
 VYHALECAHEKDISPYSLEEILQKSRLPIRIKIINQNTLLEKIEVKVLTFGAEYTLFITK
 HPIAKRFLRQKFSGYVFLETQDELHIRGDSERFWELIVTLNENRIVHNACLHFIYPNGFG
 KDSYTTLAERKLKECYKALGFIKHENFSEVKKRYLELAKTYHPDLCDLKEKKALYAKRFA
 IIQEAYSHIKKHA
 >PROKKA_00192 putative nicotinate-nucleotide adenylyltransferase
 MNSVLQYKELALYGGSFDPHKAHLAIIQDTELLPFAKLIVLPAYQNPFKKPCFLDAKT
 RFKELERALKGIDRVLLSDFEIKQERAVPTIESVLYFQKLYCPKTLYLIVIGADCLRHLSS
 WTNATELLKRVELVVFERIGYEEIQFKGRYFPLKGIDAPISSSAIRASLGV
 >PROKKA_00193 Putative nickel-responsive regulator
 MDTHNKDDSIIRFSVSLQQNLLDELDNRRIKNGYSSRSELVRDMIREKLVEDNWAEDNPN
 DESKIAVLVVIYDHHQRELNQRMIDIQHASGTHVLCTTHIHMDEHNCLETIILQGNSFEI
 QRLQLEIGGLRGVKFAKLTKASSFEYNE
 >PROKKA_00194 Biopolymer transport protein ExbB
 MGGFSMAMLKDYVDVFVFAVLGVASFALWFATERVIFYSKVNLIKAYDDIDALNLDLTKN
 LTILYVIYSNAPYVGLLGTVLGIMVIFYDMGVSGGMDAKTIMVGLSLALKATALGLAVAI
 PTLIAYNSLLRKSDVLSEKFRIMKK
 >PROKKA_00195 Biopolymer transport protein ExbD
 MKSIRRGDGLNVVPFIDIMLVLLAIVLSVSTFIAQGKIKVSLPNAKNAEKSQPNQKVVV
 ISVDEHDNIFVDDKPTNLEALS AVVKQTDPKTLIDLKSDKSSRFETFISIMDILKEHNHE
 NFSISTQAQ
 >PROKKA_00196 transport protein TonB
 MKISPSPRKLSKVSTSVSFLISFALYAIGFGYFLLREDAPEPLAQAGTTKVTMSLASINT
 NSKVKTNAESA KPEEPKEKPKKEKPKKEEPKKEVTKPKPKPKPKPKPKPKPEPKPKPKP
 EPKPKPKPEPKPEPKVEEPKKEEPKKEPKKEEAKKEEAKESAPKQVTTKDIVKEKDKQEE
 SNKTSEGATSEAQA YNPGVSNEFLMKIQTAISSKNRYPKMAQIRGIEGEVLVSFTINADG
 SVTDIKVVKSNNTDILNHAALAIKSAAYLFPKPDET VHLKIPIAYSLKED
 >PROKKA_00197 Helicobacter outer membrane protein
 MKIKKSLLLSLSLIASLSRAEDDGFYMSVGYQIGEAVQKVNTGALQNLADRYDNLSNLL
 NQYNLNLHLVNLASTPSAITGAIDNLSSSAINLTSATTTSPAYQAVALLNAAVGMWQVI
 AFGISC GPGPNLGPEHLENGGVRSFDNTPNYSYNTGSGTTTTTCNGASNVGPNGILSSSE
 YQVLNTAYQTIQTALNQNGGGMPALNSSKNMVVNINQTFTRNPTTEYTYPNGNGNYYS
 GSAIPIQLKISSVND AENLLQQAATII NVLTTQNPHVNGGGGAWGFGSGQTGTVM DIFGDS
 FNAINEMIKNAQTAL EKTQLNANENAQITQPDNFPNPTSEDKGFAQEMLNRANAQAEIL
 NLAQQVADNFHSIQGP IQQDLEECTAGSAGVINDNTYGS GCAFVKETLNSLEQHTAYYGN
 QVNQEKALAQ TILNFKEALSTLNKDSKAINSGISNLPNAKSLQNMTHSTQNPNSPEGLLT
 YSLDTNKYNQLQ NIRQELGKNPFRFRFGVIDTQSNGAMNGIGVQVGYKQFFGKKRNWGLR
 YYGFFDYNHAFIKSSFFNSASDVW TYGVGMDALYNFINDKATNFLGKNKLSVGLFGGFA
 LAGTSWLNSEFVN LNVVGNIYSAKINVANFQFLFNLGLRMNLARPKKTDSDHAAQHGVEL
 GVKIPTINTDYYSFMGAELKYRRLYSVYLN YVFAY
 >PROKKA_00198 Integral membrane protein TerC family protein
 MEFLSSLLDALSTPHGIVSLATLTLEIILGIDNII FITVMVYKLPKHQQNKAMILGLGL
 AMIARIGLLGSLFFISHLQKPLFTIAGMSFSWRDVVLFVGGAFALAFKALVELKEQIYPKE
 KHQEKAFGFFITLIEIMFLDIVFSLDSVITAIIGIAKHLEVMALAIILSVIVMMFFSKIVG
 DFIERHYRIKTLAFVFLLVGVFLFLEGLHLHVDKNYLYAGIGFALLIECLNIFIEKKMK
 KS
 >PROKKA_00199 Magnesium transport protein CorA

MVNVFFKQQKFVVIKKRFNDFNGFDIEENEVLWFELINPTPNELATLSQEYAIHYNTDHSQ
RVSSVTKYWEDSSSVTINAFFTNQDENETFHTEMATFILSNNILFTIYYGTLEIFDSIQK
KVLASPKKFEDGFDILTKEFEVYFEKGVECLEWINKQTSLLRKNIIFKETSTHDDILVRL
SNLQEFNVALRDSFFDKRRIITALLRSNKVSDTKNNLNIIILTDSSSLVESTTVNLNSLD
NIQNLFASQVNVEQNKIIKLFTVATMAMMPPTLIGTIYGMNFKFMPELEWQYGYLFALIV
MAISTILPVLYFKKKGWL

>PROKKA_00200 Phosphoglycerate kinase

MLAKMSFMQNVKNIQEVEVSHKRVLIRVDFNVPLDENLNITDDTRIRESLPTIQFCIDNK
AKDIIILVSHLGRPKGVEEKLSLKPFLKRLERLNHEVVFSQNIQQLKQALNENAPTRIFL
LENIRFLKGEEENDENLAKDLASLCDVFNDAFGTSHRKHASTYGTAKFAPIKVSGFLLK
KEIDSFYQAFNHPLRPLLLIVGGAKVSSKLTLLKNILDLDKLIAGAMSNTFLKALGYD
VQDSSVEDALIKDALELLQSAKEKKVKVYLPIDAVTTDDILNPKHIKISPVQDIEPKHKI
ADIGPASLKLFSVIESAPTILWNGPLGVHEKQEFARGTTFLAHKIADTYAFSLIGGGDT
IDAINRAGEKDNMSFISTGGGASLELLEGGKILPCFEVLDKRH

>PROKKA_00201 Glyceraldehyde-3-phosphate dehydrogenase 2

MPIRIAINGTGRIGLCAIRVASQRKDIEIVAINSTALETLLHLIRHDSVHGHFEAQLNA
DRTLNIHGSKNILVLSEKDINKLDFSAANAEIIIIECTGKFNSLEASSAHLKNSVKKVIIS
APAQNTPTFVYGVNHKNYHNESVISNASCTTNATAPLLKILDEAFKVENALLTTIHSYTN
DQNLLDTKHKDIRRARAASLNLIPTSTGVSKAISLVLPHLGPKVTGLAIRVPTPNVSLVD
LSLNFKKSVSKASVQHALKDACKHAFKGVVGIDEERLVSSDFISSPFAIVIDDQIMTIG
EKNAKVLAWYDNEMGYSERLIDMAQYIAQN

>PROKKA_00202 Uracil-DNA glycosylase

MKLFDYAPLSLAWREFLQSEFKKPYFLEIEKRYLEALKSPKTIFFPKSSNLFYALNLTPPS
AVKIIILLGQDPYHSTYLKKEQELPVAMGLSFSVEKNAPIPPSLKNIFKELHANLGVVPVC
CGDLSAWAKRGMILLNAILSVEKNQATSHKYIGWEAFSDQILMRLFETTAPLIVVLLGKV
AQKKIALIPKNKHIIITAPHSPSLSRGFLGSGVFTSVQKAYREVYRKDFDFSL

>PROKKA_00203 1-acyl-sn-glycerol-3-phosphate acyltransferase

MKSNSXXNSLRAIYRALVIAIGLAVIIVFNYFNKNNNARSSRRACSCFFSLTGVLHEKI
GTFDTDAKLIVLNHQSLDDIYLEAYHPSNICWIAKKELGEIPFYGHALTDGTGMILIDRE
DKKGIVSLLKACKEKLDQNRPLVIFPEGTRGKGGEKFLPFKQGAKIIAEKFQLKIQPMVL
INSIKIFNSKPLEAYKARTRLVMLESYTPDFNSPTWYEELEERMQKEYLKHYHELNA

>PROKKA_00204 hypothetical protein

MGFLFEKSLMSFFAHPIKILKIISLILSFLISFLVAENANEPEEIKAKVAYVKIPQLEDL
ENNPVYIGQIIGVTYDLLLFDAEFLEAKIKDGLDKTQIELLNKMPKWKKEKELFRATYY
YKIKGIKAIIPSLEVSFAFSNKDKYIDHSIAPKVILQVTDLSKNPRYANVMAKDLQVVQYK
TKDYDDKNNILVMEIAFKEATWEDFHIKEAIKQGFDNASLNQIKAKEGSVFYYCVLPKTI
QNLSFDYFSLSNRQFKTSLFSFAIPTQDTTGIQSDLPKNNFLVFSNVALLCVFFLVLF
FIFGRKLIFLGLGILCLGFVLYHLLFTQKSALLLAHKKIRILPTQNSTILGLSKNEMPIK
ILGSHDDYYKILTPHEQIGWVKKDEVK

>PROKKA_00205 Carboxy-terminal processing protease CtpB precursor

MTKRLFKGLLAVSLAVSLHGGEVKEKKPVKPKEDPQELAAKRVEAFSFRFSNVVTEIEKK
YVDKISISEIMTKAIEGLLSNLDHA SAYLNEKKFKEFQAQTEGEFGGLGITVGMRDGVL
VIAPLEGTPAYKAGVKSGDNILKINNESTLSMSIDDAINLMRGKPKTSIQITVVRKNEPK
PLVFNIVRDIKIPSVYVKKIKETPYLYVRVSGFDKNVTKSVLEGLKANPKAKGIVLDLR
GNPGGLLNQAVGLSNLFIKEGVLVSQKGKNKEENLEYKANGRAPYTNLPPIAVLVNGGSAS
AXEIVAGALQDHKRAVIGIEKTFGKGSVQMLLPVNKDEAIKITTARYYLP SGRTIQAKGI
TPDIVIYPGKAPENENKFSLEADLKHHEQELKKLDDKTPNSKEADKDKKNEEEKEVTP
KMINDDIQLKTAIDSLKTSIVDEKMDEKAPKKK

>PROKKA_00206 hypothetical protein

MSRKATNQILQKAKQLKSDEFHGENIQGYFYFIDESLNKNQNYEELQKLSVDYGVPLS
LCYGKELFENLNILQVWDEILNHLARWREILPDLPSLNFDENPLESFREIKDLVPSVYRK
LLDNDEIFNLVLILFPEQVKVLKLLVEYFKQQNKTIYQQLALKLAARLLPLR

>PROKKA_00207 Type-2 restriction enzyme EcoRI

MDKLTAAANYGMPVNQNLQCNKFIQHKDRVIMLQAISIYTQGDGQKWNLNKMFVEMFEAAK
TSLRVLGSPLFNQIVGKNNVKKSD

>PROKKA_00208 Type-2 restriction enzyme EcoRI

MGVSSIFKLIMQKLENEFKNLSFRHRSTSITKEEINSVLQGLDDELGKTLFIQNSKIKPDG
GIIEVKDDERNWRVILITEAKYQGGKDIENIQKGILVGKDSNQDLMQAGNAIERAYKNIAE

MANFMLKELHFP

>PROKKA_00209 UDP-N-acetylenolpyruvoylglucosamine reductase
MLETIIDFSRYSSVKIGTPLKVSLENDNEISQEHQIIIGLANNLLIAPSAKNLALLGPNY
DYICDKGECVEIGGAANASKIFNYFRANDLGGLEFLGQLPGTLGALVKMNAGMKEFEIKN
VLESACINGEWLEKEALGLGYRSSGFGGVLRARFKKTHGFREEVLKACKSMRKSHPKLP
NFGSCFKNPPNDYAGRLLGVLGRGYCLKRVGFAKEHANFLVNLGGAEFEEALDLIELAK
TRVLQEYGIHLEEEVKILR

>PROKKA_00210 Flagellar biosynthetic protein FliQ
MESQLMKLAIETYKITLMISLPVLLAGLVVGLLVSIQATTQINEMTSLSFVPKILAVIGV
LILTMPWMTNMLLDYTKTLIKLIPKIIG

>PROKKA_00211 Flagellum-specific ATP synthase
MPLKSLKNRNLNQRFDLSPRYGSVKKIMPNIIVYADGFNPSVGDVVKIEKSDGSECVGMVVV
AEKEQFGFTPFNFIEGARAGDKVLFLEKGLNFPVGRNLLGRVLNPLGQVIDNKGALDYER
LAPVITTPIAPLKRGLIDEVFSVGVSIDGLLTCGKGQKLGIFAGSGVGKSTLMGMITRG
CLAPIKVIALIGERGRIPEFIEKNLKGDLSSCVLVVATSDDSPLMRKYGAFCAMSVAEY
FKNQGLDVLFIIMDSVTRFAMAQREIGLALGEPPTSKGYPPSALSLLPQLMERAGKEENKG
SITAFFSVLVEGDDLSPIADQARSILDGHIVLSRELTDYGIYPPINILNSASRVAKDII
SESQNLCAKFRRLYALLKENEMLIRIGSYQMGNDEKELDEAIKKKALMEQFLAQDENALQ
PFEQSFQQLEEILR

>PROKKA_00212 Type IV secretion system protein PtlH
METLQTHRVLQALISHFTPFLESGITELIINTEQELWLYKVNNTREKRGHALFDKAFLLR
FCEQLASFRGLFFDEEHPTLNCSIPFTRYRVSANHFSITTNNQITLNIIRVPRLKPLSLED
FTFKASDPKSLKDLALKGHNILISGETSSGKTSLLNALLDCVNKDERVVSVEDSQELDLK
AFSNCVGLLVGKQENTRFNYEDALNMAMRLNPDRLIVGEIDTRSAALFLRLGNTGHKGML
STIHANSAQNTLEALSLNLSMRYTHSLDKDLMRAYFKSAIDVIVHVNRINDERQIAEVLW
TKEL

>PROKKA_00213 Isoleucine--tRNA ligase
MKEYKDTLNLNTTTFSMKGNLSVNEPKTYAKWQEQQAFKRMQNRKDNHGDFTLHDGPPYA
NGHLHLGHALNKILKDIVVKREYFKGKKIYYTPGWDCHGLPIEQQILERLEKEKTSLENP
TLFREKCRDHAKKFLEIQKNEFLQLGLVGFEDPYKTMDFKFEASIYRALVEVAKKGLLK
ERHKPIYWSYACESALAEAEVEYKMKKSPSIFVAFDLKKESEKLEKLVKKASLVIWTTTPW
TLYANVAIALKKDAVYALTQKGYLVAKALHEKLAALGVVDNEITHEFNSNDLEYLKATNP
LNQRDSLITLGEHVGLDGTGAVHTAPGHGEEDYYLGLRYNLEVLMSVDEKGCYDEGIIH
KKLLDESYLGEHVFKAQKRIIEQLGDSLLLEQEIHSYPHCWRTHKPVYIYRATTQWFILM
DEPFIQNDGSQKTLREVALNAIEKVEFVPSSGKNRLKTMENRPDWCLSRQRKWGVPLAF
FIDKRTNKPCEFESEVLEHVANLFEKKGCDVWVWESSVKDLLPPSYQDNAKHYEKIMHILDV
WFDSGSTFKAVLEDYHGAKGQSPSDVILEGSDQHRGWFOSSLLIGCVLNNQAPFKKVITH
GFIVDEKGEKMSKSGNVVSLDNLLKKHGSVDVRLWVAFNDYQNDLRVSQTFFIQTQEHY
KKFRNTLKFLLANFSDMDLKNLERPHDFSPLDHFIEALETTSGVNGAFEEHDLVKGLN
ILMAFVTNELSGIYLDACKDSLYCDSKNNEKRQAIQMVLLAVASQLCYFLAPILTHTIEE
VLEHSQVLCAFLQAKDVFDLKGINILEKLHLKEFKKPENFEAVLALRSFNEGLDFLKKS
GMVKNLSLECGIKLQEYASKEQMELIEELLMVSFVGIAKEKLSETPAFTLFKAPFYKCPRC
WRFKSELENTPCKRCEQVLKER

>PROKKA_00214 S4 domain protein
MRIDKFLQSVGLVKRRVLATDMCNVGAVWLNGSCAKASKEVKAGDVISLHYLKGIEEYTI
LQIPTLKNVPRKDTHLIYIAPKTKE

>PROKKA_00215 hypothetical protein
MKFQIVSLLAVLLLASCLPPKGHHSGLVNMVIAHQGQSVRTYWRKVVDREIVAKHNETLEK
NPKAKLKDPRGPLFMLGSNRFMLLWKNRYALAKVQSFRLEPGFYLLDSFSVETQKGILQS
APGYSYTKNGYDFKNKRPFFLAFEVKPDSKITLPSVELSLIKTPRGFLGVFLFDNNEKGA
NTQWVEGSLNLKLNASFKDAWGLEQ

>PROKKA_00216 Helicobacter outer membrane protein
MSTYSFFMDFLHNVYENEKFFIGYFIGVGLGGESVTANVLKDFGNILAKLVQFQGYGSLG
LRMGDKRHTLELSASVHGDAPSCSLKKLKSCESVRVLQAKIPRGIFESYITWSADYVYRF

>PROKKA_00217 hypothetical protein
MLGFNLSHAIDSELNKLVGKKVQEALAKEKAKAHQRADKSAFFAGLGYSGMFSWNLAY
KDSFVLLPHYDFNFGGRLLERKDPSTILNGFNFKVGYQYAAPVKIKKMFGVGRVRLILL

>PROKKA_00218 hypothetical protein

MKNNTAVITTTITTTHTTTTTITAVNTTITTTALIMKKVVAALAIVIIKKKVATGITSNIG
VARGNLTRVVSGF

>PROKKA_00219 putative dual-specificity RNA methyltransferase RlmN

MKASIYDFTLKELSQLLKPSFRAKQLYLWLYAKYKTSFKDMQNNFSKDFIAYLEREFTLR
TIEIAHVRKSVSDGSKYLFKSLKDNHTFEAVLLKMKDKKIDGETNAVLEGEKYTVCVSCQ
IGCQVGCSFCFTQKGGFVRNLKASEIIQQALLIKEDNNLPIEKAFNIVFMGMGEPLNNLD
EVCKAIEIFNTGMQISPKRITISTSGVADKIPILAGKNLGVQLAISLHAVDDKTRSSLMP
LNKKYNIECVLNEVRKWPLEQRKRVMFEYLLIKDLNDGLDCAKKLLKLLNGIKSKVNLIL
FNPHEGSKFERPSLESARMFADFLNSKGLLCTIRESKALDIEAACGQLREKKLSQQI

>PROKKA_00220 Arabinose 5-phosphate isomerase KdsD

MPILFDCNATAIQVLRDEASALLESVGQFQKPNDEAIVKLILKSQENGGLVIVGVGKS
ALVAQKIAASMLSTGNRS AFLHPTTEAMHGDLGMVEKNDVILMISYGGESLELLNLVSHLK
RLSHKIITFTKSPNSSSLKLGYYLSLKIKKEACPINTAPTSTTLTALGDVLMACLMR
AKNFSQEDFASFHPGGLLGKFLVKVDLLQTTNPLILPSTSFKDALIEMSEKRLGSAI
LVNEANELVGILSDGDVRRALLKGLSLESEVKHFATLKPKSFKNLDALLLEALEFLERHK
IQLLVCVDDCNKVLGVLHLHLQLLELGLKA

>PROKKA_00221 Ribonuclease J 1

MTDNNHYENNESNENSSSENSKVDEARAGAFERFTNRKKRFRENAQKNGESSHHEAPSHHK
KEHRPNKKPNNHKKQKHAKTRNYAKEELDSNKVEGVTEILHVNERGTLGFHKELKKGVET
NNKIQVEHLNPHYKMNLSKASVKITPLGGLGEIGNMMVIETPKSAIVIDAGMSFPKEG
LFGVDILIPDFSYLHQIKDKIAGIIITHAHEDHIGATPYLFKELQFPYGTPLSLGLIGS
KFDEHGLKKYRSYFKIVEKRCPI SVGEFII EWIHITHSII DSSALAIQTKAGTIIHTGDF
KIDHTPVDNLPTDLYRLAHYGEKGVMLLLSDSTNSHSGTTPSESTIAPAFDTLFKEAQQ
RVIMSTFSSNIHRVYQAIQYGIKYNRKIAVIGRSMENLDIARELG YIHLPYQSFIEANE
VAKYPDNEVLIVTTGSQGETMSALYRMATDEHRHISIKPNDLVII SAKAIPGNEASVSAV
LNFLIKKEAKVAYQEFDNIHVSGHAAQEEQKMLRLIKPKFFLPVHG EYNHVARHKQTAI
SCGVPEKNIYLMEDGDQVEVGPAFIKKVGTIKSGKSYVDNQSNLSIDTSIVQQREEVASA
GVFAATIFVNKNKQALLESSQFSSLGLVGFKDEKHLIKEIQGGLEMLLKSSNAEILNNPK
KLEDHTRNFIRKALFKKFRKYP AII CHAHSF

>PROKKA_00222 Ribosomal RNA small subunit methyltransferase A

MVVAKKS LGQHFLTDESFLDRIVNALPPLNPLK LIEIGVGLGDLTLKLLDRYPLKTYEID
SNLCEKMRSKLKAQKKPFKLELVEKDALFLKEE EPYFLISNLPPYYIATRLVLNALKDPKC
RGLLVMTQKEVALKFCAKDSQNALS VLAHTIGNATLLFDVPPSAFSPPP KVFSSVFEVIK
EPLKEKALASLLQVPFFEEALQKGFETLED FLKACFSSPRKTL SNNLKKS VSYREKLDKV
LDFLALENQPTS VRASEIKDYLLKLE YLLKG

>PROKKA_00223 hypothetical protein

MAHHEQQQQQANSQH HHHHHH AHHHHHY GGEHHHHNAQQHAEQQAEQQAQQQQQQQQQQAQ
QQNQY

>PROKKA_00224 Putative type I restriction enzyme specificity protein MPN_638

MLPSNWQRVRLGDIGKPCMCKRVMKHQTTRYGEIPFYKIGTFGNTADAFISKKLFLEYKT
KYSFPKKG DILISASGTIGRAVIYDGKPAYFQDSNIVWIDNDET LVKNDFLYTYSHVKW
NTEHTTILRLYNDNFRNTLIPLPLNEQIAIANILSDVD RYLYNLDALILKKEGVKKALS
FELL SQRKRLKGFNQAWQVRLGDIAEIKRGVRITKNELDVFGKYPVVS GGVGFLGYTNN
FNRYENTITIAQYGTAGYVNFQKNKFWANDVCFCIYPNKDIIKNIFLYYFLKVNQNYLYE
ISNRNATPYSISKDKILD FEILLPPLNEQIAIANILSALDNEIISLKNKKRQFENIKKAL
NHDLSAKIRVLK

>PROKKA_00225 Formyltetrahydrofolate deformylase

MLEFILKI QARDSKGLVSAISSTIANKGYNIVKNDEFVDPLKQ RFFMRLKIQKEIKPLNT
EIKEQEERSLKTALFKALENFSELLIEVILTHKKNIILLATKESHCLGD LLLRVYGGELN
AQILGVISNYEVLRLPVEKFDIPYFYAPCVDQILHEKEVLAI IKDLELKHKVSADLLVLA
KYMRI LSHDFTKRYENQILNIHHSFLPAFIGANPYQQAFERGVKVI GATAHFVNESLDAG
PIILQDTLPINHNSVEKMRLAGKDIEKLV LARALKLVLED R VFVHENKT VVF

>PROKKA_00226 Putative signal peptide peptidase SppA

MWSFIQKIFKALIMPLDFITKYFKSFVLLLVLVVFFSAKESAPSEPPNLAKLYLNGAIF
STEDFDKEVDKILKTPSIKGVLLIDSPGGAVSASVELSEKIADLKQKMPVLAYARGVMA
SGSYYAGMQASEVYVSKASLIGSIGVIFSGANVENLLNKVGVATQGVHAGEYKEIGTFTR
AWKPNEKEFLQNLVNEQYQMFVDDVAKARKLDAKDYK DFAEGKVFS AQNALK LKLIDKIS
TIKQAQDRLMELSKVKAYWLEKSPMERFIEKATQSASNIITQALGYQLLMR

>PROKKA_00227 hypothetical protein
MACKFCPKIRKTDWIFILVAALGFYLVNKLGYAPQFTPTSKPSHSLSRPIEKPDNMTEEE
RKKRFIELQKACLLHKDKKACEEVF

>PROKKA_00228 hypothetical protein
METYIIDADNIDGDLFFYNLTRNSNDFSMLPVFELDHIAQKIRNILKKHGSRKDIILNHR
ELQEAFSPFPKQLKNAQVFLSHSHADKNKALEVKNYLEE

>PROKKA_00229 hypothetical protein
MAKYDDTISKIEDAFTLILRESLQDMIEKCPYFVFLQSKNSVSLNQDLLGITYSAWIYEEL
KIAHSISSKSHLTIVMESFQVSHDISPFLESFETITLKKLSQQINS

>PROKKA_00230 hypothetical protein
MALYLAGCGGDSGKPLVGIEKNSFNSTVKIISKTDNIEIQDLKLNRCNCEHDQNFLVKLI
QETANAYLVFVSEKEKAIKNHQAKIARLQKDLEELTQHVQQSNNLDKLLNEGLFVSGHDY
KYTKDDNPIYVVKRMLDNLDSEYKE

>PROKKA_00231 hypothetical protein
MWPSVGKFYFAYLKEVIPRQAKEDTEKQISSEKSIKANQVKLTEATQDIDKMEKLIKDL
ESKNTLSVYLFKFGESFTAHYKCQNLIIEVGVKTDKGTWTFNFKR

>PROKKA_00232 hypothetical protein
MENTPIDRAKILIEELKILQGVINRMAQNSLECKWTLALAVGVLSLKIEAISNSYGLCV
LGVLLLCFWFLDAYYLTQEKLFREQYQWLIENRLKTDERLFEVFPAHQTCQPMQLLSAMF
SISPSSISLLPLLGVSSLFGLWFVFNVSVMKHWL

>PROKKA_00233 hypothetical protein
MKNKYHESHIMKSIDEVGFRETDKIVEENPKLKTENAELRLENQELKAENATQKQLIVTQ
KQLIDALKQLNESKDETIKAKDQTIEILMKCSTTEERYRDEESEFSKFCLNLQ

>PROKKA_00235 hypothetical protein
MAFWHKRLAVGCCIVLFSCMMNANNIQIVRDDPPLDPTLPAWIYSIALLKVYFSDGTYKE
GYATLLKNGRYIASSETLYSNGLYPKMILAKMQDSSAKELIXIASLRLEAMDRNQGLSL
KTADFRDDYCHKREESYHARIYAKYAQTFHSPYTHQKTPNSDLYYPALNEGNSFSIQI
TGISVAELLKSKKFLSLDSSFKKGSVLWGGRPYFSEVGEFMMASSTLENHESLVIIPKE
KIVQFLSALKNQNIFQNIP

>PROKKA_00236 Peptidyl-prolyl cis-trans isomerase B
MDPIKTYEIKEEELAKTAYATIKNKGNIALFLYKDAPQAVSNFVTLAKEGFYDGLNFH
RVIAGFVAQGGCPYGTGTGGPGHRIKCEVAHNPNKHKRGSISSMAHAGRDTGGSQFFLCFV
DLPHLDGEHTVFGKIKNAESLNVLDKIKQGDIIESVVFSSSL

>PROKKA_00237 hypothetical protein
MLILSRKVNEGIVIDDNIHIKVISIDRGSVRLGFEAPESTLILRAELKEAIVSENQKASV
CVDESLENIKKVIKP

>PROKKA_00238 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
MHVFEVYPKVNIFLKLHKEGAYHKLISRMCIVKDKLKDIISVKNALSFSKGFDFDCPLE
ENSLFKALQILKNFLKSKNFSHSIIKSLDTLAEVEKNIPTQAGLGGGSADAGGLLYHLN
QIFDCRLSLKELYSMSGSLVGADTNFFISQYKSANATSYGEVIENFEEEEPLENRLEIYAPT
HVFCSTKAIYQVYKPETCFSSQAKEWLKKPSLECLKTYDRSELNDLLKPALCANPALKDIE
SQLSKEWFFSGSGSAFFRLKPTLKGGE

>PROKKA_00239 SsrA-binding protein
MKLIASNKKAYFDYEILETLEAGLALLGSEVKALRQTRVNLKDNFVKIIEGEAFLFGVHI
SYLDTIHAYYKPNERRERKLLLHKKQLLKWQIEASKERLSIVGLKLYFNQRNRAKIQIAL
VKGKRLHDKRQSLKEKALNKEILADLKHHFKG

>PROKKA_00240 Biopolymer transport protein ExbB
MKEMVDYGIIGFLIFLSVIVIAIAIERLWFFATLRVDDYTDNRKLELALHKRLTLVVTIG
SNAPYIGLLGTVMGIMLTfMDLGSASGIDTKAIMTNLALALKATGMGLLVAIPAIVIYNL
LVRKSEILVTKWDIFHHPVDTQSHEIYNKA

>PROKKA_00241 Biopolymer transport protein ExbD
MKKVESMNVVPFIDIMLVLLVIVLTASFVQTSKLPISIPQVDKSDTSDKDVLDKKQVTI
SISNKSFYLDDEKISFENLKHKVSTLAKDTPIVLQGDKKSNNLDNFIKVVDLLQTNNLKQ
LYILVEDKKNQKN

>PROKKA_00242 ribonuclease P
MPDELRAEKSFPSPKPYDSLKNKSEFDRVYQKGFKKHNPFFSLFVLDLSKEPPKGKEGFKD
PLSCRLKDKKTLYLLGLSVSKKVGNAVKNRLIKRRLSLALKHVALCQGLALVFVPRSDC
YHLDFWALEKHFFLEMLTSIKNYMNAKALKGLKKGITHTHAKQ

>PROKKA_00243 Putative membrane protein insertion efficiency factor
MRNNKTPFLSAIFTASIRGYQRFFSAFTFSSCRFYPTCSNYALWLLYFENPLSAMGKIAI
RILSCNPFCSGGIAYPTTRLKRPSLLQSHKDFNRNFKTITFWLVPTAKSRTTYIIKV
>PROKKA_00244 Membrane protein insertase YidC
MNKNNNTNLILAIALSFLFIALYSYFFQKPNKPTTQTTKQETTNNHTATSPNAPNAFSAT
QTIPQENLLSAISFEHARIEIDSLGRIKQVYLKDKKYLTTPKQKGFLEHVGHFSSKENMQ
PPLKELPLLAADKLKPLEVRFLDPTLNNKAFNTPYSASKTTLGPNEQLVLTQDLGAFTII
KTLTFYDDLHYDLKIAFKSPNNLIPSYVITNGYRPVADLDSYTFSGVLENNDKKIEKIE
DKDAKEIKRFSNTLFLSSVDRYFTTLLFTKDPQGFEALIDSEIGTKNPLGFISLKNEANL
HGYIGPKDYRSLKAISPMLTDVIEYGLITFFAKGVFVLLDYLYQFVGNWGWAIIFLTIIV
RIILYPLSYKGMVSMQKLKELAPKMKELOEKYKGEPOKLQAHMMQLYKKHGANPLGGCLP
LILQIPVFFAIYRVLYNAVELKSSEWILWIHDLSIMDPYFILPLLMGASMYWHQSVTPNT
MTDPMQAKIFKFLPLLFTIFLITFPAGLVLYWTTNNILSVLQQLIINKILENKKRMHAQN
KKEN
>PROKKA_00245 hypothetical protein
MQNFIEIKAKTLEEALIQASIALNCPIINLQYEVIIQTPSKGFLSIGKKEAIIAGVKESV
KEVKEESVKETNTKENHQNNIEKKQKLETETPQEERITPKPPKKNLKEESHNGDKLHEI
KQELKELFSLHPYKINKVEVSLYEPGVLLIDIDGEDSALLIGEKGYRYKALSYYLLFNWIH
PAYGYSIRLEISTFLQNQEKVMEAQLQSVIMTVHEVGKGQMKAPDGVLTYYIALKKLRKAF
PNKYVSIKTNLNDEKYIVINDFNNE
>PROKKA_00246 tRNA modification GTPase MnmE
MNDTIAAIAATPLGKGAISIIKISGNNALNVLKQLTQKQDFTPRYAYVCDIFSNGVLLDKA
LAIYFKAPYSFTGEDVCEIQCHGSPLLAQNILQACLNLGARLAKAGEFSKKAFLNHKMDL
SEIEASVQLILCEDESVLNALARQLKGELKIFIEEARGNLLKLLASSEVLIDYSEEDIPS
DFLNEVSLNLEKQIASFKDLLDFSNAQKQRNKGHALSIVGKPNAGKSSLLNAMLLEERAL
VSDIKGTTTRDTIEEVIELKGHKVRLIDTAGIRESADKIERLGIEKSLKSLNCIDIILGVF
DLSKPLEKEDFNLDITLNRACKPCIVVLNKNLDLPKLELEILKSYLKIPYSLLETNTLNS
KACLKDLQKISTFFPKLDTQNKLLTSLAQKTALENAIFELQNAKNHLETLELFSYHLL
SAIENLNLTRPYETSQMLDSMFSEFCLGK
>PROKKA_00247 Helicobacter outer membrane protein
MLKLASKTICLSLIGSFTALEAYQKHQKDGFFIEAGFETGLLQGTQTQEQTIAATTQEKPK
PKPKPKPITPQSTYGYKYYISQSTVLKNATELFAEDNITNLTFYSLNPVYVTAYNQESAEE
AGYGDSSLIMIQNFLPYNLNINIELSYTDNQGNLVSGLVIETIPKQSQIILPASLFNDPQL
NADGFQQLQTTTTKFSDASTQNLFDKLSKVTTNLQMTYINYNQFSSGNGTGSKPPCPPYE
NQENCVAKVPPFTSQDAKNLTNLMLNMMAVFDSKSWEDAVKNAPFQFSDNNLSKPCYSDY
STCVNPYNDGLVDPKLIAKNKGDEYNIENGQTGSVILTPQDVIYSYRVTNLYVNLLPTR
GGDLGLGSQYGGPNGPGDDGTNFGALGILSPFLDPEILFGKDLNKVAIMQLRDIIEHYGH
TLGYTHNGNMITYQVRMCEENNGPEERCKGGKIEQVNGQEVQVFDNGKEVRD TDGSFYDV
CSRFGGQNQPAFPSSYPNSIYTDCSQVPAGLIGVTSAVWQQLIDQNALPVDYTNLSSQTN
YLNASLNTQDFATTMLSAISQSLSSSTKSSATTYRTSKTSRPF GAPLLGVNLKMGYQKYFN
DYLGLSSYGIKYNYAQANNEKIQQLSYGVGMDVLFDFITNYTNEKNPKNNLTKKVFTSS
LGVFGGLRGLYNSYYLLNQYKGSNLTGGLNRYRYKHSKYSVSGISVPLVQLKSRVVSSD
GTTTNSITLNEGGSHFKVFFNYGWIF
>PROKKA_00248 hypothetical protein
MNAKEQKQKQNNETEKSSNELKENTQAMREQKERTLTLMCKTELKPLTREPKKSVLEPG
VKANF
>PROKKA_00249 hypothetical protein
MKKIIILACLMAFVGANLSAEPKWYSKAYNKTNTQKGYLYGSGSATSKEASKQKALADLVA
SISVVVNSQIHQKSRVDNKLKSSDSQTINLKTDDLELNNVEIVNQEAQKGIYYTRVRIN
QNLFLQGLRDKYNALYGQFSTLMPKVCKGVFLQQSKSMGDL LAKAMPIERILKAYSVPVG
SLENYEKIYYQNAFKPKVQITFDNNSDTEIKNALISAYARVLTPSDEEKLYQIKNEVFTD
SANGTTRIRVVVSASDCQGT PVLNRSLEVDEKNKNFAITRLQSLLYKELKDYANKEGQGN
TGL
>PROKKA_00250 hypothetical protein
MKRLAIALALVLGVWAGKSLPKWARDCSKEVQVEKTQTKDEKFLVCGMSDILLSDMDYSL
SSARQNALEKVMFAFKGDKIEIKASELKATFIDTDKVYVLLKITKKHVALTNE
>PROKKA_00251 LPP20 lipoprotein precursor
MKNQVKKILGMSVIAAMVIVGCSHAPKSGISKSNKAYKEATKGAPDWVVGDLKVAKYEK

YSGVFLGRAEDLITNNDVDYSTNQATAKARANLAANLKSTLQKDLENEKTRTVDASGKRS
 ISGTDTEKISQLVDKELIASKMLARYVGKDRVFLVGLDKQIVDKVREELGMVKK
 >PROKKA_00252 hypothetical protein
 MLLKTKLKIISSVILSTLLWVGCSSEMATYQNVNDATKNTTASINSTDLLLTANAMLDLM
 FSDPNFEQLKGKHLIEVSDVINDTTQPNLDMNLLTTEIARQLRLRSNGRFNITRASGGSG
 IAADSRMVKQREKERESEYEQDTTVEKGTLLKAADLSLSGKVSSIAASISSSRQRLDYDF
 TSLTLNRKTGEEVWSDVKPIVKNASNKRMF
 >PROKKA_00253 Thioredoxin
 MLEVINGKNYAEDIAHQAVVVNVGASWCPDCRKIEPIMENLAKTYKGKVEFFKVSFDESQ
 DLKESLGIRKIPTLIFYKNAKEVGERLVEPSSQKPIEDALKALL
 >PROKKA_00254 Ribosomal large subunit pseudouridine synthase B
 MRINQFLAHYTKHSRREAELVLEGRVKINHEHAKLASVVKENDKVFLDKRLIKPLKNKK
 FSVLVYHKPKGELVSKADPLKRRVIYESLEKKYAHFAPVGRLDFASEGVLILSDSKAVVS
 ALMHANLEKEYLVKIQGFVTREMENAMQEGKLKLENATKGAHQKTPIKSMEFAPFIGYEII
 KNHAKYSKLRVVINEGKNRELRRFFAFFNAGVLDLRRVRYGFVNLNALPVGKMRFLNRQE
 YNELHAFMANRANTKGD
 >PROKKA_00255 DNA polymerase III subunit alpha
 MKENKAFTHLHLHTEYSLLDGANKIKILAKRVKELGMKSVSVTDHGNMFGAIDFYTSMKK
 EGIKPIIGMEAYIHNDNLSSKETKQRFHLCLFAKNQEGYENLMFLSSMAYLEGFYFPR
 INKKLLKEHSGKIIASSACLQGEVNYHLNTNERNRKYGAKGYDEAKKIACEYQEIFEDD
 FYLEIMRHGILDQRFIDEQVIKMSLETGLKIIATNDTHYTMPNDAKAQEVAMCVAMGKTL
 NDKGRLLKHSVHEFYIKSPEEMAKLFADIPEALENTQEIADKCVLEIDLKDDKKNPPTPPS
 FKFTKSYAKEEGLNFEDDASYFAHKAREGLKERLILVPKEKHDQYQERLEKEIEVITNMK
 FPGYMLIVWDFIRYAKEMGIPVGPGRGSAAGSLVAFALKITDIDPLKYDLLFERFLNPER
 VSMPTDITDFCQRRRKEIIIEYMIKEYGKYNVAQVITFNKMLAKGVIRDVARVLDMPYKEA
 DDFAKLIPNRLGITLKGYEKNGEFIEGAWELEPKIKELVESNDTARQVWEYSLNLENLNR
 NAGVHAAALVVDSDQKELWHKTPLFASEKTGGIVTQYSMKYLEPVDLIKFDLGLKTLTVI
 DDALKIIKTQHKISVDFLSLDMDDPKVYKTIQSGDTVGIQIESGMFQGLNKLRLPSSFE
 DIIAIIALGRPGPMESGMVDDFVNRKHGIEPIAYALKELEPILKPTYGTIVYQEQVMQIV
 QTIGGFSLGEADLIRRAMGKKDAQIMADNKAKEFVEGAKNLGHDGQKAANLWDLIVKFAGY
 GFNKSHSAAYAMITFQTAYLKTYKHEFMAAMLTSESNKIESVARYIDEVRALEIEVMPP
 HINSSMQDFSVAEFKNQKGELEKKIVFGLGAIKGVGGEPIKNIIEERAKGDYKSLEDFIS
 RVDFSGLTKKSLEPLVKSGSLDNLGYTRKTMLANLDLICDAGRAKDKANEMMQGNSLFG
 AMEGGIKEQVILDMIDLGEHDAKTLLECEYETLGIHVSGNPLDEFKEEIKGFKNLVKSID
 IEELEIGSQAYLLGKIMEVKKKIGKRSKPYGTADILDYKGFELMLFEKQLNALEELDI
 NKPLVFKCKVEEQEEVARLRLFEILDLESAREVKIPKARYKDLDKQKEDVREIPPIEMLA
 SSSCSLAIVLENDVKKEFLRQIKESALKHQGRPLYLIIDKDKDKQFKIQSDLMVNEKIKD
 DFKGLEWRDLA
 >PROKKA_00256 Cytochrome c551 peroxidase precursor
 MKKSILLGVCLAFSCTHALNDLELIKARESQLPMPMGKALKEYQIKKTRDVGIGTKNS
 EIMTSAQVELGKMLYFDPRISTSYLVSCNTCHNLGLGGVDLVPSAVGSQWKNPHLLSSP
 TVYNSVFNDVQFWDGRVTHLNEQAQGIQSSFEMGADPKVVVEKINSMPGYVKLFRKAYG
 SKVKIDFKLIADSIAMFEATLITPSRYDDFLRGNPKALSKEKEGLDLFISKGCVACHNG
 INLGGTMQPFQVVKPYKFANVGDFKGDKNGLVKVPTLRNITETMPYFHNQGFWDVKDAIK
 EMGSIQLGIEISDEEAKKIETFFGALRGKKPKIIYPELPIMTDKTPKPSF
 >PROKKA_00257 META domain protein
 MNLRLAGASVLTACVFSGCFFLKMFDKKLSSSDWHIQKVEMNHQVYDIETMLADSAFREH
 EEEQDSSLNTALPEDKTAIEAKEQEKEKRKHWEYLFKKKPKPKSSMGEFVFDQKENRIY
 KGKGCNRYFASYVWQDRHIGIEDSGISRKVCDEHLMAFELEFMENFKGNFTVTKGKDT
 LILDNQMKKIYLLKTP
 >PROKKA_00258 hypothetical protein
 MRATAIKIFSLSSALALLQGCLSINLKQMLPEIRTYDLNASSFEITQCPKPLTEVRLIS
 ILSADLFNTKEIVFKAQDGGQITHGKHQKWIDLPRNMLKTMFMQEAQKACLGVALPPYGAG
 APTYAVRFTILSFSLEKENSTYRAEFALGYDISVKGDHSGVVIKHENISSLENKTTKT
 SKNGNQDFQESAIQSLQHVSTQAMQEAISLIKKAIEAQSVSPLKK
 >PROKKA_00259 hypothetical protein
 MGLKFLALEQSHNEEFYSGDGKGERILIFKEGLMDRLSGDANQVVQEVMAIRNVNRILD
 DENVEKFKHILASVDDLIANLDSRKTQFDSLINNANNLVSNVNNVALDVKRLKQGQYDF

KAMFTPLIMQAQLSLRNIDNFVEKGSALIDKFDANPYKTIFGERK

>PROKKA_00260 mce related protein

MERHVNYTLIGGLFFLCLVCMVGFILWLGLHGLDDGKYEYVYVYTDKDLGGIATNSPVNY
KGIQVGNVIKVGFAKDKVGVVRLDLMIKSSVKIRKDSKVAVSSRGVYGVKIFSLRAKPO

>PROKKA_00261 Methionine import ATP-binding protein MetN 2

MMNAINNQVLIEVKDLHSAFGNTIIHRGVSFVHKGEVMAILGGSGSGKSTLLRCMILLN
RPTKGEVLLFGEDIWKLKEREQQKIFNRCGICFQFGALYSSLTVLENVGVMLEQYGAYSK
KIVEEISKMWIEKVGLPPRAYHLYPYELSGGMKKRVGIARAMATNPEILFLDEPTSGLDP
YSAGKFDELIMTLKESLQLTVVMITHDLDVHDCVDRFIMLKDGLLFNGDLKDFIKKAQ
TQGLDEGNLNFSTRGEKFWKGM

>PROKKA_00262 putative phospholipid ABC transporter permease protein MlaE

MKTEKQKFLEMRKDGANSVLILRGDWDFKTSVFRLELKKNLDDHQSLKMDFGSCQKVD
FVFGMFLFDLIKERSLNIELCNVSENNACALKVVKDWLEKEEDLESKKAGKKYELLITKL
GKSIVETYNFTLNAFNFCGMILFYFIKSVFNPKRFCITPLLYHINESGFKVLPVSILTTF
IVGFAVALQGVLLQLQDMGAPLMSVEMTAKLALREIGPFILTLVVAGRSASSFTAQIGVMK
ITEELDAMKTMGFNPFEFLVLPRLALVIVLPLLVIADAFAILGGMFAIKYQLDLGFSS
YIDRFHDTVGNHFLVGIVKAPFWGFAIAMVGCMRGFEVKGDTESIGRLTTISVVNALFW
IIFLDAVFSIVFSKLNI

>PROKKA_00263 Helicobacter outer membrane protein

MFKKIIFFGVFLIGGFVIPPLDAMPILHDKTPKKNYQEAHEKLYRSIINRQKLTRKKS
YFLGGFGAVEAIKDYQGQEMKDWIATLNLKTGVQSFFKKYIGIRGVFAWDLGSGKVNYQS
HKDPTNSFFTMLAVGLDVIMEFPLGSYKHYLGAFGGARGALVVYTDKQNFKFFKHSVVS
GLAINGGVMLTLFLRHRIELGFKILPTARLLSSSKRFETSPLFYAAYSYKF

>PROKKA_00264 Branched-chain-amino-acid transaminase 1

MANLENLDWKNLGSYIKTDFRFIATYKNGSWSQGELVSENVLQLSEGSPVLHYGQACFE
GLKAYRSQKGKALLFRPLENAKRLQTSCERLLMPKVSEELFLRACAEVVKANQKWLAPYK
SGASLYLRPFVIGVGDNLGVKPAKEYLFVIFCAPVGAYFKGGIEKGGARFITTIFDRAAP
KGTGGVKVGGNYAASLLAHKMATDQGYDDCIYLDPATHTKIEEVGAANFFGITHDNAFIT
PHSPSILPSVTRKSLMVLAKKEYLKLNIEEREILMDELGAFREAGACGTAAIITPIKEIAH
NNKSYYFEAPGHITKQLYDLLLSIQQGEQEAPKDWIFEVG

>PROKKA_00265 Helicobacter outer membrane protein

MLKRIILLGALGVLASAEESAFAFVGNYQVSMIQNQTKMVNDNGLQKPLIKFPPYAGAGF
EVGYKQFFGKKKWFGARYYGFFDYAHNRFGVMKKGIPVGESGFIYNSFSFGGTTLTERDS
YQGQYYVNLFTYGVGLDTLWNFVNKENMVFGFVVGIIQLAGDSWATSISKEIANYAKHHSN
SSYSANFQFLWKFGIRTHIAKHNSLELGIVPTITHQLFSLTNEKGYTLQADVRVYAF
QISYLRDF

>PROKKA_00266 DNA polymerase I

MEQPVIEGTLALIDTFAYLFRSYYMSAKNKPLTNDKGFPTGLLTGLVGMVKKFYKDRKN
MPFIVFALESQTKTKRAEKLGEYKQNRKDAPKEMLLQIPIALEWLQKMGFVCVEVNGFEA
DDVIASLATLSPYKTRIYSKDKDFNQLLSDKIALFDGKTEFLAKDCVEKYGILPSQFTDY
QGIVGDSSDNYKGVKGIGSKNAKELLQRLGSLEKIYENLDLAKNLLSPKMYQALIQDKGS
AFLSKELATLERGCIKEFDLSCAFFPENPLLIKDELKEYGFISTLRDLENFSTPLILD
NAPALDSAPTS DNAPRKS RMIVLENAAPLSAFLEKLKNSNARVFMRLVLNKEKKVLALAF
LLQDQGYFLPLEEALFSPFSLEFLENAFSQILQHACII GHDLKPLLSFLKAKYQVPLENI
RIQDTQILAF LKNPEKVG FDEV LKEYL KEDLIPHEKIKDFKTSKAEKLELLSVELNALK
RLCEYFEKGGLEEGLLALARGVETPFMKVLMGMEFQGFKIDVPYFKRLEQEFKDELKVLE
RQILD LIGVD FNLNSPKQLGEVLYEKLGLPKNKSHSTDEKNLLKILDKHPSIALILEYRE
LNKLFNTYTTPLRLKDKDDKIHTTFIQGTATGRLLSSHSPNLQNI PVRSPKGLLIRKGF
IASSKEYCLLGVDYSQIELRLLAHFSQDKDLMEAF LKGRDIHLETSKALFGEDLAKEKRS
IAKSINFLGVYGMGSKKLSETLNIPLNEAKSYIEAYFKRFPSIKDY LNMKKEEILKTSKA
FTLLGRYRVDFDNGVNDYVKGNYLREGVNAIFQGSASDLLKLGLMKVSERFKNNPSVRL
LQVHDELIFEIEEKNAPELQQEIQIRILNDEVYPLRVPLETSAFIAKRWNELKG

>PROKKA_00267 Type I restriction modification DNA specificity domain protein

MWQSFRLGDLFEVLPYKKRFDANKVNIYASKTKDITYPVVRTSLNNGIRGYLKENTNFLN
AGNTISFGQDTATMFYQEKPYFTGDKIKILRCKNPFNKINALFFITSLTKAFRNFSWGS
ASFVSIIENQINISLPTNPHGKIDFNFMRTLINALMKQTIQGVVQYSGAKIQATKEAISQ
ETPVQKDSL

>PROKKA_00268 Restriction enzyme BgcI subunit beta

MIGPLSGQLNAIKWGEFRLGDLFEIYTGSLLSQEDLAKGDIVRISVKSDNNGVYGHFDTL
KNKKARHFENFISVNNFFGNCFYHPYLASVEMKVHVLKLNRTLTKRVGLFLANQLNKCFY
GQFTYGTQLSSSKLKHNNFKIQLPLKPTAKIQTLEDIDDFDFMEKFIAELEQCRLAELEAY
LKATGLENNTLSNDEENALNVFNNSMGG

>PROKKA_00269 Restriction enzyme BgcI subunit alpha

MQAFRLEDDVDDYVKKELTNLGLMKNKDFNVKSQMSSSLKNALLNASKTKDKTSYGEPDF
SLEKYTHPKNKESVIPIIIEENKLYAKNLKKLKNSTLQNDDNSISKYAVNGALHYAQNILR
DKEYKKECIAIGIAGDDEENLLIEVYYVFASGINSHKLTSAKNLHFLENQESFNAFYKEC
TLTEEEKHSILIKTKADLNETAKKLNRLMHNHNITASQRVLYVSGMLLSMQEIKGKKGGL
KPSDLKGEITDTSRDGVLVFNQISEFLKTKNLSEEKRDMLASFKEISKDPQRDKETS LD
KAISMLLEKDSSITKQIFTFLYEFVHKPINESDNTGHLDIMGELYSEFLKYALGDGKELG
IVLTPPYVTMMSEELLGVNAKSFVMDLAAGSAGFLISSMVLMIEDIEKTYGKNTTIANEK
IKCMKTTQLLGVELNAEMFSLATTNMILRDDGSSLIIGKNTFETSKKIYEDFKPNILLN
PPFSYEENGMPFIKFGLEHMQKGALGAI IQDSAGSGQALKSNVEILKKHSLLASVKMPT
DLFMPQAGVQTSVYIFKAHEPHDYEKPVK FIDFRNDGFKSTKRGLNETSNPTKRYEEI IK
IYKAGLNAKVS KELWGD LKTIYIEDFIAKPCENKHAKDFNF EAHQKNETKPELEDFKRTI
ADYLSYEVSLILKNQTPPK

>PROKKA_00270 hypothetical protein

MRCLTCLKLSFKPLCPNCLNDLPLSLKVRVLEGVSVYSFYAYSEIEELIKSKYALIGSRI
LPLLSQKAGKEFVKILQEQLNIPFYGIAIDDKIKSFYSHSAALLKGFCQGNLKPTYGTL
RATNAVSYAGKSLEFRANNPRDFTFKGDENLDYFLDDIITTGTTLKEALKYLKALNIKA
HFAIALCSADE

>PROKKA_00271 Thymidylate kinase

MYVALEGVDGAGKSTQVGLLKTRFKNALFTKEPGGTKIGESLRRIALNENISELARAFLF
LSDRAEHVESVIKPALKEKRLIISDRSLISGMAYSQFSSLELNLLATQSVLPEKIIILLVI
DKEGLKQRLSLKSLDKIENQGTEKLLTIQQKLKTHAYALREKFGCEVLELNAKESVKNLH
EKIAAFIECVV

>PROKKA_00272 Phosphopantetheine adenylyltransferase

MQKIGIYPGTDFDPVTNGHIDIHRSSSELFELIVAVAHSSAKNPMFSLDERLKMMQLATT
SFTNVECVAFEGLLANLAKEYHCKVLVRGLRVVSDFEYELQMGYANKSLNHELETLYFMP
TLQNAFISSSIVRSIIAHKGDASHLVPEKIYPFISKA

>PROKKA_00273 putative aromatic acid decarboxylase

MKLVLGISGASGIPLALRFLEKLPEIEVFVASKNAHVVALEESNINLKNAMKDLRPSA
TFFNEQDIHASIASGSGYGIHKMAIIPASMDMVAKIAHGFGGDLISRSASVMLKEKRPLLI
APREMPLSAIMLENLLKLAHSNAI IAPPVMTYYTQSKTLEAMQDFLVGKWFDSLGIENDL
YPRWGMN

>PROKKA_00274 flagellar basal body P-ring biosynthesis protein FlgA

MKILILFFICLNVLFALDSNAKTEIKETYLKEYKDLKLEIETINLEIPERFSHASILSY
ELNASNKLKKDGVVFLRLNEPNLRLPVRYSVIGSMQAFKSISA IKK DENITANNTKKER
VLFGTLSNPLLEG AIDRVSAKNFIPD TLLSADKTQALIIVRKNDIITGVYEEGQISIEI
SLKALENGALNQIIQAKNLESNKILKAKVLSSSKVQIL

>PROKKA_00275 ATP-dependent DNA helicase PcrA

MGFEKSILDNLNGAQKIATSHIQGPLLILAGAGSGKTKTLTSRLAYLIGVCGVPSENTLT
LTFTNKASKEMQERALKLLKNQALIPPLLCTFHRFGLLFLRQHMSLLKRACDFSVLDSDE
VKTLCQQLKISNFRANISQIKNGMMDLSVQDSECHKAYELYQNALKKDNLVDFD LLLFLS
LKILQDNETIAKETSERYHYIMVDEYQDTNALQLEFLKKLSCTHHNLCVVGDDDQSIYGF
RGADISNILNFSKHFKGAKIVKLETNYRSSAEILACANSLISHNQHRHVKT LQSFKGSHK
SVVCKEYLTQKEESLDVAYQIKALLKKGENLENIAILYRLNGLSR SIEESLNALNIPYRL
IGAVSFYERAEVKDALAFMHLVAKKDDRFFIKRVLNKPPRGLGKITQEWIFSLDEEDLN
LEEALKLGAFKDKLNPKN EYALKKFMTMIGRLREAFEISVEEFCARFLEETNLLKSYEKE
DNYEEREGFVKELLSLVKEHFKNPTHSLD FLNESVLDVHNAEN AQKVS CSMVHMSKGL
EFKHVFVIGLEE GFFPHRGFNQESDLEEERRLAYVAITRAKEGLQLSYVKERSYFGRKIS
CSPSVFLEE AQLLKQDKPPKQNHQKDTPIKVGD LIRHKIFGTGRVLGVEKGLSGVCLKIN
CGGNVYDKISEKFVEKVDNEF

>PROKKA_00276 Tetratricopeptide repeat protein

MLNEEQNSLEEKGGENKNEKETPLKGIHSKIPSLKQALEQTISKIKSSKEFFKQLSRNKK
KLYIALGILLSLIALIVILSLLLGHKKENKQTS LQTN SATTNNETPNNTN NANNTEAEGQ
IENLDLPDLIGKDSLKRSDENQVDAMMKASLLYEQQGQKDEALHLFDKAASF SQGIASHN

LGVIKFKEKDFNGALDLDSSIASKENASVSAIDALVSAYHLQDEDLYYHYLKIVRDTLY
KDYKKSFYSYAYALKSYAGEYFEALSPLMHPNSNAFLKPNARLASKLFLMFKDEDNAYK
QLQKSANAQDELALGLLQARLGHYKQALEHLQHYLHNPYKDLNALMALELVSLKKGDTLK
ASEALKLASHTKEDSVLANSFYPIKPTINPVFLDKERAKERFWNTQYFEGKRDFIYRLLF
YYAPFKVLDSKETLGVIEEGLFLLSDAQKDLEGASLAFKRGRMLAIADKNALQGLKALE
NKCLKKALSFFDLSLKNSPNNALLHYNTGLIYAQLENYHKAYFHFLRAFHLNSADYLSAI
FAILASHFTHEDTTEFLREITENFYSHDFSSPTQKALLSSLIAYLNYRTNWDMDWLKNAP
KKLPFYALEAVFAKESKDKKLMVQSFGNLKKMLPKDLISNIFYEIVSYDASIRHTLSI
YTLLDSRKISWDQTMQGPILGRHFYTYMGFMVNDLDHQERLLEQKIASLEEREAPNDWLE
NLALVSLFQGGYKASALYQNLISDLKDNETRLKILAGLTYIAQSNYNNAALWLELGKLD
DPNNENIRYALGLLYQQRKGDLSALNHFLAIKTSDFSSPYFDDEIDANLLKERLNQEEKG
EFLE

>PROKKA_00277 Serine--tRNA ligase

MIDRKLLQLQDFEMVALSLKKRNSAMDDGLERLREVITRYKKQLIELEGLQAFQNKVSKEF
GIKMAQKADASDLKKELENNKIKLNELSKSVGELEQQIDLKLSIIPNLVDEKTPLGASEE
DNIEIKKILTPRNFTFKPKEHFELAQQNGWIDFESGVKLAKSRFSVIRGFGAKIYRALIH
LMLDFNEKNGFEIITYPALVNEKMLFGTGQLPKFKEDVFKIENENLYLIPTAEVTLTNLY
NDTIISVENLPIKMTAHTPCFRSEAGSAGKDTRGMIRQHQFDKVELVAITHPKESDVMQE
YMLSESASEILKALELPHRFVQLCSGDLGFSASNTIDIEVWLPGQNCYREISSVSNTRDFQ
ARRAKIRFKENQKNQLAHTLNGSSSLXFFSGRQDDGRFNGKPPASGWEHSHS

>PROKKA_00278 C-N hydrolase family amidase

MRVFALQLESFKETLMQSLFNSIPERSVIVLPEYVINPFFHHNMGLDVNEISNQSKRAIE
FLSQKCEELDLIISAPVLEERSKIYKKIALISKENIQYYTQQRLLIPYPHWDEESFFDNE
KSAFKELLVFERDGLRIAPLFGFEAHFDEIWVQAKNQGVVVLLSSVATFESNERWRLLC
QMR AFCASCVVVRANRIGAYRQIIVEEDQKNEFLWKFYGDSFVALPNGAIEDSLEGKMGA
LNAQIDKNEIDEWAKLWHFRTIKEG

>PROKKA_00279 exodeoxyribonuclease VII small subunit

MQDELFEFETEKAPQKNTKNTKNAKNTPKKSFEHVSLEQAIDRLNDPSLSLKDGMDLYKT
AMQELFLAQKLLLENAYLEHEKLQTPDKKA

>PROKKA_00280 Ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE

MKKEKHLKQEKIINMFDDIASSYDQANRLMSFGLDVKWRERACEHAFLFLENKKALRLVD
VACGTGDMLVAVQKSALNCAIEFKECLGIDPSNNMLALAEKKLENKASFIQAQAKDLKGV
GNNSVDILSIAYGLRNVERQEALKEFFRVLPKRGVLVILEFLKKDNPTWLDKISGFYTN
KVLPLVGGAIISKNYGAYSYPQSIEGFLSLEGLKHELKNAGFEILRTEDSIAQISTTMLV
RKS

>PROKKA_00281 Copper resistance protein D

MGFLNGYFLWVKAHVIAVISWMAALFYLPRLFVYHAENAHKKEFVGVVQIQEKKLYSFI
ASPAMGFTLITGILMLLIEPTLFKSGGWLHAKLALVLLLAYHFYCKKCMRELEKDPTR
NARFYRVFNEVPTILMIFIVILVVVKPF

>PROKKA_00282 IMPACT family member YigZ

MPFDDFEKTLQLKKEHFKAHFVTAFRYSLEGKITEGFSDDGEPKGSSGMPMLSVLRRK
DLINIGLVSVRYFGGTLLGVGGLMKAYAKSALLCVENAQKENALKDFVELETLSAHYSYK
ELDALQREIKKFSQLSKNFSNQSVVEVEISGARENLAFLQONKIN

>PROKKA_00283 Inner membrane transport permease YbhS

MNFFKILLMELRAIVSHKGVLLILIGAPLIYGLLYPLPYLRDIVTQQKIALVDEDNSFLS
RQLAFMAQSSNELEIAFFSPSMLEAKKLLKEEKIYGILHIPSHFEANIHKQVPVTIDFYA
NSNYFLIYGALANAVVESINALNDEIRFKRNAQIEEAELGTDGIKIRPIALYNPSEGYLN
YALSSVFIFILHQVMLIASSMFTSSRLELAFLDKKQIALRLCARLLVFMGAFSVFILLY
FGALFSFYGIERHGSALMVFLNSLIFMLATLSLGSFLGAWIKNEAHTTQIVLISSPLIF
MMGFVWPFESLPSYLQVFVQIVPAYHGISLLGRNLNQMHAEFIDVSVHFYALIAIFIVSFI
GSVFKLSSLKKACENA

>PROKKA_00284 ABC-2 family transporter protein

MFRLISAWVLQDKFLFVVCFILPFCLGVLGTQIFKQEIPRQLPIVVVDLKTSTTSHQVAF
ELGATSALQIKHQVASLLEAKRFLNSAEAYGALVLPRLDLEKKIKMGRKVDLPFYNAEYV
LVGKTLKNAFLQTALTDAKTALATKALVRDSNLILAKAQAMPIVLQLHALYNEENNYTQY
LLSVMLPCMWLIFIAIGMLNFIQKISNMRELLISILANACVFSFWMGMMAFYFNFIGMEG
NYAHLVVFLAVVLMALIMSGFVVVLYGISKSVIETAGAGVYTAPSAFAGVTYPQNNM
EIFGDFWSHCLPISHFMKFFLQEAAYKTDFTESLNSLMPLVFFLIFLALGLLVFYFSFKK

DKASA

>PROKKA_00285 Multidrug resistance protein MdtN

MSNSMLDKNKAILTGGGALLLGLIVLFYLAYRPKAEVLQGFLEAREYSVSSKVPGRIEKV
FVKKGDHIIKKGDLVFSISSPELEAKLAQAEAGHKAALKALSDEVKRGSRDETINSARDVWQ
AAKSQATLAKETYKRVQDLYDNGVASLQKRDEAYAAYESTKYNESAAAYQKYKMALGGASS
ESKIAAKAKESVALGQVNEVESYLKDVKATAPIDGEVSNVLLSGGELSPKGFPPVLMIDL
KDSWLKISVPEKYLNEFKVGKEFEGYIPALKKSAKFRVKYLSVMGDFATWKATNNSNTYD
MKS YEVEAIPLEELNFRVGM SVLVTIKP

>PROKKA_00286 Outer membrane efflux protein

MKKTTLFVLSLLFNSSSLNAV D K V S K T D L S S L N L A E D S T P L N H P D A Q K L S L K N A W T R V L S N
HEGLHAQEYAIKRASKMKLAALKSFLPQIDLSAFYVYLSNPIKMDFTSQKQPGVQKATNQ
IHQGIQNIQQNIPPQVLTLPQIQAGMQGVMQGFALSSSTLEAPLLFSKQNVVLGALSIIYP
LYMGGARFTMVRIADLMQKDANEVYRLKKLSTFQELVSVYYGMVLNAEVAETLEEVEKGH
YKHFQNALKMQKVGQIARVETLGAQVAYDKAHIASVKAKDVLVLSQLSFNSILSSKDDLA
PSSKLEIHTKLNLPDLSFFVASTLNSYPALKTLENQVQISKENTKLQIAKFLPQVSFFGS
YLMKQNNVVFEDMIPSWFVGAVGRMPILSPTGRIQKYQASKLAELQANSEQIQAKKNMEL
LVNKTYKETLSYLKEYKSLSSVELAKENLKLQEQAFLQGLSTNAQVIDARNTLSSIIVE
QKSVAYKYIVSLANLMALGDHIDLFYEFVY

>PROKKA_00287 Magnesium and cobalt efflux protein CorC

MGRNQGAYLDPSESI L M L M V A F L L V L L N A F F V L S E F A L V K V R K T R L E E L V K I G N S N A K L A
LKMSQRLDTYLSATQLGITLSSSLALGWVGEPAIAKLLAALFESMDLRENPIFIHMSVVI
AFLSITFLHVVLGEIVPKSLAIAKSEKAALFAARPLHVFWVVFYPVVR LFDVIAHFFLKK
VGINPK EHDGTHSEEEKIIVGESLREGIIDSVEGEI I K N A V D F S D T S A K E I M T P R K D M V
CLDEENSYEENIDIVLKS R F T R Y P Y C K G S K D N I I G M V H I R D L L S R S I F T P K M H D F K Q I V R
KMIIVPESASISQILIKMKKEQIHTALVIDEYGGTAGLLTMEDIIEEIMGEISDEYDLKQ
EGINKLEEGVFELEGMLDLESVEEALHIEFDKECEQVTLGGYVFSLLERMPMEGDTIVSH
GYAFEVLSVDGARIKRLKAVKQDQGENEA

>PROKKA_00288 Sulfate permease CysP

MEIKNIKEFEKASKKLQKDTLKIALLFLIGAALLALIFGQANSKGLLLIFA AVIGGYM
AMNIGANDVSNNVGPVAGSKAISMGGAILIAAICEMLGAI IAGGEVVSTIKGRIVSPEFI
NDAHIFINVMLASLLSGALWLHVATLIGAPVSTSHSVVGGIMGAGMAAAGMSAINWHFLS
GIVASWVISPLMGALIAMFFMLIKKTIAYKEDKKNAALKVVPYLVALMSLTFSWYLMIK
VLKRLYAVGFEIQLACGCILALLIFILFKRFVLKKAPQLENSHESINELFNVPLIFAAAL
LSFAHGANDVANAIGPLAAISQTLEDANSPMGNTLSSVPLWIMVVGAAGIALGLSLYGPK
LIKTVGSEITELDKMQAFICIALSAVITVLLASQLGLPVSSTHIVVGAVFGVGFLRERLRE
QSRRRFARIRDNIVAAHFGEDLEEIEGFLERFDKANLKEKSLMLES LKKS KNTAIALELK
KKEKKS LKKVYKEEVIKRSILKKIVTAWLVTVPV SALLGALLFVALGFIEKYF

>PROKKA_00289 Fe/S biogenesis protein NfuA

MIEFSDEDLQKPVRIVIEKIRPYLLKDGGNIEVLGVKSMKIYVALEGACKTCSSSKITLK
NVIERQLKMDIHPNLEVVCLENAKEFDKL

>PROKKA_00290 hypothetical protein

MQKF D Y E F K K R A L I K E G F L A F K Q A H Y A E A L R L F S E V L F L D K D N Q K A K V G A L L S D I A K D F P
KEAHSFYELYQSLIAMQKRSLKNQAEEQIINLIASFDEGLNQMAEKIDAQISQKSEELNG
ILYADFKRLSLERGFKEAFEDLMFSSRVIFDNKEDFYEF L K E L N H Y G Y E L A I N Y I E N M H
EDSFIYDKFLRSLLEDALKSNKA

>PROKKA_00291 UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--L-lysine ligase

MKLKKTLT Y Q N H T Y S F L S D N T H E V L E N P K E I L F V K T P L N E K Y S P F I A E K N L A I L D F N E L K
NYFDFKIKIVGITGTNGKTTTASLMYSLLLDLNKKTALLGTRGFFINDERIKEKGLTTP
LLELYSDLEEAMRLGCEYFIMEVSSHAIVQKRVAGLDFALKILTNTSDH L D F H E S I E N Y
RDAKNSFFKDEGLKVINRDETNA L F N P I N A H T Y A L D K K A H L N V Q A F S L T P S I S A S L C Y Q Q
DLRDPNLKEIALMHSP LLGRYNLYN I L A G V L G V K L L T Q L P L E T I A P L L E N F Y G V K G R L E I
VH S K P L V V V D F A H T T D G M Q Q V F E S F K N Q K I T A L F G A G G D R D K T K R P E M G A V A S Y Y A H K I V
LTSDNPRSENEEDI I K D I L K G I S N S S K V V V E K D R K K A I L S A L E N L K D D E V L L I L G K G D E N
TQIFKDKTIFFSDQEVVKSYYQHLKQG

>PROKKA_00292 Transaldolase

MQEFSLWCDFIERDFLENDFLKLINKGAICGATSNPSLFC E A I T K S A F Y Q D E I A K L K G K K
AKEIYETLALKDILQASSALMPLEYEKDPNNGYISLEIDPFLEDDAAKSIDEAKCLFKTLN
RPNVMIKVPASESALEVISALTKASIPVNVTLVFSPKIAGEIAQILAKETQKRAVISVVF

SRFDKEIDPLAPQNLQAQSGIMNATECYYQISQHANKLISTLFASTGVKSNSLAKDHYIK
ALCFKNSINTAPLEALNAYLLDPNTECQTPLKITEIEAFKKELKTQNIDLENTAQKLLKE
GLIAFKQSFEKLLSSF

>PROKKA_00293 50S ribosomal protein L25

MLEGVIRE\$ITKANAKALKKDGyliANVYgKGIENVNGAFKLNPFiKYLKEKKHLIFPVK
LGDKTFEVVVQ\$EYQKNPVTNELIHVDLLAVTKGVKFKFKVPVKHQGTPVGLKNKGILMLS
KKRISVECAPEHLDPDHYLVDVAPLDVNESILVRDLEKHENVKILDHDSIAVIGVIKAK

>PROKKA_00294 Peptidyl-tRNA hydrolase

MTLLVGLGNPTLRYAHTRHNAGFDILD\$LI\$ELNLSFTTFSPKH\$NAYLCVYKDFILLKPQT
YMNLSGESVLSAKNFYKTEELLIVHDDLDLPLGVVRFKNGGGNGGHNLKSIDLLCSNSY
YRLRVGISKGIGVIDHVLSKFHKNEEPLKNAVF\$HAKNALKFFIESHDFNAMQNRFTLKK
PLKIES

>PROKKA_00295 putative permease YjgP/YjgQ family protein

MRLFRFVGWYYFKYFLIVLLALELFFVGIDSLKYADKMPDSANMIILFFTYDILFALNYT
LPISLLLAMVLFYISFIKSNQYTALLSIGFSKCQILSPIFLISLFFTA\$VYVGLNATPFVY
MEEKTQNLIIYKD\$NLSVSEHLLVKYNDDYVYFDEINPLLQKAQNIKVFRLKDKTLESYAE
AKEAFFEDKYWILHDTTIYEMPLSFELGANALSTTRLKTFKTLKNFRPKVLDTIYQNKPA
VSITDALLSLHALVRQ\$ADTKKVR\$FLYVFAILPFFVPFLSVLIA\$YFSPSLARYENLALL
GLKFIIITLVVWGLFFALGKFSISGILIPEIGVLS\$PFFLFLSLSLWYFKKLNKRL

>PROKKA_00296 Modification methylase FokI

MEKFSLK\$NRRYIGSKTKLIEWVFGNLKLN\$NIKSVC\$DIFAGSGVVAGQFATIPNVKNI\$IN
DILFSNEVIYH\$AFFMGQAD\$FVKELKEYFNQALKLEENYFSQH\$FSDKFFSYKDCVKIG
SIREHIENLNL\$DKLNKDILLTSLIYSMDKIAN\$TVGHY\$AYRKKEILQDRFIFELISPIKH
DKNIMIERKNANELAKTLKIDLVFIDPPYNSRQYSR\$FYHLYENLVQWKKPKLYGTALKPS
CENMSEYCRSNAK\$ELNDLIEKLDCKRIALTYNNTYNSKSSSSQNKIGFKDLVEILSRKG
KLSVKEKAHSFFNSGKTD\$FKEHKEFLFIVEVKP

>PROKKA_00297 Modification methylase FokI

MKKGIRSPFFYVGDKYKLMPQLNKLFPNNINQFIEPFVGGGSVFLNTKAKRYLANDIDTN
IINLHKTLSKFNTCELFDELSKII\$IH\$YGLSFS\$FKGVTAPDELKKQYIKTY\$YAKYNKIAYE
KLRADFN\$NQNMLYLYLLLIYGFNHMIRFNSKGLFNL\$PVG\$NVDFNENVYNALKNYIDFI
QQNTIIFHNGDYIDFLNHTTYLKDDYVYFDP\$PYLISNSEYNKLWDS\$NEIALYGVLD\$LD
KKGVLFGITNLVYHKGETNFILKEWAKKYIFNIKSNI\$SYNDNTIKEDS\$QEIFVTNYR

>PROKKA_00298 AlwI restriction endonuclease

MTKKPARKILSFSTMRNPKRIGQFLAVLGKFENQILQSSTIMQIIKS\$VLAHRLYRPTSI
NQN\$KELKEKFDSNKYIFSDEELERII\$EIS\$PQQHKEMGF\$EHGWESRFDTWYKLMCEFGFCY
YAKYEKILISDSAKMLILAYYDKENDTFKESVDES\$VGAIFLNALSKYE\$VGNPYKKNL\$NH
NNP\$FKLLLSLLKRLKNAHLTPLSVKEIPILLCWKDDNAKGLYDYIIRLRQEVVTINKTEF
SYSDEFIYEKCLNLLKSVN\$KTRFKMSQITNEAVDEYIRKMRITGLISLRGN\$RFIDINTN
ENNKIDYILQTHKAFKGDYLN\$DTQANKL\$AFFNYMAIVDSFLVSVTPISTDES\$VKSSKLNE
LATTYTKDFIKQELLIACNKQ\$ESKDSFLRLIDKPLRLEFLSAIFLKQH\$FENLSVIPNYKS
DDEGLPVYTASDNKPDIVAMDTKAQSYIEVSLIRDRSQSTLEMIPIARHLKELIKNSTDI
REKFSV\$FVAPNIHDDAKE\$YAEFAQFKDNINICCYAINDFIKKVEN\$AEWLQINDDLKA

>PROKKA_00299 Helicobacter outer membrane protein

MVGKNFNRKESV\$LIAQSLDISKKG\$SVILGALLSSLWLTNPLNAHEKNGAFV\$GISLEVGRA
DQKT\$NAYKNGELFQVPFGDVSANDDGKVPD\$GQTGGCQ\$PASGTPGTPGYTKANCVVNWTSR
TMLSTNKNIPGRNQPMYGLG\$VMTGYKHF\$IGKKRW\$FGLRYYGFFDYGHTN\$FNS\$RAANAIS
PFYLS\$DQKADMYTYGFGTDMLFNIIDKPKATAGFFLG\$VNFAGNTWTNNRVGYFKDGYVYG
VNTDADAYMTNADGTITCGDTPASCNVGINLNSVYATGKLNAKVNHTIFQFLVNVGVRT
NIFEHHGIEFGIKIPTLPNHFFKGSTTIRAKKQ\$PLENGQPTTITGAETN\$FSLTQTLRRO
YSMYLRVY\$YTF

>PROKKA_00300 hypothetical protein

MDAIYPYVLVVHLLCAIIFIGYLF\$FDGVI\$FPNVK\$MFGEEFANKANTGITQRAIKIMPLC
VLGLVLTGGMMLSQYMG\$DGKWCETPFQKILMLKVILALSIFLLVLFSL\$CKFLGKKNPI
GKYIHPIALT\$FGFLIAILAKTMW\$FV

>PROKKA_00301 Copper-exporting P-type ATPase A

MKCSHCQLEFKESELFKEVIN\$H\$KELH\$FCTG\$CARVYALLLDLNL\$ESFYDKLNDSTLAPVT
PQDSMNALELEQALEENNKGDFILNLLLEKTHCNA\$CLWLNQKVLERLKG\$VKKVSVNFTTH
HLQIVFDKSLNPKEIIQKIESLGYGAKIYN\$AKNYALKAQKEQ\$RSYLLT\$LSVGFFATM\$NLM

FIAIAKYASYGGASYGGASYGAGMDKLMQRNLDLVSLLLSLLVLVVVGRFFIKGAFYGLK
NGVLGMDLSVSFGALSAFVYSVYAMLVSQETYFEASSTILTLVFGSKFLELKARLFANEK
CLALESHEIHSVIVVEKDKQIEKHPKDVAIGSVVWVPSGAKIALDGA LLNNASVDASLIS
GEFKPLELGVNDPILGGYVNVGAPFSYQVSATFQNSRSLG LLET LKKS FLEKPLIESSAN
QIADIFSKAVLFLAFVSFLLWQFGLGGNFEKALMVCISVLVISPCAFALATPIALVIGV
FKNPLIVFKEALFLET LAKVEKIFIDKTGTLTQKEVLLKEKIIYEK FDERLLKSLLKVRE
HLAHSAILKSLDGDEVNLEKIEFFAHGLKANYQNETLLVGSLKFLNAMGVNIKAKESANI
MVGFAKNKTLCALFILEERLKNNAKEVVQALQNQGLELEILSGDNESSVKECAKKGISK
YHANLTPEDKAQIISSYKVCAMIGDGNNDALALKQASVSLGFEKSALS KSACDILLLEE
DLSLLKKA FDNAQKVYQVVLQNI VLSLIYNAVLI PVAMLG YINPLIASLSMSASSLLVLV
NSRLRKRS

>PROKKA_00302 tRNA1 (Val) (adenine (37)-N6)-methyltransferase
MDRKLRLRYQPLNAYSNSDSLFLYDFSRPFIKNSGAILDIGSGC GILGLLCARDNPLAS
VHLVEKDSKMAFCSQKNALKFPNAQVFEGDFLDFNPPILYDVIVCNPPFYALGSIKSQIK
GHARHQSELDFASLVAKVKCLKPKGYFIFCYEALSLCLVIEGLKSVELTLET LRFVQSF
KDKNAHMLMGAARNNSKSALKVLPPLITHHSKNQSDNTKEVLSIYQICNTYSIKALLN

>PROKKA_00303 Riboflavin biosynthesis protein RibD
MRLYESLLEMCLNKAWEQTLALENPSVACMVLDKHHEILSLETHKKAKTPHAEVLA AKS
ALKILRPSLKN DLEKLEDPKT L SDF LKTHHDNAFKDCVFLIT LEP CNSY GKT PACSELLE
ILKPKRVVIAAEENEAKKGG LARLQKARIETIICHNLENKAKD LLLPFRIMEQKGRFNLF
KLALRMNGDYHHGKITGQKSVIFTHNQRAICNTLIISGKTIRTDNPLLDARFCDHFYQNK
NPNIAILSKRSIDPNSKVFSAPNRLVNTFHDPKDLPLEKGFNFIEGGWELFESLRNKIDA
LLLHSHASMIGEAFSVLALKTPFKGRLLHAQILENEALLWIENS

>PROKKA_00304 Sodium/glutamate symport carrier protein
MQEIKLDIYATLVCMVLVLLLGRYVISKVKFLRDYDIPEPVVGGVLVAFFIMLARQFYNF
GLQFDSS LK DPLMLTFFITIGLSADFKSLQKGGKMLAVFLLAVAGFVVCQNMVGISIASL
LGVNPLMGLLGGSIALVGGHGTSAAWANFFTQPPYNFSSSLEVGMVCATFGLVSGGIIGG
PVAKYLISKYKLEPKDTKEKDTLEGVASKGFETPKEQRLITASSFVETLALIAIALLVGT
FLSHLMPKSFTLPTFWCLFVG VILRNALSFFKIH SVFDREVS VIGNVSLSLFLAYALMS
VNLELLKLAVPLAVILSVQVVFMI LYVVLVTFRVCGKDYDAAVLCAGHCGFGLGATPTA
MVNMQTITNHYGPSHVAFIVVPLVGAF FVDIINALAIKGFLLLPFFPS

>PROKKA_00305 Carboxynorspermidine synthase
MHTVLQIGAGGVGSV VTHKMGMNRDVFKNIVLASRSLDKCYAIKESMLKKGLGEIGVEQV
DADDTQALVALIQKYKPKVVVNVALPYQDLTIMQACLET KTHYIDTANYEHPDLAKFEYK
EQWAFDKAYKEVGILGVLGAGFDPGVTNAYVAHAQKHHFDTIHTLDILDCNAGDHKRPFA
TNFNPEINLREVSSKGRYYENGKWIETKPLEIKQVWAYPQIGEMDSYLLYHEELES L VKN
IKGLRRARFFMTFSQNYLTHMKCLENVGMLGIKEIEHQGVKIVPIQFLKTLLPDPATLAK
DTTGKTNIGCYMTGIKNNQDKTLIYINVCDHKKCYEEVGSQAISYTTGVPAMCAAKMICN
DTWSADHFRAGVFNIEELNTDPFMEELIKQGLPYEVIER

>PROKKA_00306 Putative electron transport protein YccM
MLESSSHFLKS FRLKRYIGFLLISLALLITPFI RIDGAHLFLISFEHKQLHFLGKIFSAE
ELQVMPFPMVILLF IGIFFITSLGRVWCGWACPQTFLRVLYRDVIETKIFKLHKKISNKQ
ESPKNTPSYKIRKALS VLLFAPVVAGLMM LFFFYFIAPEDFFMYLKNPSEHPVAMGFWLF
STAVVLFDIVVAERFCIYLCPYARVQSVLYDNDTLNPIYDEKRGGALYNNQGH L FPLPP
KKRSPENECVNCLHCVQVCPTHIDIRKGLQLE CINCLECVD ACTITMAKFNRPSLIQWSS
TNAINTRQKVCLVRLKTIAYMGVIAIVIALLAITSFKKERMLLDINRNSDLYELRSSGYV
DNDYVFLFHNTDNKDHEFYFKVLGQKDIQIKKPLNPIAVKAGQKIKAVVILRKPLKSNAI
EYKRAKDALIP IIIQAYSTDDKNITIERESVFITPSED

>PROKKA_00307 Glycerol-3-phosphate acyltransferase
MESVLNFLTININVI FTLLGYLIGGIPFGYALMKIFYGMDITKIGSGGIGATNVLRALQSK
GVSNAKQ MALLVLILD LFKGMFAVFLSKLFGLDYSLQWMVAIASILGHCYSPFLNFNGGK
GVSTIMG SVVLLIPIESLIGLTVWFFVGKVLKISSLASILGVGTATVLIFFVPYMHIPDS
VNILKEVGTQTPMVLI FIFTLIKHAGNIFNLLAGKEKKVL

>PROKKA_00308 Dihydroneopterin aldolase
MKTQGVHVHNLVFETILGILEFERLKPQKISVDLDFYTEL PNKAYLDYMEIQELIQKM
MQEKQYLLIEDALEDSLQILKTRYSEISELYLKISKLEISPDSQVGASVKIRYESNL

>PROKKA_00309 hypothetical protein
MVLKAHPINPLLEPLYFPSYAQFLNLEPHFVVKKKRAYRPFQWGNTIIIKRHDLEERQSN

QPSDIFRQNAEINVSSQTFLKEMSNASSQIALDSTAQ

>PROKKA_00310 Heme transporter BhuA precursor

MLRNQFRIVFVSCIVASNLQAQETHTTLGKVTTKGERTFEYNNKMYIDRKELQQRQSNQI
RDIFRTRADVNVASGGLMAQKIYVRGIESRLLRVTIDGVAQNGNIFHHDANTVIDPNMIK
EVEVIKGAANASAGPGAVAGKLSFTTIDANDFLRKNQTYGAKAEAAFYTNFGYRMNATAA
YRGKNWDILAYYNHQNIFYYRDGNNAFRNVFHPNYDLQDPSNSEMSVGTGPSEVNSVLAKI
NGYINETDSISVSYNLTRDNSTRLLRPNTTSALSANDPGSQPAPFVIDFGKELAHTINF
NHNLSLKYKHEGGPNFNPRIESTAFGLVVRGGNYPVVPNFAYNSNEPANPDYIPEVREW
CNNPDNISQCTQGAIRPSNGGYQIGYGAPNSINWQGDSDSSGGAQAGYGQLSAISTSANV
YHGLVPKNPDYDMTPPNAQNPANANDWTLGNADEGLARRIFLINSNVNFKVTHPISEY
GNVFEYGMIIYQNLVSFSGLDKGNNGYYKNNIDPNDPNGPGLPYRHYYTDQSSQYPQNLNT
PNPLYRNMPQNSHAIGNIVGGFMQANYNLSNVIVGAGTRYDIYTLDDKNGRTHVTSGFS
PSATVLYNPIESIGLKVSAYVTKGALPGDGVLMRDPTVIYQRLRPAIGQNVFNVDFN
SKYFNVRGAIFYQVINNFINSYGQDTSKNGGNNATAKNMSGNLPETINIYGYEVSGNVRY
KNFLGTFSVARSWPTARGHLLADTYALAATTGNVFILKADYDVRWGLTLTWLSRFVTNM
YYEGYSIYYYPQYGLIKIHKPGYGVHNVFINWTPPSKKWQGLRISAVFNNILNKQYVDQTS
VFQASADAPASDMI PKGKRMALPAPGFNARFEVSYQF

>PROKKA_00311 L-seryl-tRNA(Sec) selenium transferase

MANITTKETLLITPDLKSPYQKIINASASVFDEKHGRSFFSPQFYEKIEPYLKEVLTQP
IGLECDLNTAKKTNRLTPLKQLFKACFGTEEILIVNNNTSAVFLIANALAEKEIIVSYG
ELVGGDFNLKDILLSSGARLHLVGNINRAYLRDYRLALNENSKMLFKTHNPAFKKDTPFK
DLQALAKEHDLIDYYNLGDVLLDRTALEEILALKPSLLSFSADKSFNSVQAGIIMGQKE
WVEMLNHPLYRALRAGKITLTLLFCSLNAWVNHQEEITIHALLNQTKDALLQKALKLYA
LLKPLELNVSIIASSFSKIGNLFGRELESFCVKIQPKNTRALNGEKLYLRLFQKGVITRIS
CEFVCFEVFSLNEKDFEKIALVLEEILNKA

>PROKKA_00312 hypothetical protein

MEKISDLIECIAYEKNLPKEMISKVIQGCLLKMAQNELDPLARYLVVEENKQLQLIQLVE
VLEDGDERLINDPSKYISLSKAKEMDPSVKIKDELSYLSLESMSKQGAINRLFKDLQYQL
EKALEDSEHFAEFQKRLNSVLMGQVILVDHNQNTFIEIEQQFQGVLSMRHRIKGESFKVGD
SIKAVLTQVKRTKKGLLLELSRTTPKMLEALLELEVPEIKDKEIEIIHCARIPGNRAKVS
FFSHNARIDPIGAAVGVKGVRRINAISNELNKENIDCIEYSNVPEIYITLALAPAKILSVE
VKKIPIEELSAAEEKESVQERFIVNNHLQAKVRLLDIEKSKAIGKGGVNVCLASMLTGYH
IEFETIPSVKENAENENEKETPKVGVEALESFKN

>PROKKA_00313 hypothetical protein

METKEEYKDKKLEEMIVSLCEKGDLSQTDQIIKDLKEIYQGEYRHKYSKITTIILNSTR
DREQAFMTLTQNIRTLKEIQDNKEVESIKPKLEKLYDHMNLECIRLQDFDEKMSRIKDV
IRLNDLNKNYKKLSEELNKQQTQYITILGIFASIVLTFVGGGLAFSTSVLSNIDKANAYRL
VFVMAFMALFFGNILYLLFSFLSKISLSKEEKDKQEIFSKNPCFGLI

>PROKKA_00314 hypothetical protein

MVNALELAKYILKHSDKELSNLELQKTLTYFTELDYIKKFDKHLVSDDFEAWRYGSVAREV
YYEYRNYGANPIDKPKETLSQNLRKDELETINRSIEKCNKKS YWDLVKESHKEDGAWHK
SFKEDRKEIISKDLIRQEAQTNGN

>PROKKA_00315 hypothetical protein

MTDEPRFIYDNKGFLTNTQTCYFIARDDKYLFAILNSKVMYFFMRQMASNLGEGAFRWIKQ
FIERLPIPIQITEKNQELADKIIALVDKILALKEKDPKANTQRLEKEIDALVYQLYHLTDE
EIKIIEDGQ

>PROKKA_00316 Modification methylase PaeR7I

MIRFAHIPLKDFIKKHNPQEPKKETIENFEKEINSLLNAPRQDDEEFQKNEINKFLKNT
YNHDCNTNKKVDSAIYVGEEVWVLI EVKALNNRNEFPKNRENPLSKAFCQMVLYFLEEIE
KEKNNSLKHAII CNAHEFFLFDCDLLNLFQNDKRITKLHKDCAKKEGTD PSTKRFYSDL
EEYLKDKDFKELRYTHFNLSDDFKELPLIYQVLSQEVLLKQRKTL DANTLNKDFYEELLY
ILGLEEQNDKGKILIKPSRTKNSLS DALKEQYKNLDDEEVMALLIAWNNRILFLRLLES
LISFNHFNHFNHFNHFKESFLTTFENFKDFNALNTLFFEVLAKKNSERLKEIKENKILQK
IPYLNSSLFDKTPLESKGHEIKLLDNKPLEIYPKSVLKKNEDYQKEKDLPLLKYLFEFLR
LHKFTTTPKDIKDNTDTS ENLNINPSVLGLVFEKLNKYKEGSFYTPSFITSYMCKESITP
IVRDKFNQKYNIECEKLEELKNYLKSYKEDKRKEYLQLLLTLRVCDPAVGSGHFLVSALN
EMVLIAYELGLIASLHRYDLKLENDEII IHYKPTGEIFNYKKPNSENDPHHQIQKELFEL
KKSIIENCLFGVDINPNSCEITKLRLWIELLKYSYIIFEEGKNTNNLETLPNIDINIKCG

NSLISRFAKDKALLKTEKNKNLEYIIAEYKELVKIYKDPKILESLTRPIKDSNAVRKYA
KERLYQELAQNPNDKFKKALNDRIEKIKEAFKLTLEPPQKELKFKKFLKEHLELYGKSIL
EEANDNGLELEALALEKKMAHEGLFHDYTPYPKLDKTDKVVGLEHFNRYVLTSYKDLQDE
NERYANALEWRFEFPEVLDDDEGNFLGFDCIIGNPPYIRQEHKDLKPLLEKQYQDFYNSS
SDIYTYFFALSYYHLLKEKGFSAFITSNKYARAKYGAKLRELLLKKTTLVSYMELNALKVF
ESAADVTSIIHFQVFPKESAFNYEPTQNDKENLNKTPSLLMKQNALSTESFIFANAT
LLDLRDKIESVGTPLKDWDIQIYRGILTGANEAFFIIPTEKRDAILNACKTQEERERTERL
IKPILRGKDIKRYLMSGRICGLSTPIMATLLISNLKSLPLI

>PROKKA_00317 Type III restriction enzyme, res subunit

MKIKFKRLDYQEQRDQILGVFKGIDLREPENDAQRIANPVFEIGAIDVLLNENIENLRS
KQKITQGSVGIEKSLNCDILMETGTGKTFCFLECVALHKNYHLSKFIVLVPSNAIKLGV
LKSVEITREFFKSEYSTHLESYEDVERFILASNHKCCVLVMTFSAFNSEKNTINQSCLEN
TNLFNGAKSYMQUALASIRPIVIMDEPHRFLGDKTKKYLEQLNALLTLRFGATFKDDYNNL
IYALDSKKAFCALVKSISVASVGESDECFLLEKLGIEKRNGYEAAINYTDLENKTQSVKV
KTHDNLGVVTQISALEDYIVEKITKTEIRFLNGFNLLLDQKEPFSHLEGEQEVMLEKAI
KSHFEREEGLFKKGKICALCMVFISGVNSYLSENEKPAKLALLFEKLYQQKLEEVLLKKPLD
ENYRAYLERTKDAILKVHGGYFAKSKKESDEAQVIALILKEKEKLLSFESDLRFIFSQWA
LQEGWDNPNVMTICKLAPSHSNITKLQQIGRGLRLAVNDKGERITKEHADFDVNELVVI
VPQVEGDFVGAIQQEISEHSLIKQAFSAEELEKSGIVEKGRYGFLFETLESILGFGEKTGD
ENFKLTLNQNEFLEKEPELERLKDEKYLDLEKLKDFLKDRLIGNSRVRKKNERKSEKIKI
NKENFKKFETLWESLNHQARIAAYAIIDSESLIDEIVKNINSSFNVKSIVSVTTHKKVETM
GNNAKTEMFERESACVWSLHEFISALSNNKVKLSFKSVAKVLEKIDENKFNEIKKNEQEG
RRLEELFLEIIYQNIKDKISYQMRETTIKNRKNDIFYDEKGEIREFLDGSGLGVDKYEIKN
SSAQERCLYENFMQVDSEIEKGMIEESNDTKIIVFGKLPRVKIPIGLNQTYSPDFGYVVE
NNDKKVLLIVETKGVEHENELRDEEKCKISTAEKFFEALKKQGVNIEYKTKMKKDQLSAL
INEILNRKD

>PROKKA_00318 hypothetical protein

MNAAFKKRHFILVQLDEKIDPKDKSTHDFCLNTLTKSTSPSIFDITEERIKRAGAKIKEA
CAHLDVGFRAFEIIDDETHANDKNLSQAHQKDLFAYFNPKKWETQTILIKLLGCEGLELT
TPIICLIENALYLAENTAFIVGDIEMSEVLENLKDKGVEKISMYPAPISNDRCLCELGSN
LFDLKLESGLDKIRG

>PROKKA_00319 putative methyltransferase

MQNKEISGEKSVNEKNVEVFNRYFPGCLSIENDNQLTLDTRRLKALLGDFSEIKEEGYGL
DFVGKKIALNQAFKKNNKILKPLNESTSKHILIKGDNLDAKILRQSYSEKIKMIYIDPP
YNTKNENFIYGDDFSQSNEEVLLKQLDYSKEKLDYIKNLFGSKCHSGWLSFMYPRLLLAKD
LLKQDGVIFISIDDNEAAQLKLLCDEIFGERNFLSSLTWLKNGNAQNDQYFQNNYENILV
YAKHVEALSLNRMASKKEVKVFCENDKYYYEGAGLTGTGGAGGTLNTRANLGYSIYYNPKT
LDLGLVDDYDKKLAKSSNNENLVYNDRQDLLKQGYKIIIRPPKKGVGLGRWTWALDTFNSN
KNIISIKKNSRNEYVICKKEFLDKKQVKQNGENGEFYAILDKSSPTRNVNIENIGGGNGTKE
VNALFNQKIFNPNPKPTKLISRLIELSTNESDIILDFFAGSGTTAHAVLESNKSDYQKLSE
GGGGGYLMA

>PROKKA_00320 ATP-dependent DNA helicase RecG

MQETDNLKLTNLVKSLEALLVYTPKGYKDLNLLERFETGLSGVLEVGILEKRNYAKVLK
IFAYSKRFYKNLELVFFNYSAFHHSQFKTGENLFIYKLEQSSFNQAYIINTPKILTEFG
KISLIFKKVKNNHKKIQENLQKLISLENLKKEGVKENIAHLLLEIFFPTPHFVKDFETNKN
FPSQHLNALKYIEMLFYMNLERKKLQFSAKIACPNNSERLDNFIAFLPKLTRDQQNAI
KEIQSDLTSSIAKRLIIGDVGCCKTMVILASMLAYPNKTLMAPTSILAKQLYNEALK
FLPPYFEVELLLLGGSHKRSNHLFEKITHVIGTQALLFDKRDLENEFALVITDEQHRFGT
KQRYQLEKMASSKGNKPHSLQFSATPIPRTLALAKSAFVKTTMIREIPYPKEIETLVLHK
RDFKIVMEKISSEIAKNHQVIVVYPLVNESEKIPYLSLSEGASFWQKRFRKNVYTTSQONK
NKEEVIEEFRELGSILLATTIEVGISLPRLSVMVILAPERLGLATLHQLRGRVSRNGLK
GYCFLCTIQEENERLEKFADELDFKIAELDLEYRKSGDLLQGGEQSGNSFEYIDLAKDE
ILIAEVKRDFLKAASVSRGTFEN

>PROKKA_00321 hypothetical protein

MKSKITHFIAISFVLSLFSACKDEPKKSSQSHRNNTKTTQNNPINQANKDIKKIEHEEED
EKATKEVNDLINNENKIDEINNEENADPSQKRTNNVLQRATNHQDNLNSPLNRKY

>PROKKA_00322 Helicobacter outer membrane protein

MKCFSKDLFKKVTPFLSVYFLNPTLTQAKSRFYVASQYQVGKMIMKKYNDLKRTIEGAS

FSLGWEINPTNYWFYSRYFFMDYGNVILNKRTGAQANMFTYGFGGDLIVEYNKNPLYVF
SLFYGMQVAENTWTISKHSANFIIDDWRSIQGFSLKTSNFRMLGLVGFKFQTVLFHHDAS
IEVGIKWPPFAFEYDSAFVRLFSVFISHTFYL

>PROKKA_00323 Exodeoxyribonuclease

MKLISWNVNGLRACMTKGFMDFNSVDADVFCIQESKMQQEQNTFEFNGYFDFWNCAIKK
GYSGVVTFTKKEPLSVSYGINIEEHDKEGRVITCEFESFYLVNVYTPNSQQALSRLSYRM
SWEVEFKKFLKALELKPKVIVCGDLNVAHNEIDLENPKTNRKNAGFSDEERGKFSELLNA
GFIDTFRYFYFPNKEKAYTWWSYMQQARDKDIDGWRIDYFLCSNPLKTRLKDALIYKDILGS
DHCPVGLELV

>PROKKA_00325 hypothetical protein

MKKTLCLSFFLTFSNPLQALVIELLEEIKTSPHKGTFAKVLDKSKEPRQVLGVYNISPHK
KLTTLTITHISTAIVYQPLDEKLSLETTLPNRPRTIPRNTQIVFSSKELKEAHAHQMPSLN
APMQKPQNKPPSSSQQPSQNFYSYEPKLGSKNSKNSLLQPLATPSKISPTNEVKTPTNDTK
PPLKHSSSEDQENNLFITPPTKTLPNDTSNADINESNESNENRDNVEKQAVRDANVKEFA
CGKWVYDDENLQAYRPSILKRVEDKQTTTIDITPCDYSTAENKSGKITTPYTKISVHKTE
PLEEPQTFEAKNNFAILQARSSTEKCKRARVRKDGTTTRQCYLIEEPLKQAWESEYEITQ
LVKAIYERPKQDDQIEPTFYETSELAYSSTRKSEITQNELNLNEKFMEFVEVYEGHYLND
IVKESSEYKEWVKNHVRFKEGVCMVLEIEEQPRAKSTPLSIENSKVVCVKKGNLYLFNEV

>PROKKA_00326 Chromosomal replication initiator protein DnaA

MDTNNNIEKEILALVKQNPVSLIEYENYFSQLKYNPNASKSDIAFFYAPNQVLCTTITA
KYGALLKEILSQNKVGMHLAHSVVDVRIEVAPKIQISAQSNINYKAIKTSVKDSYTFENFV
VGSCNNTVYEIAKKVAQSDTPYPNPVLFYGGTGLGKTHILNAIGNHALEKHKKVVLVTSE
DFLTDFLKHLDNKTMDSFKAKYRHCDFFLLDDAQFLQGKPKLEEEFFHTFNELHANSKQI
VLISDRSPKNIAGLEDRLKSRFEWGITAKVMPDLETKLSIVKQKCQLNQITLPEEVMEY
IAQHISDNIRQMEGAIKISVNANLMNAPIDLNLAKTVLEDLQKDHAEGSSLENILLAVA
QSLNLKSSEIKVSSRQKNVALARKLVVYFARLYTPNPNTLSLAQFLDLKDHSSISKMYSSV
KKMLEEEKSPFISSLREEIKNRLNELNDKKTAFNSSE

>PROKKA_00327 purine nucleoside phosphorylase

MLLCAGRNETLKKATPIGVGLIGSAINLTRICLKNPDTENLIFIGSAGSYSPMELLSVF
ESIQQYQIEESFSLNSYTPLDNFIHIETEEALFERVGVNSSNYIHTSEMFAKKMAQKG
VLLNMEFFSVLSVAKAFSLKAKGIFCVSNHVGLENAHKEFKENHAKVKQILENIIDSLII

>PROKKA_00328 hypothetical protein

MFEKIRKILADIEDSQNEIEMLLKLANLSLGDFFIEIKRGSMDMPKGVNEAFFTQLSEEVE
RLKELINALNKIKKGLLVF

>PROKKA_00329 Glutamine--fructose-6-phosphate aminotransferase [isomerizing]

MCGIVGYIGDSEKKSVLLEGLKELEYRGYDSAGLAVLSNDCLEVFKTQGKLENLKSELKN
KEFLDFGVSIAHTRWATHGKPSSANAPHFTENLALVHNGIENYASLKKELENKGHAFL
SQTDEVIAHLLLEETLKSESDLLKAFEKSISLLKGSYAILMLHKRAKESLFYAKSSSPLV
VGKGKEGVFFASSLSVLAPKVDQFVILEENSVGQISLENFKDLKHIENMKDYAFENKDYS
KGDFRNYLEKEIYEQHSSLLECLEGRLEALSVCYCEIDPEFLENVNEITLCSCGSSYHASL
ASVYLFERLAKIRARAILASEYRYAHFKSNPNELFIAISQSGETADTLEALKLAKAQLK
TISLCNAPFSMMSRISDHVLLIRAGVERSVASTKAFSSQVMLLWLLSVYIGKQLGTISKE
EERIQAKNMLNSVNAMKVESKLHEKIKRLSKRYLHGHHFFYIGRDVFYPLALEGALKLKE
ISYLHAEGYASAEMKHGPIALVDSNLFTIALLSKHLDFDKTKSNIEELSARDSTICVLSS
EILEIADDFIQLEESESMEEFFRMNLAMQLLALEIAMRLNHDVDHPRNLAKSVTVE

>PROKKA_00330 Thymidylate synthase ThyX

MEVICKHYTPLGIASQAIRTCWQSFEYSDGGGCKDKELIHRVGNIFRHSSTLEHLYYNFE
IKGLSRGALQELSRHRIASLSVKSSRYTLRELKEVESFLPLNETNLERAREFLVFVDNEK
VNAMSVLALENLRVLLSEHNIKNDLAKYAMPESYKTHLAYSINARSLQNLTLRSSNKAL
KEMQDLAKALFDALPGEHQYLFEDCLKH

>PROKKA_00331 EcoKI restriction-modification system protein HsdS

MFALFYPLLNEQIAIANILSDLDRLYNLDALILKKEGVKKALSFELLSQRKRLKGFNQA
WQVRVLGDICEFGNGEAYETLIVENGDFKLISLNSIDIDGNLKNMTKRVNFYDNSLKQDD
IVMVLSDVAHGDFLGLCAVIPSNDYVLNQRMGRLRIRNDCINILFLRLYINANQKYFKMQ
GQGSSQLNLSKKAIEDFEIPLPLNEQIAIANILSALDHEIASLKNKKRQFDNIKKALNH
DLMSAKIRVLKK

>PROKKA_00332 EcoKI restriction-modification system protein HsdS

MDALTTPSNWQVRVLGDIGITISGLAGKTKQDFINGNAKYITFLNVLNNVIIDTSILENV

KIYPNEKQNSFKKYDLFFNTSSETPKEVGMCAVLLDDIDQVFLNSFCFGFRIFDKAVDSL
FLSYLINSEIGRKAFENLAQGSTRYNLSKSGFNNVCLILPPSKRTNRIR

>PROKKA_00333 Type I restriction enzyme EcoKI M protein

MAIKKSELYSSLWAGADSLRGMDASEYKNYVLNLLFLKYISDKARNNNFSEIEVPQGCF
YEDILALEGDKEIGDKLNKIIAKIAERNDLKGVIDSVDNDNTKLGENKAMIDTLSNLVK
IFADLSLGAHGALDDDLLGDAYEYLMRHFASESGKSKGQFYTPSEVSLLLSSLLLGIDENT
RQDKSIYDPTCGSGSLLLKASSLAGKNGLTIIYGQEKDISTTALCKMNMILHNSADADIAK
GGSSTLSNPLFTTENGMLKTFDYVVANPPFSLKNWTDGLSIDPKSKQVINDSFNRFEDGT
PPEKNGDFAFLHLHIKSLKNTGKGAVILPHGVLFRGNAEGVIRKNLLTKGYIKGVIGLAP
NLFYGTSPACVIIIDKENAHARKGVFMIDASKDFKKDGNKNRLRDQDVQKMIDTFNAYK
EIPYYSKMVSLEEISANDYNLNIARYIAAKPESEKDLFALINSPSYLPKNEIKAYDPYFQ
VFKEKNTLFFKSDKEGYALKTECENIKELIIQSLEYQTFHASVLSAFDRLELSTTFND
LEPGFNPKTLIESVCSKVLYEFEKIEILDKYGVYQLFKDYINEVLQDDWFLLSFNGFESA
KNLRKLNPLKDKSKKANYLEEDFVIQKTHYKSDLPKNLIKQRFEEKESKELEELNTL
NEKEAHFEEFIEEHSNEEGLFYELKINESVLKKEKLNATDLEDKEILKTALELLEAKNKA
LKMKNKAYEELELKAHFHQYKNLKLDEIKDLIIQDKWLNSLKNALENKILKRINAFTSALN
TIISNYSNSLLELDKEVKESESKVLEHLKDLGLMG

>PROKKA_00334 Type-1 restriction enzyme R protein

MKTEKEVQKQVIETFKSMGYAYLGLDTSKNKNINKESLKAWLIKNOKINDERWQRIEQK
INNALKNDLYEANQKFYELLIYGVKTKISQNTTQTTYFIDWKDVSKNEFSVAEEVSVKG
PNTKRDPDIVLYVNGIALGVLELKSSVSVESGIRQNLNDQKKEFIRDFFKTIQLVMAGNE
SQGLKYGVIEETEEKYLSWKEEGVLKNLFETIERFLKKERFLEFIHDFLIFDKGKKKCAR
FHQYFAIKKTQEFIQRKEGGIIWHTQSGSKSLTMVWLTRWLRRNKERARVLIIVTDRRELD
AQIQGVFEGIGEDLYRADSKKDLLSVLFENKEFLVSSSLVHKFDDNDLEDLKKQPVLEKEWV
VLVDECHRTQSGKLHKAMKSLLPNAIFIAFSGTPLLKQDKKTSQEVFGDYIHCKYKFNEAV
SDKVVLDLNYEARSVDQYVSSPEKLDEYFELKTQGLNDIAKAELKKKWNVLQKVFSTKDR
LEHIAQDIVLDMAKLPRLRSEKGNAMLVAESVYNACRYFELFLETELKDKVAVITSYEPN
IADLKDCGSNESEESYKYRIYRKMLQNFFNEKDEKKALNQIKEFEKVKERFINEPNRMK
LLIVVYKLLTGFDAPSLTYLYIDKKMQDHKLFQAVCRVNRLDSEDKDFGCIIDYSDLFDS
LQEAHSDYTNGAFENYEREDIQGLISDKAQKIKKKLEEARDQLRSLGESVKEPKDEMYYI
AYFCGNDLEKNAQKRRLFYQLVGAFLRMFVELNNAEKPVSKEEMQQIKQEAQFYRHLQK
AVSLSSGSDVDLKSYSSEDMRRILDAYIKATDSKTLIKIEDQGLCEVLAQMDINDFNKELS
QAFKNESMAESIANNKKRIIEKEASDPKYEKLSSLLNDLINQFREKKLTYLEYLQOI
QHLLAQVIHKEDRNYPKKINTNALKTLYDNLNDENDNLDKKEALALEIDACIRGNKKDQGWV
GHNQKEKNLKIALLRKIINDEGLLENVFNLAKNIEEYR

>PROKKA_00335 hypothetical protein

MNAYCLTLNNTNIAIEKKDIKHLHISVCPDGSVHVSCPLALNDESLRLSLIKRPLWIK
QQQNFNLQNRQSQRGMLESHYLF GKRYLLKIEHTTKKHFLVQSPKYLILHVQKKTSLE
NRLKVLNYYRQVLREKIQTICINQYEKILNESIQSFRIQKMKRIWGSNCIAKRTLLFNLE
LVKAPRKGIEYVVVHEILHLKVRHHNEYFRDLSLYLPNWQRAKASLKETYLEYS

>PROKKA_00336 Fe(3+) dicitrate transport protein FecA precursor

MHKKVLLALTASLICQESLFAKEKDYTLGKVSTAGKKDKSDYSGQVNLGYSGITAPKSWQ
DEEVKKYTGSRVTISNKALTQQANQSIEEALQNVPLQIRNATGVGAMPTIQIRGFGAGG
SGHSDATLMLVNGIPVYMAPYAHIELDIFPVTFQAIDRIDVIKGGGSVQYGPNTYGGIVN
IITKPIPTQWENQAAERITYWAKARNAGFAAPPDKTGDPSFIKSLGNNLLYNTYVRSGGM
INKHVGIIQAQANWVRGQGFRDNSPSNISNYWLDGIYDINESNGIKAYYQYYDFAIAQPGS
LSEQDYKINRFANLRPLNQKGGRSQRFQAVYENRFGDLDRVGGTFSFTYYGQLMTRDFQV
SSSYNSANMVTCTFSEAACRAAGLPAGYNLAVPYATNYNGWAEVENPVRSINNAFEPKVN
LIVNTGKVKQTFIMGLRFMTTTFLOQRYLNTNECATKTSSEGAGFLCEGANVMMSGWKPHI
KHGVYRNWNNWRNNYTAVYLSDRIEAWDGRFFIVPGLRYAFVQYNNENASNWMQIPEKDL
RKIKHMNNWMPSTNIGFIPVQGDHNVLTIFYNYQRSFVPPQLDVLSSYGGAEYFTQHFDTVE
AGARYTYKDKFSFNADYFRIWARDFATGQYSVYTS GPMKGNVRPINGYSQGVLELEYR
IRGLQFHAAFNIDTRVTSHGPLTDLNGDVLKGTSYNKHFPFVSPFQFIFDARYNWHKTT
IGISSYFYSRAYSGISNSAAGGYGMQYSSGNNYESVLNSGYQCEAWCMTQHEGLLPWY
WVWNIQVSQIFWENGRHRVTGSLQINNIIFNMKYFTGIGSSPAGLQPGPGRSVTAYLNYT
F

>PROKKA_00337 Arginase

MILVGLEAELGASKRGTDKGVRLREVLSATHGDVIKGMQTITQERCVLKYKEFRYAKNFE

DYYLFCKENLIPCMKEVFEKKEFPLILSSEHANMFGIFQAFRSVHKDKKIGILYLDAAHAD
IHTAYDSDSKHIHGMPLGMVLNVRVSGFNRMSESEEKAWQKLCSLGLEKGGLEIDPKCLV
YFGVRSTEQSERDVIKELEIPLFSVGAIRENMQGVVQTKELLKAVDIIYLSLDLMDG
KLFTSTGVRENGLSFDELRLQLLGLLLESFKDRLGALEVTEYNPTVSTKHNNEEKQVLE
ILDLIINSCKIKDKKPSFIMSY

>PROKKA_00338 Alanine dehydrogenase

MTIGLVKESMDLESRVALVPDDVALIVQKGEVLVENSAGANSNGYSNEAYESVGAKIVDT
QTAWGQDLVVCKEPLHEEYPLLKEKATLFSYLDLAYQKSLCEMFIDKKITSICTETIAG
SKNDYPILAPMSVVAGRLAAHLVQHLLALEHVKGFMGKGVMLGGLSGAQRAKIVVGGG
VVGMEASAKVLSQMGAKVTILELDYAKLQNHPIYHLYDLEVLVNEANI IQALNGAVGLVG
AVLVTASQTPKVILRRHLKYMQKQGVVIDVACDLGGCIETIRQTSNSNPVYVEEGLLHYG
VPNMPGIVAKTSSTAYSHASVPYLLYLEHGLKGLTANTKIVANTLGGLSAYNGYITQE
GIAKTFNLAFKSSSEVLKEL

>PROKKA_00339 hypothetical protein

MAKIGSEDSYLEDILKDKLYYQIPIYQRPYQWTEENCEKLLDDLFFNYENDRESDYFCGS
LVLIAIGTDSATNAKTYDVVDGQQLSTFILLAKVLATLYSDLPENQEFLLQASWSDRHE
NKKKKKKRRLDFELVGSNAKKYFQDALKFFDDL DASKGEDSKSNDFPSKGKNNYLKNAICL
KNYLEKKRLETLTISLSGCILMSYLSPSLALMQTKR

>PROKKA_00340 hypothetical protein

MNATDIFKGELLKKL TEEKEREELATRWEDLRQKCLDNGFTMEILFSQYLVYLESKTSKE
KMEKMLVTWFKKLNKTPLEYLDGVEDFYNAYCEVLEMQDRHAYLLSYKNDDHLCVMLCAS
LLHRYSESEIETLKELLVKFYQDWVAGQTKTTRSQTCCNIINALKEKKSVENIASIVKK
YFKDKNITQRFKENLQDSNLYAKFYFDGKSAKKKFMAQTHSHFSGIFHER

>PROKKA_00341 hypothetical protein

MSDDPKPKRIEKNDFHVEHILPQNPDPSQWVKDFSEERERYTHSLANLTLLGGKKNTK
ALSQVLNQDFKEKKEIYMGKPVALDNKKTFKVMTCYDMTKNDVCRYTEWTPKSLEKRKEE
LIKVIESVLEL

>PROKKA_00342 Helicobacter outer membrane protein

MKKFFSKPLLALIVSVNALLAMDGNVFLGAGYLQGGQAMHADINSQKQATNATIKGFDA
LLGYQFFFEKHFGRLRYGFFDYAHANSIRLKNPNYNSEVAQLAGQILGKQEI NRLTSLAD
PKTFEPNMLTYGGAMDLMVNVINNGIMSLGAFGGVQLAGNSWLMATPSFEGILVEQALVS
KKATSFQFLFNVGARLRILKHSSIEAGVKFPMLKKNPYITAKNLDIGFRRVYSWYVNYVF
TF

>PROKKA_00343 putative inorganic polyphosphate/ATP-NAD kinase

MKDSHQITIGVFVRPTHYQNPLFEELEQAKEWVLKLEDEGFESFMIDSLDGAKDERLIEK
AYAFCLCGDGTILGALRMTHSYNKPFCFVGRIGNLGLFLSAVELNDLKDFLQDLKHNRIKL
EEHLALEGRIGKTSFYAINEIVIAKKKALGVLDIQAYAGHTPFNTYKGDGLIIATPLGST
AYNLSAHGPIVHALSQSYILTPLCDFSLTQRPLVLGAEFCLNFCAHEDALVIDGQATYD
LKANQPLYIQKSPTTTTKLLQKNSRDYFKVLKEKLLWGESPSSKKR

>PROKKA_00344 DNA repair protein RecN

MRDFNNAQITRLKVRQNAVFEKLDLEFQDGLSAISGASGVGKSVLIASLLGAFGLKESNA
SNIEVELIAPFLDTEEYGFREDNHEPLVISVIKKEKTRYFLNQTSLSKNTL KALLKGLI
KRLSNDRFQNELNDILMLSLLDGYIQNKQAFSPLLGTLEEKFTHLEKLEKERRLLEDK
KRFQKDLEERLNFEEKMLERLDLKEDEYERLLEQKLLSSKEKLN DKIALALEVLENTHK
ITHALESVGHSAEFLKSALLEASALLEKEQAKLEECERLDIEKVLERLGM LSGI IKDYGS
IMHAKERLGHVKNELHNLKEIDHHCETYHKEIEQLKAECKLCEEISGFRKEYLAGFNAL
LSAKAKDLLKSPSLVLEDAPMSEKGAQKLVHLQNSQLETLS SGEYSRLRLAFMLLEME
FLKDFKGVLVLDEMDSNLSGEESLAVSKALETLSHSQIFAISHQVHIPALAKNHILVFK
ENHKS LAKTLSNEERVLEIARMIGGSENIESAISFAKEKLKAQE

>PROKKA_00345 hypothetical protein

MKFFLLKKFSEFLNAQTHFSLKRLSASSFLLETFSKEKHAFVVDLNPYIGLSKKPPESV
LKNTLALDFCLNKFTKNAKILQANI IDNDRI LEIKGAKDLAYKSENFILRLEMIPKKANL
MILDKEKCVIEAFRFNDRVAKNDILGALPPNIYEHQEEDLDFKGLLDILEKDFLSYQHKE
LEHKKNQI IKRLNTQKERLKEKLEKLEDPKNLQLEAKELQTQASLLLT YQH LIHKHESRV
VLKDFEDKECAIEIDRSMPLNAFINKKFTLSKKKKQKSQFLYLEEENLKEKIAFKENQIN
YVKGAEESVLEMFIPVKNSKIKRPMSGYEVLYYKDFKIGLGKNQKENIKLLQDARANDL
WMHVRNIPGSHLIVFCQKNTPKDEVIMELAKMLIKMQKDAFNSEIDYTQRKFVKI I KGA
NVIYSKYRTISLKDT

>PROKKA_00346 hypothetical protein
MGVSALGNVTYINQNAPLNSAIQQGARSMDMLDLAQSFQSKLELAQETRALENMQSISDKD
PKEGSKNRYPN SQRAKAFGLEEEIEELWHEEHLLDIVG

>PROKKA_00347 Putative beta-lactamase HcpC precursor
MLRGVKKAVFRVLCIGALCLGGLMAEPNPKELIFSGITIYTDKDFTRAKEYFEKACRLND
ADGCAILREAYSNGGVILKGSARESIEKALEHTATAKACKSND AEKCKDLAEFYFNANDL
KNALEYYSKSKLNNVEGCMLSATFYNDMIKGLKKDKKDLEYYSKACELNYGDGCAILGD
IYHN GEGVTQNFKKAFKYYSKACELNNGEGCSKLGGDYFLGESVTQDLKKAFGYYSKACE
LNEALTCTLVGEFYRDGEGVTKDLKKAFEYSAKACELNDAKGCYALAAFYNEGKGVAKDE
KQT TENLEKSCKLGLKEACDILKEQKQ

>PROKKA_00348 DNA polymerase III PolC-type
MLDSLHLHKDIQALIA RLKHQDLSLGMLEKSLSR LIYDEINLEYLKACGLNFVETSEN LIT
LKNLKTPLKDEVFSFIDLETTGSCPIKHEILEIGAVQVKGEI INRFETLVKVKSVPDYI
AELTGITYEDTLNAPSVHEALQELRLFLGNSV FVAHNANFDYNFLGRYFVEKLHCPLLNL
KLCTLDLSKRAILSMRYSLSFLKELLGFGIEVSHRAYADALASYKLFEICLLNLPSYIKT
TMDLIDFSKCANTLIKRP PKARYQEISSPFSLFERTKGLFNIIKATS

>PROKKA_00349 Ribulose-phosphate 3-epimerase
MKVAPSLLSADFMHLAKEIESVS NADFLHVDVMDGHYVPNLTMGPVVLEKVTQMSQVPLD
VHLMVENASFFVGLFAPLKPQII SIHAENEKHPHRVLQLIRNSGITPGIVLNPHTHEESI
KYLLESVGLVLLMSVNP GFGGQKFLDLVLEKCLKV KELIRHYNPSC LLEV DGGVNDKNIF
ELQQAGVDVVVSGSYIFKSKDRKLAIEGLQNA RQPLA

>PROKKA_00350 Fructose-1,6-bisphosphatase class 1
MDYKHFCKKHANIVIEIISLLEKGVKKAQEILEKPDAGSYTKLENSSGDTP IKADLALDK
FLEENFLSLENIKSVFSEEKETPITKENG SYLIAYDPLDGSSVMEANFLVGTIIIGVYEKD
YKAQNLVASLYVVF GHKIELVVALDKVYRYAFYQNK FHF IETIVLENKGKIIASGGNQKD
FSLGLKKALEGFFAENYRLRYS GSMVADVHHVLIKKGGMFSYPQKKLRKLFEVFP LALMV
EKAKGEAFYFDKG VKRLLDQSVENYHEKSECYLASSHEAHILEKYLKGE

>PROKKA_00351 hypothetical protein
MPNSAKKLEYEERFNDALLKLKACQEEKQVASC LKCEKVLKCEIRNSYVNAAYESMSLGE
AGGFDFN

>PROKKA_00352 hypothetical protein
MVSNTTLQKNLDAFYTHPKIARFCLDLLKDLINQNLGLDLN AFHFLEPSAGSGSFVDVLK
ELGIADWEALDIAPKAQGIQKKDY LLELIEFNKKRIIIGNPPFGHRGKLALDFLNKSLNE
APIVAFILPNLFKRSIQKHIDKRAKLVLNAPLEKNAFIFNERPYDVKCVFQIYMHKNIA
LNLKDERIIAPPKIRHND FITYIHNNTPHTLKYFNKEKYQWDFAVVRQGFYDYNEKITNA
NLLIKNRQYFFIKAH SKEALMIIHKIDFNKLAHKNTQVLGFSTYDFVEEYCKLKEMHA

>PROKKA_00353 hypothetical protein
MLEKVFQEITNKRKFFASSSTGEQFENQFRNELKKLFSEINGDLTEKLSHIEEKPNKEIK
TAFNQLKKQVLEKNHPDTLXPF SNLTSHFLYQPFQS QNYHDFLVFIFDHVVGIEIKFSK
NDKGEKNLQTSRPMWNSNL PKPNAIYVYG VANANITFLKAQIF

>PROKKA_00354 hypothetical protein
MSYETREVLLKYFDTL DKDEESLKNALKDLENPF GFAPYIRKAYEHKKEFSNHHQIESFF
SHNHILREQNVLEFLKTLTH

>PROKKA_00355 Elongation factor Ts
MSGISAQLVKKLRDLTDAGMMDCKKALVEVAGDLQKAIDFLREKGLSKAAKKADRIAAEG
VVALEVAPDFKSAMMVEINSETDFVAKNEGFKELVKKTLETIKAHNIHTPEELLKSPLDN
KPFEEYLHSQI AVIGENILVRKIAHLKAPSSHIINGYAHSNARVGVLIGIKYDNEKNAPK
VVELARNIAMHAAAMKPQVLDCKDFSLDFVKKETLALIAEIEKDNEEAKRLGKPLKNIPA
FGSRIELSDEVLAHQKKA FEDELKAQGKPEKIWDKIVPGKMERFIADNTLIDQRLTLLGQ
FYVMDDKKTIAQVVADCSKEWDDDLKITEYVRFELGEGIEKKTENFAEEVALQMK

>PROKKA_00356 30S ribosomal protein S2
MVTMKDLLECGVHF GHQTRRWNP KTKKFIFGVRKNIHIIDLQKTLRYFRYTYNIVRDASA
QGKSIMFVGTKKQANETLKEFAESI QVPYVNYRWLGGM LTNFSTIRKSVRKLEIIIEEMEN
SGQIDLLTKKEKLMILRKKEKLDKYLGGVRHMKKIPDMIFVIDVAKEKIAVAEARKLHIP
IVAPLDTNCDPDLVDYPIPGNDDAIRSIRLFC KEMSEAILEGRELMQE EIVHADENSEEI
EFVSNEEKEEMLAEIQKEITQGAE

>PROKKA_00357 ATP-dependent helicase/nuclease subunit A
MDTKRQCMALKASAGSGKTFALSVRFLALLFKGANPSEILTLTFTKKATAEMKERILDYL

KILQKENLENEKEKSQNILKELEEKYHLDPSLVRNSAPKIIYQRFLEVRISTIDAFFQS
ILRKFCWFVGLSANFEVNEDTKVHQRLNEGFLSALNNEQLEELSAFIVQCLSYDSYTS
SILERLRLFLKNKLYLFDPNKKEPVFDEESFLEKLRSLNNQIQSIETASDRAKEAIKCDSF
RGFLNSSLTWLEKKSEYIYFKKLKDEIPTLESECEEIENDLKRYEARETALFKKFPKFI
QLYDKATSKIQALDFDAIKDKVHVLNNGYEEMPAEFFYFRLDSKIAHILIDEFQDTS
LNDYKILAPFIDEIKAGIGQAKWHRSVFFVGDVKQSIYAFRGSFSSLFESVSKDFYHDNLQFN
HRSAPLIINYVNTIFKKAYQNSPTAYLEQKYPKTSQNKHVTEGYVKVSLVADERELLLDQ
VLQEAQNLLDHHIDPKDITILCATNKDALEIKNYLQENLSAIRPSTESSAKLSQFVESKI
IKNALRYALAEPPYKPFYKHSVLKLAGYLHDDVIALAGFNPKKESVAGFVWKIMELFELY
GEPAQICLELAVGCEDADGFLEKLEAKSIASFNAKGAQIMTIHKSCKGMQFPYVIVCERLG
KPNSSHSNQLLEEDGAEALLRLYRMKNREVVDKDYARALDKEEAADHEETNVYYVAFT
RAELGLIVVAKDKKESKESKNTMREQLDLVPLEEGEIMPVISPOKEPLITSALIKPHA
YGEQVQEIEEEPSDYKNNQDAINFQIALHKGLEYQYAYNIPKKSLEYLNYHHGFYGL
DYQALEESLELFENDAEIQALFKNHALKGEAAFLFEGVVSRIQVLLWWDKGQNLQVLDYKS
SQNYQQSHKAQVSHYAEFLKTQAPHFKIQAGIIYAHKRLLLEKLWV

>PROKKA_00358 Na(+)/H(+) antiporter NhaA

MNLKKTENALSLTLKNFIKSESFGGIFLFLNAVLAMVVANSFLKESYFALWHTPFQIG
DFFIGFSLHNWIDDVLMALFFLMIGLEIKRELLFGELSSFKKASFPVIAALGGMIAPLI
YFFLNANTPSQHGFIPMATDIAFALGVIMLLGKRVPTALKVFLITLAVADDLGAIVVIA
LFYTTNLKFAWLLGALGVVILAVLNRNLNIRSLIPYLLLGVLWFCVHQSGIHATIAAVI
LAFMIPVKIPKDSKNVELLELGKRYAETSSGALLTKEQQEILHSIEEKASALQSPLERLE
HFLAPISGYFIMPLFAFANAGVSVDSINLEVDKVLLGVILGLCLGKPLGIFLITFISEK
LKITARPKGISWWHILGAGLLAGIGFTMSMFISNLAFTSEHKDAMEVAKIAILLGSLISG
IIGALYLFALDKKAALKK

>PROKKA_00359 preprotein translocase subunit YajC

MGQIKDILTTLPLVVLFLIFYFLIVRPQRQQQKKHKEMIEGLTKGDKIVTQGGGLIVEVL
KAEANFFSVKLNDDTTAKLSKNYVAFKLDELQDQGLAEPIVIQQGREEISAKLSGAKTLK
QRQITTK

>PROKKA_00360 preprotein translocase subunit SecD

MKLFNSRLIVFICALLLGVGFSVPSLLETGKPKITLGLDLRGGLNMLLGVTDEALKNKY
LSLASALEYNAKKQNILLKDIKSSLEGISFELDEDEAKKLDALLLELQGHSSQFEIKKEA
EFYSVKLTPLEQEELRKNTILQVIGIIRNRLDQFGLAEPVVIQQGKEEISVQLPGIKTLE
EERRAKDLISRSAHLQMMMAVDEEHNKDAMKMTDLEAQKLGSVLLSDVEMGGKILLKAIP
LDGEMLTDAKVVDQNNQPVVSFTLDAQGAKIFGDFSGANVGKRMIVLDNKVYSAPVIR
ERIGGGSGQISGNFSVAQASDLAIALRSGAMSAPIQVLEKRIIGPSLGKDSVKTSIIALI
GGFILVMGMFLVYYSMAGVIACLALVNLFLIVAVMAIFGATLTLPGMAGIVLTVGIAVD
ANIIINERIREVLRENEGIAKAIHLGYINASRAIFDSNITSLIASVLLYAYGTGAIKGFA
LTTGIGILASIITAIVGTQGIYQALLPKLTQTKSLYFWFGVNKRA

>PROKKA_00361 preprotein translocase subunit SecE

MELFKQTRILSFMRYSNYGVIVSAILALLALGLLFFKGFSLGIDFAGGSLVQARYTQNA
PKEVRDLFEKEARFKGVQVSEFGSKEEILIKFPFVETAENEDLNAIVANILKPSGDFEIR
KFDTVGPRVGSSELKEKGILSLILALMAIMVYVSFRYEWRFALASVIALVHDVILVASSVI
VFKIDMNLEVIAALLTLIGYSINDTIIIFDRIREEMLSQKTKNAIQIDEAISSTLTRL
LTSLTVFFVVLILCVFGSKIIIGFSLPMLIGTIVGTYSIFIAPKVALLLGFDMGKYEN
EARKIKKAQEKEKMRRLYEGGQV

>PROKKA_00362 hypothetical protein

MDWGRVVHVLFSLSLTLTIAGFLYEPNTVVLFVALALNLISVTLKIGVIKRFASELLASS
LATVHLHLPFAFVFLQILNNLVTAIYMLMIGALISNAFSLIFLLIESVVTSETD

>PROKKA_00363 Leucine--tRNA ligase

MDFINIEKKWQEFWWKNKSFEPEKDDFNLPKKYILSMLPYPSPGEIHMGHVRNYTIGDALAR
YYRLHHYNVLHPMGFDSFGMPAENAAIKHGIHPKTWTYENIEAMQKEFEALGFSFSKNRE
FATSDPDYTKFEQQFFIDLWEKGLIYRKKAMLNWCPNDKTVLANEQVIDGRCWRCDTEVI
QKELYQYYLKITNYAEELLKDLETLENHWPSQVLIMQKNWIGKSSGLQFGFKIADECLKA
CNNIQEIEVFTRADTIYGVTYIAIAPEHPLVEHAIKLVSQEDSKMIKAILNTTQRERAL
EKKGVFLGIYAIHPLTKQKIPVWVANFALANYGSGALMGVPACDERDFEFANLYHIPIKV
ITQGPQNLPHTKEEVLKNSGEWSDLSSSVAREKIIAYFEKENLGKRVINYRLQDWGVSRQ
RYWGAPIPMIHCKHCGIVPETQLPVTLPEDIVIDGEGNPLEKHASWKFAQCPKCHKDALR
ETDTMDTFTIQSSWYFLRYTTPKNQRENQAFDQNYLKYFMPVDTYIGGIEHAILHLLYARF

FTKALRDLGYLHLDEPFKQLITQGMVLKDGAKMSKSGNVVSPKEILKKYGADAARLFIL
FAAPPAKELEWNSALEGAHRFIKRLYDKANAITPTTSKPEFKEVSLNEAQKLARKKVYE
ALKKSHEIFNKAESAYTFNTLIASCMEALNALSAQNNERILCEGYFVLLQILEPIIPHTA
WELSERLFKRENFKPIAIDENALMEDFMTLGLTINGKRRAELKVNINASKEEVIVLAKKE
LEKYLEKASVKKEIYVPNKLNVFVIA

>PROKKA_00364 hypothetical protein

MRALLFFILLLLLFKGCYKPIAAYAQNALGDSIYVKLIVNLPNPENSVEFKDLMNRLVVQ
RFQSRLASEKDADSIIEITNVTDTSTITQNKEGFTTFYRATVSVNYTYDNKRGVRKTFQ
NSGYINYAVNLQDPLNTYQNRYYAINQAVEQTLTKFVAQIAYEGKFNNEK

>PROKKA_00365 Bifunctional protein FolC

MKNSPIDXLSSAFLETKPKEYHKFDPSPRFIQIYKDFKNAFFEIQAKVIHIVGTNGKGSTG
RFTLLLLADQNFKVLHFTSPHVFEFRERFFLNGSIVKESVLENAHQQQLQSHAFSSACSYF
EYATLLAVMLAKDCDYLVLLEAGLGGEFDSSTNALEKTLVFTPIDYDHKEFLGDSLESIAT
TKLKAMGSLSIIPAPQQLVLNIAQKIAKDHAQLIVVQNEISKGVRDYIERHHLAHFLAM
NLEVALKAFAEALLPYNKQEVLEKLNKPLNLIGRCELLSPNVLIDVGHNPMSAKALKEEIKR
VFNAPIVLIYNCYQDKDAFLVLEILKPVIKKVLILELHNERIIQLEKLGILETLGLEHA
LFEDVKENENYLVYGSFLVANAFYKRYQQKRD

>PROKKA_00366 Murein DD-endopeptidase MepM

MPQNQLVITIIDESGSKQLKFSKNLKRNLIIISVILLIVGLGVGFLKFLIAKMDTMTSE
RNAVLRDFRDLYQKNYTLTKEIKNKREELFIVGQKIRGLESLEIEVKRGANGGVHLYDEVD
LENLNLAAQKHLALMLIPNGMPLKTYSAIKPTKERNHPIKKIKGVESGIDFIAPLNTPVYA
SADGIVDFVKTNSNVGYGSLVRIEHAFFGSSIIYTHLEHVNVQPKSFIQKGQLIGYSGKSG
NSGGEKLHYEVRFGLKILDAQKFLAWDLDFHQSALLENKFIEWKNLFWVLEDIVQLQEHV
DKDALISQ

>PROKKA_00367 Murein DD-endopeptidase MepM

MFLDRRLIVMVTDSKGSRYINVHILFRQIGLYALLSVVGSLLFLGISLLVLNKEIRNIEK
QHALITKEFEKKKETNEKLSLQMDDEFLLDQLSGERINDLEEVDVGNRPPEEEKEEGNFSS
RLDVAGITGLQKSFIMRLIPNDYPLESYRRVSAAFNKRIHPILHVLHNHTGLDLSTAIN
PVYASASGVVGLASKGWNGGYGNLIKVFHPFGFKTYAHLNKKIAVKTGEFVKKGQLIGYS
GNTGMSTGPHLHYEVRFGLDQPINPMSFTKWNMKDFEEVFNKERSIRWQSLITIIINRLMQK
QDQRLLSLKAQK

>PROKKA_00368 Polymer-forming cytoskeletal

MDYHLHVDGELEGVVHVKSTVVIGQTGSVVGEIFANKLVVNGKFTGTVEAEVVEIMPLGC
LDGKISSQELVVERKGILIGETRPNLQGGALLINEQEKKIENK

>PROKKA_00369 Transcription-repair-coupling factor

MIQSSLYRALNKGFDYQILACKDFKESKLAKKEVISYFKPNIAVLFPEFRAKNDDLRSF
FEEFLQLLGGREFYQALENKQETIIIPISALLHPLPKKELLESFKITLLEKYNLKDLEK
DKLFYYGYEILDLVEVEGEASFRGDIVDIYAPNSKAYRLSFFDTECESIKELDPTTQMSL
KEDLLEIEIPPTLFSLESSYKDLTKVEQSPLNSFSKDLTSFGLWFLGEKANDLLHAYK
SVISPRALEEIQELASLNELDCERFKLLKVLNAQGYEDLEIHAHALEGFIALHSNHKIT
LLAPNKTILDNLVGTIEKSNMECVIAPFVLNFKTPDGIFISLNSFERKKKRQKSKLALNE
LNPGEWVVHDDYGVGVFSQLVQHSVLGSKRDFLEIAYLGEDKLLLPVENLHLIARYVAQS
DSVPAKDRLGKGSFLKAKVVRTKLEIAGKIIELAAERNLILGKKMDVHLAELEVFKSQ
AGFEYTSQEKAI AEISKDLSSHRMMDRLLSGDVFGKTEVAMHAIFCAFLNGFQSALVV
PTTLLAHQHFFETLRARFENFGVKVARLDYASEKNKLLKAVELGLIDALVGTHAILGAKF
KNLGLVVVDEEHKFGVKQKEALKELSKSVHFLSMSATPIPRTLNMALSQIKGISSLKTPP
TDRKPSRTFLKEKNDELLKEIIYRELRRNGQIFYIHNHIASISKVTKLENLIPKLKIAI
LHSQINANESEEIMLEFAKGNYQVLLCTSIVESGIHLPNANTIIIDNAQNFGGLADLHQLR
GRVGRGKKEGFCYFLIEDQKSLNEQALKRLLALEKNSYLGSGESVAYHDLEIRGGGNLLG
QDQSGHIKNIGYALYTRMLEDAIYELSGGKKRLEKSVEIQLGVSFAFLNPELIASDSLRLD
LYRRLSLCENTDEVGQIHEEIEDRFGMDDLSAQFLQIITLKILANQLGIKLSNFNQNI
TLTYSDEHKESLKAPSKDDNDILETLKHLRAQISLKRR

>PROKKA_00370 Ubiquinol-cytochrome c reductase iron-sulfur subunit

MADIQRRDFLGMSLASVTAIGAIASLVAMKKTWDPLPSVVSAGFTTIDVANMQEGQFSTV
EWRGKPVYILKRSKKEGFNEKRDFKIGESVFTTAIQICTHLGCIPTYQDEEKGFLLCPCHG
GRFTADGVNIAGTPPPRPFIDIPFKIEGKTIKFGEAGA EYKMMMAKA

>PROKKA_00371 Cytochrome b

MAEIKKAKNLGEWLDMLRGTNKLKVLMT EYWI PKNINFLWAMGVILLTLFGVLVISGIF

LLMYYKPDAKMAFDSVNFTIMQEVAYGWLWRHMHATAASMI FV I I Y I H M F V G I Y Y G S Y K K
 GREMIWISGMILFVVFSAEAFSGYMLPWGQMSYWAAAVITNLFGGIPFIGADVVEWIRGN
 YVVADSTLTRFFMLHVFLLP I A I I L L V G V H F Y S L R I P H V N N Q E G E E I D F E L E E K K F I E G K
 K K E S K V I P F W P V F L S K D I F V V C A F M V F F F Y L V C Y H Y D F A M D P I N F E R A N S L K T P P H I Y P E
 W Y F L W S Y E V L R G F F F S A D L G L M A F G V A Q V I F F L L P F L D R S P V V A P A H K R P A F M V W F W L L I
 I D M I V L T I Y G K L P P L G I G K Y I G L A G S I T F L A L F F V V L P I I T I A E S K K Q G G V R
 >PROKKA_00372 Cytochrome c
 MKEFKIL I I L I V V V G V I Y Y G V E P Y A H S V M H P K V A P A D F A F K D L E P M D L K N G D A N K G K Q L V
 A E N C T A C H G I K S Q N I P A P M D S L S A S N S F G V V P P D L S H V A G V L N A N F L A H F I K D P V K T A K L
 S H K F N D E R S Y P M P A F S Q F S D Q D L S D I V A Y L T S I L P K N L S D K E V F A Q S C Q R C H S L D Y A K D K
 A F S D P K D L A N Y L G S H A P D L S M M I R A K G E H G L S I F I N D P Q K L L P G T A M P R V G L S E K A Q K Q V
 I S Y L E K A G D R K K H E R N T L G I K I M I F F A V L S F L A Y A W K R K V W S E V Y
 >PROKKA_00373 Phosphoethanolamine transferase CptA
 MILD SK F F V F I L F N V L S A C F L S S V S A G F I F K A F L Y S F I W L F I I Y Y A I S F I K N R F V T E S L K
 S F I L V L G I V F S L I D I F G S Y Y F H L P L S N D L G N I L F T T H Y K E S L E F L H A Y V Y P H R Y F V I G L I
 L I A I G S L K L F S F V P N K P I P L K M A S I L S V L F L I V E A P H T I A T I K K Y K E R E D L L N A D G T M E Y
 I A L A K G A Y Y F G R N I S S L R E S N N S N Q A L E K A S Y S K D Y L L K N T G S V E N V V L V F G E S L N R N F M
 G V Y G Y Q A P T T P Y L N A L K E K G S L L V F D N V I S P A F Y T D K S F T M L L T Y A N R D N L N Q A W Y Q Y K
 N L A H I L K L S D Y K S V W I T S Q G Y G L M W G N S Y Y Q V A K H F D T Y I E N D K P Y D E N L A A L F K R Y Y D N
 E R E R E R V K S V K I L L F F T
 >PROKKA_00374 Phosphoethanolamine transferase CptA
 M I G N H F E Y K S R F P K E F S R F N L N N T S Y F S K N K S L K V K N N A D K Q V V T D Y I N S V Y Y N D Y V L H S
 L I E L F K D K D S L V I Y L S D H G D D M F E S S A F N A H E C S N A S M E I P F L I Y M S D A F K Q K H P Q M V K S
 F E E A L H K P F M S D D L L H T L L P L A G I I T K D Y E K T R D L L N E D Y N D K R P R K P C D N K V Y P M S K
 >PROKKA_00375 Penicillin-binding protein 2
 M D N K N I D P S F N P E R F L E T Q K Y K G S I T A L I F L L L F I V F L M V A F K K A F F A Q V N M P N L V M S K Q
 D T A T R G A I Y S Q D N Y S L A T S Q T L F K L G F D T R F L N P D K E D F F I D F L S I Y S N I P K K S L K D A I N
 T K G Y I I L A Y N L T P N V A A N I R D L N K K F L T F G V F Q N F K D A H D K V W Q K Q G L N I E V S G V S R H Y P
 Y Q N S L E P I I G Y V Q K Q E E D R L T L T T G K K G V E K S Q D H L L K A Q Q N G I R I G K R D V S F N F I Q N H S
 Y T E V E R L D G Y E V Y L S V P L K L Q R E I E T L L D K T K D K L K A K E I L V G I I N P K S G E I L S L A S S K R
 F N P N A I K T S D Y E S L N L S V A E K V F E P G S T I K P I V Y S L L L D K N L I N P K E R I D L N H G Y Y Q L G K
 Y T I K D D F I P S K K A V V E D I L I Q S S N V G M I K I S K N L N P E D F Y N G L L G Y G F S Q K T G I D L S L E A
 T G K I P P L S A F K R E V L K G S V S Y G Y G L N A T F L Q L L R A Y A V F S N E G K L T T P Y L V Q R E T A P N G D
 I Y I P S P K P T F Q V I S P K S A R K M K E T L I K V V R Y G T G K N A Q F E G L Y V G G K T G T A R V A K N G S Y S
 V Q S Y N S S F F G F A E D E R Q V F T I G V V I L G S H G K E E Y Y A S K I A A P I F K E I T E I L V R Y N Y L S P S
 I A I Q N A L E K N R F K I K
 >PROKKA_00376 flagellar hook-basal body protein FliE
 M Q A I H N D K S L L S P F S E L N T D N R T K R E E S G S T F K E Q K G G E F S K L L K Q S I N E L N N T Q E Q S D K
 A L A D M A T G Q I K D L H Q A A I A I G K A E T S M K L M L E V R N K A I S A Y K E L L R T Q I
 >PROKKA_00377 Flagellar basal-body rod protein FlgC
 M F L S S F D I S G Y G L S A Q R L R A N L I S S N I A N A N T T R T S E G G P Y R R Q E A V F K A F D F N E I L N Q K
 I A Q N N Q I T P Y E D P L D E G D E Y P L I P I T S V V V D K I V R D D S E P L M K Y D P S H P D A N A Q G Y V A Y P
 N V N A V V E M A D L V E A T R A Y Q A N V A A F Q S A K N M A Q N A I G M L Q T
 >PROKKA_00378 Flagellar basal body rod protein FlgB
 M E C F L H A M N L Q G F K K G V L M D F S K A F G L V Y K A L D Y R S L R Q D M I A S N I A N V D T P F Y R P K D L D
 F E S V L A K K K A E I F E N Q S S K V L P L A H T N P R H L D F E N S A K D G A S L F F R D G H L A K N D G N S V D L
 D I E T S E M G K N S T M Y L A L S S A L K K Y R G V I N Y A I D S S K N L
 >PROKKA_00379 Lipid II flippase FtsW
 M T T D R N L F F C A S L L I F L G V L M S Y S L S T Y T T V V L Y H Y G E F H F F I R Q L V S A I I G I V I M W G L S
 R V D P S K W F S R L G F F L L F I P P L L I I G M F F L P E S L S S A G G A K R W I R L G F F S L A P L E F L K V G
 F T F F L A W S L S R T F V A K E K A N V K E E L I T F V P Y S V V F V A L A I G V G V L Q N D L G Q I V L L G A V L A
 V L L V F S G G S T H L F G L I V S G A F A I S V L A I V T S E H R I L R L K L W S N L Q N S L F T L L P D K L A N A
 L R I S D L P E S Y Q V F H A G N A M H N G G L L G Q G L G L G Q I K L G F L S E V H T D M V L A G I A E E W G F F G L
 C V C F I L F S V L I V L I F R I A N R L K E P K Y S L F C V G V V L L I S F S L V I N A F G V G G I L P V K G L A V P
 F L S Y G G S S L L A N C I A I G L V L S L A R Y T K G
 >PROKKA_00380 corrinoid ABC transporter substrate-binding protein
 M L V T R F K K A F I S Y S L G V L V A S L L L N V C N A S A Q E V K V K D Y F G E Q N I K L P V S K I A Y I G S Y V E
 V P A M L N V W N R V V G V S D Y A F K D D I V K A T L K G E D L K R V K H M S T D H T A A L N V E L L K K L G P D L V

VTFVGNPKAVEHAKKFGISFLSFQEKTI AEAMQAMQAQAAVLEIDASKKFAKMQETLDFI
 AERLKDVKKKKGVELFHKANKISGHQAISSDILEKGGIDNFGGLKYVKFGRADISVEKIVK
 ENPEIIFIWVVSPLSPEDVLNNPKFSTIKAIKNKQVYKLPTMDIGGPAPLISLFIALKA
 HPEAFKGV DINAMVKDYYKVIFDLNDAEIEPFLWH
 >PROKKA_00381 corrinoid ABC transporter substrate-binding protein
 MLIARFKKALISCSL GALIVSSLLGVANASNQEVQVKDYFGDQAIKLPVSKIIYLGSAE
 VPAMLHTWDRVVGISDYAFKSDIVKATLKDPKRIKSMSSDHVAALNVELLKKLSPDLVVT
 FVGNPKAVEHAKKFGISFLSFQEKTI AEVMEDIDAQAKALEVDASKKLAKMQETLDFITE
 RLKNVKKKKKGVELFHKANKISGHQALDSDILEKGGIDNFGGLKYVKFGRADISVEKIVKEN
 PEIIFIWVISPLTPEDVLNNPKFSTIKAIKNKQVYKLPTMDIGGPAPLISLYIALKAXP
 EAFKGV DINAMVKDYYKVVF DLNDAEVEPFLWH
 >PROKKA_00382 putative peroxiredoxin
 MLVTKLAPDFKAPAVLGNNEVDEHFELSKNLGKNGVILFFWPKDFTFVCPTEIIAFDKRV
 KDFQEKGFNVIGV SIDSEQVHFAWKNTPEVEKGGIGQVTFPMVADITKSISR DYDLFEEA
 IALRGAFLIDKNMKVRHAVINDLPLGRNADEMLRMVDALLHFEEHGEVCPAGWRKGDKGM
 KATHKGVAEY LKENSINL
 >PROKKA_00383 D-methionine-binding lipoprotein MetQ precursor
 MNIFKRVCVTAIVLGFFNLLDAKHHEKKEDHKITRELKVGANPVPHAQILQSVVDDLK
 KKGIKLVIVSFTDYVLPNLALNDSSLDANYFQHRPYLDRFNLD RKMHLVGLANI HVEPLR
 FYSQKITDIKNLKKGSVIAVPNDPANQGRALILLHKQGLIALKDPSNLYATEFDIVKNPY
 NIKIKPLEAALLPKVLGDVDGAIITGN YALQAKLTGALFSEDKDSPYANLIAAREDNAQD
 EAIKALIEALQSEKTRKFILD TYKGAIIPAF
 >PROKKA_00384 Stage V sporulation protein D
 MKNLRYKLLLLFVFIGFWGLLVNLFILSVKNQEYYEKL AERNMTKKEFLVPTRGDITDRN
 DEFLATNELVFGVFLPSRLKQKELLEKIEIIQKFFPNFSKEVLLNNYQKENS LYHNHLIK
 VVGFIPIYIAMQPLYAKLIQTQGI FVHPLDKRYPNNALASHVLGYVGVASLQDLKDDEEN
 QYSQIVGKTGIEKEYNKLLQGKVG YKIMRVNALNQELATLEVVPSTNNHLQLSLDKRLQ
 KEADKLFENKRGAILVMDAENGELLVAGSYPEYNLND FVGGISQDKWQKLQDDIYNPLL N
 RFANALYPGSGVVMGVL SFL ENLHITENTTIPTPPFIEVGKRKFRDWKKTGHGNSNLY
 KAIRESVDVYFYKFGLEISIEKLSKTLREVGFG EKTGV DLPNEFVGIVPDNLWKLKRFNQ
 DWRVGD TLITAIGQGSFLATPLQVLAYTG LIATGKLATPHFAINN KQPLKDPLNSFQKKK
 LQALRVGM YEVCNHKDG TAYHSTRGSKVTLACKTGTAQVVEIAQNI VNRMKEKDMEYFHR
 SHAWITAF LPEYKPKYAITILVEHGE GSKLGGLLVKMSNKLYELGYL
 >PROKKA_00385 hypothetical protein
 MLSLKQDSFFFLCLGIFGFYFYSLLRDLMPFLPPMIGFLFLFYAKKYDRFLPSLSVFGCL
 FWFESMHLKTLGLV LALLFLIYHQI IYKNSLKL FNDGFLFKTLHVFLVYYLYLSRFFSVSL
 SLKILGFLALFALIESALWGLYEKSSL
 >PROKKA_00386 putative GTP-binding protein EngB
 MIAIKDAHFLTSSSQLSQCPASLTSEM VILGRSNVGKSTFINTLLGKNLAKSSATPGKTR
 LANFFSTTWEDKENALTTTFNVIDLPGFGYAKVSKSLKKEWEGFLWELLSVRVSIKLFIH
 LIDARHLDLEIDKNAKENIQALLRPDQAYLT LFTKFDKLNKNEQHRLFLNAPKPFLINTT
 HFNALSSKYPTLEIVRQTLLKYLLTNPL
 >PROKKA_00387 OstA-like protein
 MRWWCFLVCCFGILSVMSAQKTENRGLKKERELLEITGNQFVANDKTKTAVIQGNVQIKK
 GKDR LFADKVS VFLNDRKPERYEATGNTHFNIFTEDNREISGSADKLIYNALNGEYKLL
 QNAV VREV GKS NVITGDEIILNKTKGYADVLGSAKRPAKFVFD MEDINEENRKAKLKKKG
 AKEKP
 >PROKKA_00388 hypothetical protein
 MKRSSFTSNVSNLNFVVL SFITIGLVFFF LRSQPTSVVSKENIPKIELENFKAFQINDKI
 LDLSIEGKKALQYDDYEIFFDSKIKRYDEDTIESVESPKARRQODLYFFPNGVTYKRSD
 SSFWSETGIYNHKEQNFKGGRFILTSKDSKIEGLDISYSHALAVIEAQSIQAHFLFLDEI
 KQSQKEKKKLPTFKGGF
 >PROKKA_00389 3-deoxy-D-manno-octulosonate 8-phosphate phosphatase KdsC
 MIKLLLLDV DGTLDGSLYFDENFHEIKAFNVKDGLGMVLWQKL GKKIAIITGRTSIMVK
 KRME SLGVQFVFMGVENKGVVVERLKKDLQLSAQEIACVGGDDYNDLGMFKACALS FAPFD
 AHPLLKNKAYKVLQNSGGKGAVREAI DYLLTLEGLQDEALKLYL
 >PROKKA_00390 RlpA-like protein precursor
 MGLALKVKVCF LGVIFLISACAVKKEGIKNLSYKHESLRAYENAKDYDPTTKKATYKRNF

ERHFKHSDSDQSDNTKQPLDNGMRDSSAIQIRATMRPYQVGGKWYYPTKVDLGEKFDGIAS
WYGNPNFHAKKTSNGEIIYNMYAHTAAHKTLPMTNVVKVINVDNNLSTIVRINDRGPFVSDR
IIDLSNAAARDIDMVKKGTASVRLIVLGGFVISTQYEQSFNASSSKILHKEFKVGESEK
SVSGGKFSLQMGAFRNQIGAQTALADKLQAENKNYSVKVAFKDDLYKVLVQGFQSEEEARD
FMKKYNQNAVLTRE

>PROKKA_00391 Membrane-bound lytic murein transglycosylase D precursor

MKYFNTKWLFFLMTHWFLLTSLSHAKMAFESNIDTKALEAFGVNAGFLSQMPNALKKMNE
EEWKRLVKRFDVNYQFIPIIKNMLIEASVPQEFLLAMAESKFSSRAYSRKKAVGIWQF
MPSTAKELGLKVNHYIDERRDPIKSTQAAITYLKRLYKQTGEWYLVAMAYNYGLRKVQNA
IKAAGTSDIKILLDEDDKKYLPKETREYIRSILSLALKFNSLDNLKDKEYLLNRGARVSLV
GVPFKRHTSLVQVAKNLNLSLETLSYNHQFRYNILPSKDPTYTYIIPYEKLALFKQRQL
KQNKSIQASPKSPFITHVVLPKETLSSIAKRYQVSISSIQLANNLKDSNIFIHQRLIPT
NKKLLATREF

>PROKKA_00392 putative deoxyribonuclease YcfH

MFIDTHCHLDHKDYENDLEEVLKESLEKGVTCVPIPGADMKDLNRAIEISEKFEGVFFAI
GAHPYDVESFDESLEKFEVGHQKCAIGECGLDYRLPELNERENYKSKQKEIFTKQIEF
SIQHNKPLIIHIREASFDSLNLKSYPKAFGVLHCFNADGMLELSDRFYYGIGGVSTFK
NAKRLVEILPKIPKNRLLLETDSPLYTPHPFRGTRNSPTYIPLIAQKIAEIIINIETEELA
SLSTHNAQTLEFSP

>PROKKA_00393 Riboflavin synthase

MFSGLINQIAKVKSFHNNILSIESDLNPKLGDSIAINGACLTAEISSKTHFSVELSQKTQ
NSVALENYKHLVHIEPALKADASLDGHFVQGHIDAIGVIEKIIHNANQVDFEISISKETL
LLCVEQGSIAIEGVSLTSLKVEEKGFWLTIIPYTLLENTLFKTYKLRVNIETDMLVRSV
ASILKKTGFEKNFSWNDADSLTLGY

>PROKKA_00394 Flagellar biosynthetic protein FlhB

MNKTIKAVALAYNMGQDHAPKVIASGVGEVAKRIIQKAKEYDIALFSNPMLVDSLLKVEL
DCTIPEELYESVVQVFLWLNSVENNAQMSK

>PROKKA_00395 Methionine import ATP-binding protein MetN

MVVELKNIEKIYENGFFHALKGVNLELKKGDILGVIGYSGAGKSTLIRLINCLERPSSGEV
LVNGVNLLKLKPKELQKARQKIGMIFQHFNLLSAKNVFENVAFALEIARWEKTKIKSRVH
ELLELVGLEDKMHFYPKQLSGGQKQRVAIARSLANCPDLLLCDEATSALDSKTTHSILT
LSGIQKKLDLSIVFITHEIEVVKELCNQMCVISSGEIVERGSVEEIIFANPKHAVTKELLG
IKNEHGDQKCQDVYRIVFLGEHLDEPIISNLIKRFKIDVSIISGNIEELTTKDIGYLVVR
FLGSVAETQRALEYLNALGLQVEKLD

>PROKKA_00396 D-methionine transport system permease protein MetI

MISQMLIQATLETLYMVVASFLAVVFGPLPLGVLLLVSKKGHLLNKPLLHKILDTSINMT
RSFPFIILIIILLPLSRFLIGTSIGSSASIIPLAISAIIPFVAKLFENSLMEVEHGKIETT
LSLGASHLEVIKMMLESPLSVNNITIXLISLIGYSAMAGALGAGGLGDLAIRIGYQSY
RGDVLFFYAVVVIIVLVQIIQSAGDYVVKRLRKHKY

>PROKKA_00397 hypothetical protein

MLKKLLFIALFLGLFLRAQGEHYEIIVELSKAFLKAQEVLTTHQAYKTCIETGHDRTQIR
LQSAFLENLSQTEQQFDDYFEKDFKSVEVLKTLKNLQSLKTSNKLACIAPKNAQNFEI
LEGAITQIIDLEKQMDKFINGHSK

>PROKKA_00398 lipid A 1-phosphatase

MKKLKSFLILLLLWVYPLRSEPINEGAYILEEIGDVLRLPLIFVGTVSLAMRDYRGLGEL
AVGTLVTQGVIIYGLKGAFSNAHKD GARVEFAKRPPCNSWRGMPSGHAGGVFSAAGFVYYR
YGWKPALPVIALAILTDASRVVAGQHTILQVTIGSLIAWGFAYLFTSRYKPKQWMLYPEI
SSDFKGSRYGVSFYSYQW

>PROKKA_00399 putative undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase

MLFVDNANKIQGFHHARTPRAGGLGIFLSFVLAYLFESFEAPFKGFFVFLGLLLVLVLSGF
LEDINLSLSPKIRLILQAVGVCCIISSTPLVVSDFSLPLSLPYFIAFLFAVFMVLVGISNA
INIIDGFNGLASGICAIVLLVIHYIDPSSLSCLLAYMVLGFMVLNFPGLKIFLGDGGAYF
LGLVCGISLLNLSLEQKISVFFGLNLMLYPVVEVLFSILRRKMKRQKATMPDNLHLHTLL
FQFLQQRSLNYPNPLCALILILCNLPFILISVFFRSNAHALIVIGLVFIACYLIGYAYLN
RQVYALEKRAF

>PROKKA_00400 Pyridoxine 5'-phosphate synthase

MRFGLNIDHIVTLREIRKTYEPEILEALFIAKNTHKVDLITIHLREDRRIQNEDVLRLL

EISPLPINIECSINAEITDFLCSLKNKPSKVTIVPENRNEVTTEGGIDCSLKGLGEVIRA
YHNKGIEVSLFVDPLKDALHFARENQVKQVEFHTGVYANLHNALYSNANNQIHAISTLKD
KSPKELKEELHNAFLQLRRMSKEAFFMGITACAGHGLNYTNVKELLKIPSLRELNIGHSV
VSKAVLVGLEKAILEMAQLIKR

>PROKKA_00401 4-hydroxythreonine-4-phosphate dehydrogenase
MAKKKIAIISCGDVQGVGLELILKSHKEVSALCEPLYLIDGELLERANQLLHNAYETKTLN
TLAIDAPLPLNLSSTIGKVSTQSGAYSFESFKKACELADDKEVDGICTLPINKLAWQQAQ
IPFVGHTDFLKQRYKDHQIIMMLGCSKLFVGLFSDHVPLGAVSQLIQVESLVRFLAFQK
STQAKIVQVCGFNPHAGEEGLFGEDEKILKAVQKSNQTLGFECFLGFLPADSAFAPNKR
KITPFYVSMSHDVGLAPLKALYFDESINVSLNAPILRASTDHGTAFDIAYQNKANNKSYL
NAIKYLA

>PROKKA_00402 tRNA N6-adenosine threonylcarbamoyltransferase
MILSIESSCDDSSLALTRIKDAQLIAHFKISQEKHHSSYGGVPELASRLHAENLPLLE
RIKISLKNKDFSKLKAIAITNQPLSVTLIEGLMMAKALSLSLNLPLILEDHLRGHVYSLF
INEKQTCMPLSALLVSGGHSLLILEARDYEDIKIVATSLDDSFGESFDKVS KM L DLGYPGG
PIVEKLALDYRHPNEPLMFPIPLKNSPNLAFSFSGLKNAVRLEVEKNAPNLNEAIKQKIA
YHFQSTAIEHLIQQT KRYFKIKRPKIFGIVGGASQNLALRKAFENLCAEFDCKLV L APLE
FCSDNAAMIGRSSLEAYQKKRFVPLENANISPR TLLKSFE

>PROKKA_00403 Flagellar basal-body rod protein FlgG
MLRSLYSATS GMLAQQT H IDTTSNNIANVNTTGFKKSRADFNDLFYQAMQYAGTNTSNTT
LSPDGM EVGLGVRPSAITKMFSQGS PKETENNL DVAITGKGFFQVQLPDGTTAYTRSGNF
KLDEQGNLVTSEGYLLIPQITLPEDTTQVNI GVDGTVSVTQGLQTTSNVIGQITLANFVN
PAGLHSMGDNLSITNASGDAIVGNPDSQGLGKLRQGFLELSNVRLVEEMTDLITAQRAY
EANSKSIQTADAMLQTVNSLKR

>PROKKA_00404 hypothetical protein
MSINLAGSAYRVTVPACVLIAALLKLKAEQEAEARLKAQEAEDET NKTWYLVIVFCILIGL
AVLAVFYIKGKHQSNNNP KSGVENLKNPSHKN

>PROKKA_00405 Ubiquinol-cytochrome C chaperone
MITPNTQEGSLKNCSVMGATLLLIFFRNEGVLYKEILCDVCDHLKVSYNEKSNISLIEQN
MLAKLLKDSLEQMSKKDLEKL RHELGMTNIDKVIGENKQVLIASVLT LFKAGGSHSYALV
VSADAMVKKL

>PROKKA_00406 hypothetical protein
MSNSCSNQNHNQKGD AKNQNNESCQKNSQNQANQCNANHEHKNDKK

>PROKKA_00407 hypothetical protein
MAYKYDRDLEFLKQLESSDLLDLFEVLVFGKDGEKRHNEKLTSSIEYKRHGDDYAKYAER
IAEELQYYGSNSFASF IKGEGVLYKEILCDVCDKLKVNYNKKTTETTLIEQNMLSILERS
LEEMDDEEVKEMCDELSIKNTDNLNRQALSAATLT LFKMGGFKSYQLAVIVANAVAKTIL
GRGLSLAGNQVLTRTSLFTLGPVGWIIITGVWTAIDIAGPAYRVTIPACIVVATLRLKTQQ
ASEDKKSLQIESV

>PROKKA_00408 hypothetical protein
MIYTLAVFFFPWVAFFLRDRIFSAVFSFILWIIFGYNIFAFVFSGGGVFGISLILESFFLW
VILTVWCCLVMHSDQKQEDMEQLIN VINGIVKANSKSLHLETQQTNEANDMKQSIGETS
KTNSNEADATRILICVFIFFLAIILLVAFFIKGGYQLSNNPSIDAKNP SHKNYKKN DNNR
SKSLNP

>PROKKA_00409 hypothetical protein
MATRTQARGAVVELLYAFESGNEEIKKIASSMLEEKIKNNQLAFALS LFNGLERINEI
DALIEPHLKDWD FKR LGSMEKAILRLGAYEIGFTPTQNPII INECIELGKLYAEPNTPKF
LNAILD SL SKKLAQKPLN

>PROKKA_00410 6,7-dimethyl-8-ribityllumazine synthase
MQIIEGKLQLQGNERVAILTSRFNHIITDRLKEGAMDCFKRHGGDEKL L DLVLVPGAYEL
PFILERLLGSEKYDGV CVLGAIRGGTPHFDYVSAEATKGIASAM L KYSMPVSFGVLT TD
NIEQAIERAGSKAGNKGF EAMSTLIELLSLCQTLKG

>PROKKA_00411 2-dehydro-3-deoxyphosphooctonate aldolase
MKTSKTKTPKSVLIAGPCVIESLENLRSIATKLQPLASNERLDFYFKASF D KANRTSLES
YRGPGL EKGLEMLQI I KEEFGYKILTDVHESYQVSVAKVADILQIPAF LCRQTDLIVEV
SQTNAI INIKKGQFMNPKDMQYSVLKALKTRDKSIQSPTYETALKNGVWLCERGSSFGYG
NLVVDMRSLKIMREFAPVIFDATHSVQMPGGANGKSSGDSSFAPILARAAA AVGIDGLFA
ETHVDPKNALSDGANMLKPDELERLVT DMLKIQNL F

>PROKKA_00412 Carbonic anhydrase 1
MEAFGLALEFQENEYEELKELYESLKTQKPHTLFISCVDSRVVPNLITGTPGELYVIR
NMGNVIPPKTSYKESLSTIASIEYAIHVGQVNLIIICGHSDCGACGSIHLINDETTKAKT
PYIADWIOFLEPIKEELKNHPQFSNHFARKSWLTERLNVRLQLNNLLSYDFIQERVMDNE
LKIFGWHYIIETGRIYNYNFESHFFEPIEETIKQRKSHENF
>PROKKA_00413 Orotidine 5'-phosphate decarboxylase
MQLCVALDLEKKEDNLSLLQELKGLDLWAKVGLRSFIRDGAVFLDEIRKIDGNFKIFLDL
KLYDIPYTMANAALCAKLEIDMLTVHLSSAKSALTILMQRLNALKKRPLIMGVSAITSF
SEEEFLMVYNAPLKTQAITLSAMGKESGIDGVVCSVFESLAIKEALGKNFLTTPGIRLN
KNDKEDQERVANAKEAKQNLSDFIVVGRPIYQAKEPREVVLELLKDC
>PROKKA_00414 Pantothenate synthetase
MRVLETIAALREYRKSLSKESVGFVPTMGALHRGHQSLIERSLKENPHTIVSVFVNPTQFG
ANEDFSTYPRPLEKDLCCEGLGVNAVFAFERSEMPYEIEQRLKLYAPKFLSSSLEGAV
RHGHFDGVVQVVLKMFHLVNPTRAYFGKKDAQQLLIIIEHLVKDLLLLDIEIVPCEIVRDS
HLALSSRVNLYNATQKQALAIKALENIKQAIKGEKACEKLKKLGLKILETLEVDYLE
FCNHKLEPLKTIEPTNTLILVAARVGKTRLLDNLWV
>PROKKA_00420 Helicobacter outer membrane protein
MKSQSLNLISKDIYEKNPFTLSLSLSASSLLNAEDNGFFISAGYQIGEAQMVKNTGELK
KLSDTYENLSNLLTNFNLLNQAFTNASSPSEINATIDNLKANTQGLIGEKTNSPAYQAVY
LALNAAVGLWNVIAYNVQCQPGNSGEQSVTFEGQPGHNSSSINCNLTYNNGVSGPLSIE
NFKKLNQAYQTIQQALEQDSGFVLDKSGKQVTITITTTQTNQNGANKSETTTTTTTTNDQ
LLQEASKMIDVLTTCNCPWVNTAHNSNGGAPWGLNTTGNVCQVFATEFSAVTSMIKNAQEI
VTQAQSLNNQQSNQNAPODFNPYTSADRAFAQNMLNHAQAQAKILELADQMKKDLNNIPS
QFITNYLAACRSRGTPSEGVTSNTWGAGCAYVEETITALNNSLAHFGTQAEQIKQSELL
ARTILDFRGLSNLNSTYNSITTTASNTNPSFLKNLISQSTNPNPGLQAVYQVNQSA
YSQLLSATQELGHNPFRRVGLISSQTNNGAMNGIGVQIGYKQFFGEKRRWGLRYYGFFDY
NHAYIKSSFFNSASDVFTYGVGTDVLYNFINDKATKNNKISFGVFGGIALAGTSWLNSQY
VNLATFNNFYSAKMNVANFQFLFNLGLRMNLAKNKKKASDHAAQHGVLELGVKIPTINTNY
YSLTGTQLQYRRLYSVYLYNVFAY
>PROKKA_00421 60 kDa chaperonin 1
MAKEIKFSDSARNLLFEGVRQLHDAVKVTMGPRGRNVLIQKSYGAPSITKDGVSVAKEIE
LSCPVANMGAQLVKEVASKTADAAGDGTATATVLAISIFKEGLRNI TAGANPIEVKRGMD
KAAEAIINELKKASKKVGKKEEITQVATISANS DHNIGKLIADAMEKVGKDGVI TVEEAK
GIEDEL DVVEGMQFDRGYLSPYFVTNAEKMTAQLDNAYILLTDKKISSMKDILP LLEKTM
KEGKPLLI IAEDIEGEALTTLVNVNKL RGVNLIAAVKAPGF GDRRKEMLKDIAVLTGGQVI
SEELGLSLENAEVEFLGKAGRIVIDKDN TTVIDGKGSHDVKDRVAQIKTQIASTTS DYD
KEKLQERLAKLSGGVAVIKVGAASEVEMKEKKDRVDDALSATKAAVEEGIVIGGGAALIR
AAQKVHLNLHDDKVGYEIIMRAIKAPLAQIAINAGYDGGVVVNEVEKHEGHFGFNASNG
KYVDMFKEGIIDPLKVERIALQNAVSVSSLLLTTEATVHEIKEEKAAPAMPDMGGMGMGM
GMGMM
>PROKKA_00422 10 kDa chaperonin
MKFQPLGERVLVERLEEENKTSSGIIIPDNAKEKPLMGVVKAVSHKISEGCKCVKEGDVI
AFGKYKGAEIVLDGTEYMLELEDILGIVGSGSCCHTGNHDKHHAKEHEACCHDHKKH
>PROKKA_00423 DNA primase
MILKSSIDRLQLTIDIVEISSYVDLRKSGSSYMACCPFHEERSASFVSNQVKGIFYHCFG
CGASGDSIKFVMEFEKLSFVEALEKLAHRFNIALEYDKGVYYDHKEDYH LLEMVSSLYQE
ELFNAPFFLNLYLQKRLSLESIAFKLGLCTNRIDYGIENKGLNKDKLIELGLV LKSDKK
DKTYLRFLDRIMFPIYSPSAQVVGFGGRTLKEKA AKYINSPQSKLFDKSSLLYGYHLAKE
HIYKQKQVIVTEGYLDAILLHQAGFKNAIATLTALT PSHLPLKKGEPEILL SYDGDKA
GRNAAYKASMLAKEQRRGGVILFENNLDPADMIANGQIETLKDLLSRPMAFIEFVLRM
ASSYLLDDPLEKDKALKEMLGFLKNFSLLLQSEYKPLIAALLQAPLHVLGIRERASFQPF
YPKTEKPNLAQKFAHPNTMSLEFLEKLVIRYLL EDRSLDLAVGYI HSGVFLHKKQEFD
ALCQEKLD DPKLVALLLDANLPLKKGGFEKELRL LILRYFERQLKEISKSPLSFSEKMNA
LKKARQAIMKLKQGEIVAI
>PROKKA_00424 hypothetical protein
MKKIKALALFSGGLDSLMSKLLIDQGIEVTALHFNIGFGGNKDKREYFENATAQIGAKL
LVCDIREQFFNDVLFKPKYGYGYFNPCIDCHANMFRNAFYKMLELDAD FVLSGEVLGQR
PKSQRKEALNQVRKLVREVGEARFDKVLDR TQASGEKPQFLDELLLRPMSAKLLEPTFM

EKKGFVDRDKLLDVSGRGRTRQLQMIKDYGLKYEEKPGGGCLLTDIQVSNKIKNLKEYRE
MVFEDGVIVKNGRYFVLPNARLVVARNEEENHKLDIEHPLMDKIELLSCKGPLSLVDKN
ASKEDKELAGRIALGYAKTLKNQAYLIQIGNEKRELYPLDKESAREYLFA

>PROKKA_00425 hypothetical protein

MQEFLGFGVVGNFAGHLEQAGESHSFINMKSEEKDA PKGLFPFYIPYENCYLGRCCINNH
KIILPSDLNLKVQAEPEIALECDVKYDEKHLVTKLVPNFFMAFNDA SVRNLDAAKLSQKK
NFSPASKGIGQKLPIDRFVYGGVCNNFSIASFLKYNNVWHIYGENSKLLKYEFFYQKLLD
WIKDQLNHQQDGDLSLEAIRPFLERHNFPTKMIFAIGATPYMPFAQEHFLQKGDEVVIVAY
NHLQYSFEKIQNLLLEEDTLQAKEHANLSYVYQIVE

>PROKKA_00426 TrbC/VIRB2 family protein

MSAHFLKIIFLVGMCVSSSLFAEGLEGFFNALEAQLKSPIAKGILMVIFIGIAIYVWRNLD
RWKEILFTILGVVFGIFLFFKAPSLANWFMGIF

>PROKKA_00427 hypothetical protein

MIILSASVKNLREISVKEKFLWLNAKSYLISVFAPFILLPWIDLLSAFLLYLGFALFSV
LEFFDEDIADIIVAKSIIKTKTKYYRA

>PROKKA_00428 Type IV secretion system protein virB4

MLEKLLSAIKQKVSNYFLGVLPKSYSMSSEENNILGLYDEHFLLTKNENLVGILRLEGVSY
THLSTEQLQDLFTEQMALDSLEKVVARLVVKKRKIDHQQNIQTDSKYLQAILNQFENKE
VYENQYFLVLESASLSLQGVLEHKKKSLMHANRENFKDILSYKAHFLQETLKSLEIQLKNY
APKLLSSKEVLNFYAEYINGFDLPLKPLVGGYLSDSYIASSTTFEKDYFIQESFNQKTYN
RLIGIKAYESERITSIAVGVLlyQETPLDIIIFSIEPMSVNKTGLFLKERAKFSMSNLVKN
ELLEYQELVKTKRLSMQKFALNIIKAPSLENLDAQTSVLVGLLTKENLVGVIETFGCLKG
GYFSFFPERIHLNHRLRFLTSKALACLMVFERQNLGFKANSWGN SPLSVFKNLDYSPFLF
NFHNQEVSHNNAKEIARVNGHTLIIGATGSGKSTLISFLMMSALKYQNMRLAFDRMQGL
YSFTKFFKGHYHDGQSFSINPFCLEPNLQNLFLQSFFLSMFDLAPSRDKEALEDMNAVF
SAIKSLYETLYPKAFSLDDFKETLKRTSSNQLGLSLEPYLNNPLFNALNDAFNSNAFLNV
INLDAITQNPKDLGLLAYLYFYKILEESRKND SGFLVFLDEFKSYVENDLLNTKINALIT
QARKANGVVVLAQDIYQLSGVKNAHSFLSNMGTILILYPQKNARELKHHFNVPLSETEIS
FLENTPLYARQVLVKNLGNSSNMIDVSLESLSGRYLKIFNSDSSHVNKVKALQKDYPTW
REKLLKS

>PROKKA_00429 hypothetical protein

MDLEHFNLTLYEESPKQAYEYSKQFTKKKKNALLWDLQNGLSALYARDYKTS LGVLDQAE
QRFDKTQSAFTRGAGYVGATMINDNVRAYGGNIYEGVLINYYKAIDYMLLNSAKARVQF
NRANERQRRAKEFYEEVQKAIKEIDSSKKHNINMERSRVEVSEILNNTYSNLDKYEAYQ
GLLNPAVSYSLSGLFYALNGDKNKGLGYLNEAYGISQSPFVAQDLVFFKNPNRSHFTWII
EDGKEPQKSEFKIDVPIFMIDSVYNVSIALPKLEKGEAFYQNFTLKDGEKVT PFDTLASI
DAVVASEFRKQLPYIITRAILSATFKVGMQAVANYLGFVGGVLVTSLSYSGVSTFADTRST
SIFAHKIYLMRIKNKAFENYEVRAIDSIDAFSFSCLKPCRSLNPKVIDARELLSGFVAVP
QIFCSNRHNILYIRSFKNGFVLSHLK

>PROKKA_00430 Chemotaxis protein CheV

MADSLAGIDQVTS LHKNNELQLLCFRLGKNKDLYAVNVFKIREVVKYHGNLTII SHENNS
LVEGLIIIRELTIPLIDMKKWFYYSQNKNDLRPYRIEKEKGEDDIVMICEFSRWTIGV
RIYEADRILSKKWTEMEQSAGLGGSAGNNKLVSRTYFDGRLVQVVDIEKMLIDVFPWIE
DEKHNDELTLSKIHSNQCVLLADDSPSVLKTMQMILDKLGVKHIDFINGKTLLEHLFNPT
TDVSNIGLIITDLEMPASGFVIVKQVKN DPLTSKIPIVNVSSMSGSSNEDMARSLKADD
FISKS NPKDIQRVVKQFLELA

>PROKKA_00431 Carboxynorspermidine/carboxyspermidine decarboxylase

MKKYSAIPTPCYVLESERLEKNAKILEIVRQQSGAKVLLALKGYAFWREFGILRQKLNGC
CASGLYEAKLAFEEFGGRESHKEICVYSPAFKEAEMSAILPLATSII FNSFYQYATYKDR
ILDKNKQLENLGLSPIKMGLRINPLYSEVTPAIYNPCSKMSRLGITPSGFEEKGVKEHGLE
GVSGHLHFHTHCEQNADALCRTLEHVEKHFRPYLENMAWVNF GGGHHITKSDYDVNLLIQ
IKDFKERYHNVEVILEPGEAIGWQCGLIASVIDIVQNDQEIAILDASFSAHMPDCLEMP
YRPSIFKISVENDEELIEVEKGENQGAFSYFLGGPTCLAGDFMGSSFSFETPLKRGDKIVF
QDMLHYTIVKNNSENGVPLPSLARLDQQGFKILKNFSYEDYKNRN

>PROKKA_00432 lipid A 1-phosphatase

MKKFLFKQKFCESLPKSF SKTLLALSGLLILGIFAPFPKVPKQPSVPLMFHFTHEYARF
IPTILSVAIPLIQRDAIGLFQVANASIATLLTHTTKRALNHVTINEQRLGERPYGGNFN
MPSGHSSMVGLAVAFLMRRYSFKKYFWLLPLIPLTMLARIYLDMHTIGAVLAGLGTGMLC

VGLFTSPKKT

>PROKKA_00433 Phosphoethanolamine transferase EptA

MTSLFHLKFLKPLSCLQAGLLYSLIFGVLYHFPLFAYVYKESNQVSFIAMMVVLFVCVNG
ALFLALGLISIHLMRLSAIVFSLNLSVAFYFISAYKVFLNKSMMGNVLNTNTHELLGFLS
VKLFIFIVVFGVLPGYIIYKIPLKNSSKKAPFLAILALVFIFIASALVNAKNLWFDKHA
KFIGGLILPFAYSVNAFRVSALKFFAPTIKPLPLFSPNHSHSFVVVLVIGESARKHNYALY
GYQKPTTPRLSKRLENHELTTFNATSCATYTTASLECILDSSFKNLSNAYENLPTYLTKA
GIKVFWYSANDGEKNVKVTSYLKNYELIQKCPNCEAIAPYDESLLYNLPDLLKEHSNENV
LLILHLAGSHGPNYDNKVPLNFRVFKPYCSSADLSSCSKESLINAYDNTIFYNDYLLDKI
ISMLKNAKQPALMIYLSDHGESLGEEAFYLGHPKSIAPKEQYEIPFIFYANDLFKKEHS
IIQTQTPINQNVIFHSHVLGVFLDFKNPSAVYRPSLDLLKHKKE

>PROKKA_00434 Helicobacter outer membrane protein

MKKKFLSLTLGSLLVLSALSAEDNGFFVSAGYQIGESAQMVKNTKGIQDLSDSYERLNNLL
TSYSALNTLIRQSADPNAINNARGNLNASAKNLINDKTNSPAYQAVLLALNAAAGLWQVM
SYSISVCGPGSDKNKNGGVQTFENVPANGETTITCDSFYEPGKWSAISTENYAKINKAYQ
IIQKAFGASGKDIPALSDTKELNFEIKGKKNDNAQGERWKFPWTNGEFVSIKWVNGTYT
ETKEDIKVSNNAQELLKQASTILTTLNEACPWLSNGGAGGVGGGNSLWVGIDKGDGSACG
IFKNEISAIQDMIKNAEIAVEQSKIVAANAQNHNLDTGKTFNPYKDANFAQSMFANAKA
QAEILNRAQAVVKDFERIPTEFVKDSLGVCHEKGS DGNLRGTPSGTVTSNTWGAGCAYVG
ETVTNLKNSIAHFQDAEQIHNAQNLAYTLANFSGQYKKLGEHYDSITAAISSLPDAQSL
QNVVSKKTNPNSPQGIQDNYYIDSNIHSQVQSRTQELGSNPFRRAGLIAASTTNNGAMNG
IGFQVGKYKQFFGKNKRWGARYYGFDVYNHTYNKSQFFNSDSVDVWTVYGVGSDLLVNFINDK
ATKHNKISFGAFGGIQLAGTSWLNSQYVNLANVNNYYKAKINTSNFQFLNLGLRTNLAR
NKRRGADHSAQHGMELGVKIPTINTNYSSLGTTLQYRRLYSVYLNIVFAY

>PROKKA_00435 Citrate synthase

MSVTLINNENNEREYEFETIECTRGPKAVDFSKLFETTGFSSYDPGYSSTAGCQSKISYIN
GKKGELYRRGHRIEDLVAKYKYVDVCKLLLTGELPKNQDESLEFELELRHRSFVHESLLN
MFSAFPSNAHPMAKLSSGVLSILSTLYSTHQNMMTEEDYQTMARRIVAKIPTLAAICYRNE
VGAPIIYPDIARSYVENILFMLRGYPYSRLKHTTQGEVEITPLEVEAFDKILTLDHAGQ
NASSTTVRNVASTGVHPYAAISAGISALWGHHLGGANEKVLLQLEEIGDVKNVDKYIARV
KDKNDFNFKLMGFGRVYKSYDPRAKILKGLKDELHQKGVKMDERLSEIAAKVEEIALKDE
YFIERNLYPNVDFYSGTILRALKIPVRFFTPVFVIGRTVGVCAQLLEHVKSPQARITRPR
QVYVGD

>PROKKA_00436 Isocitrate dehydrogenase [NADP]

MAYNPKILQKPKEGEEITIKDNKLHVPNHPIIPFIEGDGIGSDITPAMIKVVDSAVQKAY
KGEKKIAWYEVFVGEKCYQKFKDHKELSPEEQWLLPDTIEAINHYKVSIGPLTTPIGEG
FRSLNVALRQKMDLYVCLRPVRWYGSPSPVKEPQKVDMVIFRENSEDIYAGIEWQEGSAE
AKKLIHFLQDELKVKKIRFPESGIGVKPISKEGTERLVRKAIEYAIIDNDKPSVTFVHKG
NIMKYTEGAFMKWGYALAQKEFNAQVIDKGWPWCSLKNPKNGKEIIKDMIADAFLOQILL
RPSEYSVIATMNLNGDYISDALAAMVGGIGIAPGANLNDTVGMFEATHGTAPKYAGLDKV
NPGSIILSAEMMLRHMGWVEAADLIVSAMEKAIKSKKVTYDFARLMDGAKEVKCSEFASV
MIENM

>PROKKA_00437 hypothetical protein

MIKFVRNVVLFILTAIFLALMLIVSYCMPHYSAAVISGVEVKRMNENENTPNNKEVKILA
RDVYFVQTYDPKDKKSVTVYRNEDTRFGFPFYFKFNSADISALAQSLVNQQVEVKYYGWR
INLFNMFPNVIFLKPLKESADISKPIFSWILYALLMGFFISASSVCTLFKSKAH

>PROKKA_00438 ATP-dependent dethiobiotin synthetase BioD

MLFISATNTNAGKTTCARLLAQYCNACGVKTILLKPIETGVNDAINHSSDAHFLQDNRL
LDRSLTLKDISFYRYHKASAPLIAQQEEDPNAPIDTDNLTQRLQNFTKTYDLVIVEGAGG
LCVPITLEENMLDFALKLKAKMLLIHSDNLGLINDCLLNDLFLKSHQLDYKIAINLRENN
TAFHSVSLPYIELFNTRSNNPIVIFQQSLKVLMFALK

>PROKKA_00439 hypothetical protein

MLLNYDFLEFVDEPQRNNSLSLTASIDKALADRKLARQNKPRVRIKDNLSLRATLDAFGE
ELSRLKREQNTETIKGSTIESANLNSIVNKIRSGLPFGTISAFRPFKDAFIKDFTEKQOK
ALISAYKSGYTKADADKITQYWQTKPTKIDLHKPIAKDFFKGNTNIYRTLRLNLFQKFM
DSYIAPKSEATMKDFMSSDKFVKYRYTQKDNMKRTQQLKSLLDKSRHFLGYIQVIGYWK
EDIKDDLIPDKETSFFVFTNEPSSFTDLASYLLLLARLFNQEAICYCDNAKTDKVELVSA
LPNDFGDVWAKFTDITFTIPLTQISITRLHNKVYTFEKKNMENYGVSFEELESTTKLKAMG

MPKDIEGFNISCKAYIERHIKTSIYSSSLGVGRNYPGRFTDMDKVQQYELQAIKRLDLQRC
 SKSKSFKASYNHNIKVNNLVKALRQGKKVSRTLIAKVLANTIDTDAGYCFISPTDLATQL
 GNISPRLSKSIVTAIEQAQGLRLTYALIDKVSYNLHNILSFIFDIDNPLSDQAFEELVI
 EVPREALKNVKLPQIKNVLTQSIFDGTYYQFRN
 >PROKKA_00440 Universal stress protein family protein
 MNILFGISDTQECYNAIKFAVKLAHSLKEVRFTLLHVSMEVFIYSESGMMDYGQTEALEE
 EKAKALLKQFEDAFKKENIECESVLKSGDLIDVVLDMADYDLLLLIGASESNLLYRLFIS
 HQNSLVEQSSIPVVIK
 >PROKKA_00441 ATP-dependent Clp protease adapter protein ClpS
 MKMYNIPTPTMAQVIMVDDPITTAEFVISALKDFFDKSLEEVKALTSSIHGCGVCGVY
 PYDIARHRAAWVRDKARELDFPLKLLVEEIK
 >PROKKA_00442 ATP-dependent Clp protease ATP-binding subunit ClpA
 MAKFNQDLNEVLNQALNLALDLNHALCTTEHVLLVILEHESGEKIIIGALERDDYDKLKQI
 LKDYLLQYVPLKSDPAKMPARSFVLLRMLKRMAYASCFESVGVEELLILMLDHPDCYASKL
 MDSFGIARLYSNPALDLDNHHGIPNGINDNEEAPKNTPLKKYAKNLSALAQDNALDPVIG
 REEEILRVIEILGRRKKNPPLLIGEAGVGKTSIAEALALKIAQKEVPEFLQEYEVYSLDL
 ALMVAGAKYRGDFEKRLKKTLEIQQNGRIILFIDEIHTLLGAGSSNAGSLDAANILKPV
 LTDGGLKCLGATTFFEEYRSVFEDKAFNRFSVIKVEEPSKEACYLILKKIAPLYEEHHQ
 VRYDESVFACVDLTSDYMHDKFLPDKAIELLDEVGSRKKINPKKGKKIGVDDVKETLAL
 KLKIPKMRLSSDKALLRNLEKSLKNKIFAQAEAISLVSNAIKIQYCGLSAKNKPVGSLF
 FVGPSGXGKTELAKELALNLNLHFERFDMSEYKEAHSAKLIGSPSGYVGFEQGGLLVNA
 IKKHPHCLLLLDEIEKAHSNVYDLLLQVMDNATLSNLDGNQASFKHVILIMTSNVGSKDK
 DALGFFSAKNTKYDKAVKELLTPELRSRIDAIVPFNALSLEDFERIVSVELDKLKALALE
 QDIALKFHKEVLKFIAQKSYQTTLGAREIKKIIHNEIKTKLSLLLLQSFKKPCKIACLL
 EKNQLVLKEIKRAQKVKENDF
 >PROKKA_00443 Aspartate 1-decarboxylase precursor
 MTFEMLYSKIHRATITDANLNYIGSITIDEDLAKLAKLREGMKVEIVDVNNGERFSTYVI
 LGKKKGEICVNGAAARKVAIGDVIIILAYASMNEDEINAHKPCIVLVDEKNEILKKG
 >PROKKA_00444 Nucleoid-associated protein
 MDFSQGLGSLGLLDGMKKEFSQLEEKNDTIHTSKSGGMVSVSFNGLGELVDLQIDDSL
 LEDKEAMQIYLSALNDGYKAVEENRKNLAFNMLGNFAKL
 >PROKKA_00445 hypothetical protein
 MFHKAFTFIVLWFFLNLGAYDFKHCQAFFKKASLQKGGVALKELPKGVYLYYSKTYPKH
 ARVIKSDPFVGLYLLQSAPSEYVYTLRDLKDALIRPMASIGDKEALEAQLLVGQKGYDR
 YAQISQKIQKNGVISNICYQMLGLGVGGNGFIETKFIKRFLNQKEPYYGDIGVRLEEYNK
 RLVVVQFDPFFPKNPFLKNDEILAINHQKIRSLAEFEWVSVNLKYQSLVKVEIKRNHKIK
 EVTLKANKRYGGFLLKDTFLERYGIALDERFIITKIGSHLPKGLDFLKLGDRLWVNRKN
 VASNPKALREALSAPKIELLVWRKGFEFYIKVR
 >PROKKA_00446 TrbL/VirB6 plasmid conjugal transfer protein
 MKNDAYEIIILSWFITPLTAILGRFAEFFLYTLHAQLVFNSVVALAFMLFAYRSLKEQNFF
 SASALTEALWFGFFALFNALKNPMHFYEYFQNAIFIAPNMIAQSLSQSLSNFSNHALS
 LDFIFNHGFFYALSFISDLSHNEMSVWLFLSILQGLFLSVLFAIIILVYLEVHVWCSLGLV
 FLAFGFFKTWRSVVVICLKCFALGFYKPFLLLVGFLNVSVTKALIDAHMQEKQDLSLLL
 VVALFLCCVFIIGVPPFINALFRVQNSLKETYKLATNLSANLSQNALNSLQYITTPAPS
 SVSSSMNESVSKEKETHSPTFKVETTQLDVKIPNFKQKKVKKDTINTKNEI
 >PROKKA_00447 hypothetical protein
 MRIFVIMGLVFFGCTSKVHEMKSPCTLYENRLNLA
 >PROKKA_00448 VirB8 protein
 MKEKPFNSEQLIYLEELLHQEKHLENKLSGFSVNDLDMQSVFRLERNRLKIAYKLLGLM
 SFIALVLAIVLISVLPQKTEHHFVDVLNQDKHYAIIQRADKSISSNEALARSLIGAYVL
 NRESINRIDDKSRYELVRLQSSSKVWQRFEDLIKTQNSIYAQSHLEREVHIVNIAIYQOD
 NNPIASVSIAAKLLNENKLVEYKRYKIVLSYLFDTPDFDYASMPKNPTGFKITRYSITEI
 APTNRND
 >PROKKA_00449 Conjugal transfer protein
 MRKFLYALMGFLLVFSVLKADDFLEENETAPAHNLNHPMQDLNAIQGSFFDKNRSKMSNT
 LNIDYFQGQTYKIRLRYAMATLLFFSKPISDFVLGDKVGFDAKILESNDRIILLIKPLQIG
 VDSNISVIDNEGKIFSFYVFSTTFTSSKHPNLQVFIEDKNYYSNAFMKPQNKENAFENAP
 TNNKPLKEEPLKAGKEEAKEKEETIIIGDNTNAMKIVKKDIQKGYRALKSSQRKWYCLG

ICSKKSKLSLMPEEIFNDKQFTYFKFDKKLALS KFPVIYKVVDGYDNPVNTRIVGDYIIA
EDVSAKWTLRLGKDYLCIRFVKKGKDE

>PROKKA_00450 Type IV secretion system protein PtlG

MNKWLKGALVFVGGFATIITISLIYHQKPKAPLNNQPNLLNDDEVKYPLQDYTFTQNPQP
TNTESSKDATIKSLQEQLKAALKALKSKEMNHSKEETFKSPPIDLKANTTPPKKDFSLKQ
WDLLASRITPFPKQSPKNYEENLIFPMDNPKGIDGFTNLKEKDIA TNENKLLRTITADKMI
PAFLITPISSQIAGKVIAQVESDIFAHMGKAVLIPKGSKVIGYYSNNNMGEYRLDIVWS
RIITPHGINIMLTNAKGADIKGYNGLVGELIERNFQRYGVPLLLSTLTNGLLIGITSALN
NRGNKEGATNFFGDYLLMQLMRQSGMGINQVVNQILRDKSKIAPIVVIREGSRVFISPT
DIFFPIPRENEVIAEFLK

>PROKKA_00451 Alginate biosynthesis protein AlgA

MTRKPMKIKNILLSGSGSKRLWPLSRSLYPKQFLKLFGNKSLFELSFKRNASLVDETLVV
CNEKHYFLALEEIKNEIQNKSVGFLLLESLSKNTANAIALSALMSEREDLLIVTPSDHLIK
DLQAYENAIIKKAIGLAQKGFLVTFGVSIDKPNTEFGYIESPNGLDVKRFIEKPSLEKAIE
FQKSGGFYFN SGMFVFQAGVFLDELKKHAPTILKGCERAFETLENAYFFEKKIARLSEKS
MQDLEDVSDIALMQQSHKIKMVGLNAKWSDLGNFNALFEEVANEPKENVSLNQTPIFAR
ESANNLVFSHKVSALLGVENLAIIDTKDALLIAHKDKAKDLKALVSEIEINNQELLQTHT
KVYRPWGSYEVLHESGCYKVKILEVKPNARLSLQKHFRSEHWVVISGMASVELDHKMFE
LQANESTYIPKNTLHRLANYGKIPLIIIEVQVGEYVGEDDIVRIDDDFN RQNQNA

>PROKKA_00452 GDP-mannose 4,6-dehydratase

MKEKIALITGVTGQDGSYLAEYLLNLGYEVHGLKRRSSSINTSRIDHLYEDLHSDHKRRF
FLHYGDMTDSSNLIHLIATTKPTEIYNLAAQSHVKVSFETPEYTANADGIGTLRILEAMR
ILGLEKKTRFYQASTSELYGEVLETPQ NENTPFNPRSPYAVAKMYAFYITKNYREAYNLF
AVNGILFNHESKVRGETFVTRKITRAASAIAYNLTDCLYLGNLDAKRDWGHAKDYVKMMH
LMLQAPTPQDYVIATGKTTSVRDFVKMSFEFIGINLEFQDTGIKEIGLIKSIDEKRTNAL
NLNLSHLKKDQIVVRIDERYFRPTEVDLLLGDP TKA EKELGWVRECDL KELV KDMLEYDL
KECQKNLYLQDGGYILRN FYE

>PROKKA_00453 GDP-L-fucose synthase

MNETILITGAYGMVGQNTALYFKKNKPDVTL LTPKKSELCLLDKDNVQAYLKEYKPTGII
HCAGRVGGIVANMNDLSTYMVENLLMGLYLFSSALDLGVKKA INLASSCAYPKYAPNPLK
ESDLLNGSLELTNEGYALAKLSVMKYCEYVNAEKG VFYKTLVPCNLYGEFDKFEEKIAHM
IPGLIARMHTAKLKNEKEFAMWGDGTARREY LNAKDLARFIALAYDN IASMP SVMNVGSG
VDYSIEEYEEKVAQVLDYKGVFVKDLSKPVGMQQK LMDISKQKALKWELEIPLEQGIKEA
YEYYLKLLEV

>PROKKA_00454 Hydrogenase isoenzymes formation protein HypE

MDSVTLACNGGGKETNALIERVFMPYLKEWIVAFDE DAPKFKASGEYCVSTD SFVITPLI
FNGGDIGKLCVCGSANDVSVQGGEPLYLNMGF IEEGLEISLLKQILQSIQKELFKANLK
LLSLDTKVVPKGSVDKLFINTTCIGKIIKPGISSCHLKQGQAII LSDTIANHGASLFAMR
HEIKLKTNLESDCQLLYPLLKPLFLSDLKIHALRDATRGGLASVLNEWANSSRVKIVIEE
EKIPLKEETKGICEILGLEPYTLANEGV FVLALHQKDAPKALKILKSNEKAKNACVIGEV
LENPYPSVVLKNAWGFERILETPEGELLPRIC

>PROKKA_00455 Carbamoyltransferase HypF

MCNDATLLNQITLFGVVQGVGMRPFIIYTLAQKLELVGFTRNTQA ALEIILPAHKTESFLN
ALKKGLPPLALVEKIIISPYDKVLHFNDFRILESKNHPLNLLSQIPKDLGVCKDCLREIR
DKNSPYFH YAFNSCAKCGARYSLLNAMPYDRENSAL KPFKLCKFCASIYQDPTNKR FHIQ
GISCKKCGIALTYKRFKNDDALLECAKDLQKGKII ALKGLGGFVLLCDARNFQTIERLRL
LKNRPLKPFVLMFKDLNSAKQHAFNLNALECESLNSASAPI LLVRKKPNTPLAPNIAKNSP
FYGVILPYTPLHALLLDLLDFPIIFTSANFNSLPLASNEKEIDSLHFIFDFKLTHNRAII
HRIDDSIVQRVDNIIRPMRLARGFAPLYLTLPKRSNGSPKKI LALGAEQKGHFSLDSET
SILLSPFCGDL SVLENEKHFKETLNFFLKTYDFKPTILACDKHKNYTTTQMAFD FNTPL
LQVQHHHAHFLASVLDALLQDPHLNHPF IGIVWDGSGAYENKIYGAECFVGDLERIEEIA
RFEWFLLGGQKAIKEPKRLVLEIALKHQLNKLKRVQKHFKEDELEIFQQMHDKKIQSI
ATNSIGRLFDIVAFSLDTGTISFEAESGQVLENLALQSDESA FYPF TIKNSVVG LKDFY
QAFEKDLGVLEPERIAKKFFNSLVEIITALIVPFKEHV VVCSGGVFCNQLLCEQLAKRLR
GLKRQYFFHKHFPNDSSIPIGQALMAYFNPTIIKKG

>PROKKA_00456 Agmatine deiminase

MKRMLAEFEKIQAILMAFPHEFGDWAYCIKEARESFLNIIQTI AKHAQVLVCVHTSDTIG
YEMLKNLPGEIARIDTNDTWARDFGVISVENQGVLECLDFGFGNGWGLKYPSNLDNQVNS

KLKSLGFLKHPLKTMPIYLEGGSIESDGAGSVLTNTQCLLEKNRNPHLNQNGIENMLKKE
LGAKQVLWYSYGYLKGDDTNSHTDTLARFLDKDTIVYSACEDKNDHYTALKKMQEELKT
FKKLDGTPYKLIPLLEIPKAIYDENQQRLPATYVNFLLCNNALIVPTYNDPKDALILETLK
QHTPLEVIGVDCNTLIKQHGSLLHCVTMQLY
>PROKKA_00457 Modification methylase DpnIIB
MILKSPVVSGLKTKHPTQKSLALMEKIISIHTNPNDIVLDPFMGSGTTGLACKNLERKF
IGIESEKEYFQTAQKRLNLF
>PROKKA_00458 hypothetical protein
MSCNTINNILKIKLFFCNCDDARVSNFFKYFFVHFQELGLKKLISACYMEPEISSNSNKI
SKGFYCEYEGKKDWQTTIDGLKFFKGDDFRSEESIQLLKEADIVVTNPPFSLFREFVAQ
LIEHSSKKFLIIGNINAITYKNIFTLIKNNKVWLGMHLGRGISSFIVPRHYELYGTETKID
NLGNRLISPNNCLWLTNLDYKKRHEILPLTKKYNKNKYELYDNFDFGINVNRTIDIPLDYR
GSMGVSITFLHKFNPKQFEIIGFRKGNDGKDLSVNGKCPYFRVLIRHKKDY
>PROKKA_00459 Glycosyl transferase family 8
MDTQTPQNQIIPFMSFDKNYALGAGVSLYSLSHASRHTSTIDFSPINQNNQLLGANIV
YKIHCLIKGVTLQKQKLLKTLEPFKTFASLEFIDIDSLDNSIESYLNESCSKRYGGLLV
LCRLLLASLFPNYSKIIISIDVDTVFLGDVASAYFALDNNPTKSLGMVKDTFSHLSFEAFC
DFCKRACKNFKIDFSRFSLELKRHHQGFNMGFLVAHLDRWRQDEFERIALFLKTRGKD
LFYPEQCLVMVFWERILELPIYYNCYSDFKEHYPKNIIMLHFQYKPWRSVSSLNGLR
ICYEAEASFWLANLFCPTPFKNDFFERLEMAKDQQMQSFKTHIRSKTIRDYFCYRIKNIL
KSFCCKTH
>PROKKA_00460 tRNA dimethylallyltransferase
MDAEIFSLSLSIYKDINIASAKPSLKERKDIKHYALDCLSIDEKNNALFKTLLEDAMR
VSSKEILLIVGGSSFYLKSILEGLSRMPKLSNEEVVKIEREIATLSDPYIFLKSIDPNMA
FKIHPNDTYRIHKALEIFYATHTPPSKYFKANPKKPFHAISLFAVSVEKNALHSNIKQR
TKNMLHSGLVVEIKALYAKYPKDSQPFKAIGVKESVLFLEKQLTLKELEETITSNTIKLA
KRQNTFNKTQFNPLYAGSVKEVRHAILKHSKSAIKG
>PROKKA_00461 Ribosomal silencing factor RsfS
MNQRIEMITALLDEKKAFFDITHIDLSKTPYLVEDVVIATTLANKHALSLLDALKNTLKPL
GEVFYQIDESNEEWIILDGLDIIHLFTEECRKKFDLEGFLNAYKRGLPYQNA
>PROKKA_00462 NADPH-dependent 7-cyano-7-deazaguanine reductase
MTPENLNKSLGAKTPYIFEYNSQLLEAFNPNNPLDPLITLECKEFTSLCPITSQPDFGV
IFIRYIPKDKMVESKSLKLYLFSYRNHGSFHESCINMILLDLIQLLEPKYLEVYGDFVSR
GGIAIKPFVNYAIKEYQDFKEKRLLDK
>PROKKA_00463 hypothetical protein
MLENRNVENCYYYDTFDNMNLGNLTLFCKLIKLYARGDEEAQDLGGIWAEMLPNTNKEKALF
VAQLLCDGGINKYGLSCAGLTENLLKDIFNSIGLATAKEVDEHLANMDKEEAELKE
>PROKKA_00464 hypothetical protein
MGFQENENKLKVGALVKATINNKVVEAKVISIGFNRVTLRSEKGNEATYAFNSDKFLKWFN
KVPLTEHMLRHAKNEKSGEDLLKGIKVVTSGSSVKDRTSTPKEPKEKEDKFKLAFGFEEK
SESSFVVAKYDFTDRKSRIGALLTPMFYQGGKHQASAIILTALSCEGLNKHSDLE
>PROKKA_00465 hypothetical protein
MLLDYDFLLLLNDESGNPTRYYYLLQDFEKDFVASKVAQNKTKRFVKEIIGSEKASKTKN
SAIEVSNTKASAMKNETIGSGDLKKVCEKIKSTLPFGIISAFKPFKDAFYRDFNDNEQKL
LIGAAGSGCIQSSADKLDQLKTRLLYWQDKSVKVDWDKPILIKDFFKGNNYLYRRLCFL
GKHFMDFRLKNNAKASVKDFMSSKEFIKRYRTPKQNTERRAKKLSYLENKRDFIGFVQT
LNSLKDNPQDPFLPNEETSFLVFSNEPTIIFNLRDYLLVLAQIFNQQAICYCESKCPIEL
INASPGKDFNKTQDSFSDIKFSTPNQLEQSLNALKNLNAFFSKHPDKHNGIEFNEIAKT
QIEALYLPQSSDDFDDFRKHLEDVSKSFIRAKKNRYGFPKIFDVADIEKEERKVIEWREK
ERASKQSYKQNLQINKIANDLKHDKVVDKRTILSVIDADIDRGFIIPKDLLKQLEKISAS
LSKDIVIAIKQVEKLELNYALIDNIQHNTLDDTLDFTFIIGDSLVSQSLYVTFDLVIDMD
RPMSEQFLNHIGGLGSFESRERALEWVRLAESKLIKAPREALKNAELSQIEEILTGCIF
NGAYRLQNDLKGNRHGNFK
>PROKKA_00466 hypothetical protein
MANNFLKKYPSNRSLCIYRDNKDGYSARGDRNFLSQIKTIPLIQAGGHEEAFGLSFAK
KDFKKVIKSLQAL
>PROKKA_00467 hypothetical protein
MSGAREFKGCYDYLHYLVKHFNPINDPRLSVFNVKEFKRYNDIKKKMKVESEENAQIF

SCNKILVALLR

>PROKKA_00468 DHH family protein

MPKKEPLKMSKKRIFKDFLKEAKQHRPIVFYTDNDCDGMLAGSVLMSMCYRLGIKDDFFFF
SPLRNVHGYGFTDLAINDLLSQPYIFNPKTNQVLRLDCIKNQFQKDPLLFSADLGADLAA
DTQLQEILLERFEQCIITDHHKSFEVGLIDGNKIAYINLNDEKDANYYSGAFTSALVFSQ
IFQTQTTPLEEELIAITLLSDRIDLDHGDNLDMLNLAPVKHERIQCFKDKDLSLAQDD
LDGISNLYGFNCINYINALSRFEWG

>PROKKA_00469 hypothetical protein

MKKSIDRYDKEKERLKFLTAQNYSIFENTVLNHLEFEDNIKKSELNKKFEENTYLRNA
YYFKELLNESMENGSIISDMDDFVKWFYEHIVLTRIICFEQDSAMQIFQVLNDRGQPLSPI
DILKSSLMQEIKQDSEKRKDFITTWDKLVEACKSVEGVDIDLEDFFNMYLEYADPSASKK
RADKGLKKVFKDSKDKACGFIYEISEFMKAYTALLKKQDRYVYLLRYLPSRYWASILT
LYVKYPDFDALKKLLVSYYYQTWIAAGGTITRIKQTSINIIKNVKSNSKSVGTIKDLILN
NSYNTFDQYLYNLWNSSSVYPSKWVRPVLALANYFMADEEKPHFIVMDAETQVEHVLPQS
PKRGSQWNADFDEKREEWVNNIANLTLLKRKKNAQALNGDFDEKRKIYGGKDTSKVIS
YDITKELYSNYRKWNEKSLQERHDFLYNTITPILHIEGQEEEEEEEEFEDDFDLEWLKIAK
HQNNKEVINA

>PROKKA_00470 hypothetical protein

MAKVEVELKKLHQILVDSEYFYQVPDYQRPYVWDKDHGLGALIDDLVGSYTNNEDQYFCGS
IVIAENQKDKRWDVVDGQQRLTSFIILACTILRLYKHRLGQSKAFIEEEYL

>PROKKA_00472 ribonuclease BN/unknown domain fusion protein

MREFFKSVRGFLRLRLMIFPKRLKNAFLGLSELFFYASSLSFYTILSLSPILLFVFSFLV
SHYMQAHSGEALIFPNAPKLIGAIKDFLENFKKTDMTLGTLEEVSIILVALVLCENYR
SIASKIFQAKPRDYAHFKGKEIFLFWGFGTTLVFLFALPLVVFDDIKIQVFFEDKDSNLL
HVLRWIGTYAFFLILFTIPTNKVFKHYFWVFLWVFFTSVSWHVLKWAFTYYVLYNRTYHE
LYGSVSILWFLMSWVYVSWLVILIGMYGCKVCDAFDPKEVFKKFLGFFFKET

>PROKKA_00473 Biotin synthase

MQEIFLCSISNVRSGDCKEDCTYCTQSSHHQGAIKRYKFKDEKVVVLQEARALRELALGF
CLVTSGRELDDEKCEYIAKLAKAINQEELGLHLIACCGRADLEQLEFLRDAGIHSYNHNL
ETSQNFFPKICSTHTWEERFITCENALRAGLGLCSGGIFGLNESWEDRIEMLRALASLSP
HTTPINFFIKNPVLPIDAETLSADEALECVLLAKEFLPNARLMVAGGREVVFKDNDKQEA
KLFEYGINAVVLGDYLTTKGKAPKKDIEKLLSYGLTMATSCH

>PROKKA_00474 Type III restriction enzyme, res subunit

MFEMATGSGKTLVMAGLILECYKQGYQNFIFFVNSASILEKTKLNFTDSVSSKYLFS
NINDENTEIKSINNLNESHNGAINIYFSTIQGLFSLFTKTKENAITLEDLKDQKLVLAD
EAHHLNTETKKKLNDESEAKRNWESVVKLAEQNKDNLLEFSATIPNEKSVEEKYKNL
KAFTYTLKEFSEDKFCKNIYSLSYENKELETRFLGACVSSLYKELLAQHHNIENFKPCIL
FKSERIEDGKENQERFNAFLENLSPLDLENFFHYSRNAFFKDAKNFFDERNYTPNLVAF
QTKFKKSTQINTNNEKELEKGMLLNLSLEDNRNPKRVIFSVDKLNEGWDVNLNLFDIVRLK
NKANQKDTIKDAQLIGRGARYYPFSYNDFKPSRIEFYQKFDKSNPLSALERLDYHAIYD
SEFIAQLKKELQDLGLGLIEKEKTTIPLTPTKRFKCYASNTKDKNKNLNFNKDYSDPVEA
TLQSLHVPLFGFIVREKKVDFKEENKSDKTYIIPHTLDKIPLNYFLKALNLKNLDFKTLK
KAFKKHAFNNKVEFIERYVSSLKTNFHKSQKFDDNKTLLKLAVYIENLKDITLLKEQDKY
EVSALELKEFETHNKSLSASEWEKDIPFYEWLLFKDMRKLDSDLERKFLVFINDNKEVLD
KKFKEWCVLNRNDHFAELKVFCNIEDSPYYAQGFEPDFILFAQTHSDEFLGFTCYMEAKGE
HLEPSNAWKKEFLEMLENATLKSHNKKLHLKGLPFFTFTNNNRTLNAEFVEAFHQTFKD

>PROKKA_00475 Modification methylase DpnIIB

MLKTPLKTLILDILINHFTKERLETILQONDEKLLTFMFEHENANDYKKHFFKMIANSLVF
NEKALLECLEIKELDRSFTFRKKNIGLFSQGGFIKSSGSSVVLHFPFKDNVLLGNAKDSSA
KSNELFYHEVLHKKIDTLLKKKVLRCFEMHGEGDLESALKDKNTNYLIKGNLIALHSL
KKKFAKQVKCIYIDPPYNTGNDSFNYNDFNFHSSWLVMKNRLEAAREFLSDDGVI FVQC
DDNEQAYLKVLMDEIFLRENFVATITCKVKSAGGLTTDTEMI FNCSEYILVYAKSFENLT
YNSIKISKEIINSQSKTSEQYNSIINNVDHRKKEFICEKDNIKYYKISKGNFNIQRLSVK
TMSEQDFYRLKDEIFRLTALSGGIGKKLKNHITDFTNDEDLFMYEYAPSKGKDKGIKTQY
FLYKQTVTMLKNLIKADDNKKIYKLEPLSNIFTEDFWQGISNEGQITLKNKKPEQLI
YRILEIATNENDLVLDFFAGSGTTCVAHKKMRRYIGIEQMDYIETITKERLKKVIEGEQ
GGISKKCDFKGGGSFVYAEKVNNSGIKKQILNAKSASECLKIFNDLNARVLKRADNKMD
EIHSEEFHNLDLNEQKRKCCASLDSNEDYLNLDGIDEDAWGIDEITKKYNEIFYS

>PROKKA_00476 DNA/RNA non-specific endonuclease
MKTFTKNWLCFSLIAMS WLQAGMLDNFTKAINS YTA KKLNEIKDQVNSANPAKNHNPAHNA
NGMLTNIDCKVLKNNFYSVCYSELKNPIYGV SXLFGLVDKNNIEKRYEFKPDTRLAEY
QQAVTQDYKSGSGFDRGHFVANDASDFDFASNPLRETYRMTNITPEAKNTNRHSVLLVEKEG
RNLARKYHQVLVEELTI IKQGYRTFSSKNIAIPSGFWYHYDTS LTDSYENAKSECFYIPN
DNQNYPLQEMRRDCKEYERVEKQVVFKNKNKNTLNL PKYLNNAKKY
>PROKKA_00477 T-protein
MKAGIIGLGLMGSSLGLALQEWGRFKSVIGYDHNALHAKLALTGLGLVDECVGF EKILECD
VIFLAIPVEGII ECLKKMTSVKKSATIIDLGGAKAQIIHNIPKSIRQNFIAAHPMC GTEF
YGPKASVKGLYENALVILCDLED SGTEQVEIAKEIFLGVKARLIKMK SSEH DTHVAYISH
LPHVLSYALANSVLKQNDPEVILSLAGGGFRDMSRLSKSSPLMWK DIFKQNRDNVLEAIE
KCEKEIAQAKAWIENNDYESLAEWMAQANKLQEFM
>PROKKA_00478 Lon protease
MTEDFPKILPLLVEEDTFLYPFMIAPIFLQNNASIKAVAYAKNNKSLVF IACQKDKLNDN
EAPY YDVG VIGSIMREANMPNGRVKLLFNGIAKGRILEPAKENEQGFLEAQISPIEYLEY
DKENIQAIVEVLKEKVITLANVSSLFPPDLIKALEDNDDPNRIADLIAAALHLKRDQAYS
LFANNNTEQRLLDLIDIVIEETKTQKLQKEIKSKVHQMEQTNKEYFLKEQLKQIQKELG
TDKQRDEDLNQYYQKLES IKPFLKEEAFKEIKKQIDRLSRTHADSSDSTLQNYIETMLD
VPFGQYGKKALDIKHVREQLDKDHYSLKRPKERIVEYFATMQLLEMRRKKKPEKKDKTKG
TILCFYGPPGVGKTS LANSIAKAIERPLVRIALGGLEDVNELRGHRRTYIGSMPGRIVQG
LIEAKKMNPVMVLDEIDKVDRSVRGDPASALLEILDPEQNI AFRDHYANFSIDLSQVIFI
ATANNIDRIPAPLRDRMEFISVSSYTPSEKEEIAKNYLIPQELEKHALKPSEVDISHECL
KLII EKYTREAGVRDLRRQIATIMRKLALKYLEDNPHKKGR TKKGENEKSEDKENEKRGE
NKDFCVSITPNNLKEYLERMVFEIDPIDGENKIGIVNGLAWTPVGGDV LKIEALKIRGKG
ELKLTGSLGDVMKESALIAFSVVKVLLDNETLKAPKIPSETDAENKKKKKALKVYNAYDL
HLHVPEGATPKDGP SAGIAMASVMASILCDRATRSEVAMTGELT LSGEVLPIGGLKEKLI
AAFKAGIKTALVPVKNYERDLDEIPA EVREN LNIVAVKNIAEVLEKTLL
>PROKKA_00479 Outer membrane protein assembly factor BamD
MAMIVIGTGCANKKKKKDEYNKPAIFWYQGILREILFANLETADNYSS LQSEHINSPLV
PEAMLALGQA HMKKKEYVLASFYFDEYIKRFGTKDNVDYLTFLKLQSHYYAFKNH SKDQE
FISNSIVSLGEFIEKYPSRYPYVEYMQIKFILGQNELNRAIANVYKKRHKPEGVKRYL
ERIDETLEKETKPKPSHMPWYVLIFDW
>PROKKA_00480 hypothetical protein
MSKIKPQIKKNNPSKFNFSGWHLFGTVCMIGEMVLTLLGGAKNKKK
>PROKKA_00481 Flagellar assembly factor FliW 2
MIFDVKAPILGFETIHKMHLQKVDEIFLRLNSAEDNSVVSFTLVNPFALRKYEF EVPTPL
KILLELEGAKSVLVANIMVQTPIELSTVNYLAPLIFNLDKQLMGQVVLDSNKYPHYHLR
ENILSHTHE
>PROKKA_00482 3-hydroxyacyl-[acyl-carrier-protein] dehydratase FabZ
MEQSHQN LQSQFFIEHILQILPHRYPMLLVDRITELQANKKIVAYKNITFNEDVFNGHFP
NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDEPIAKTKIVYFMTIDKV KFRIPVTPGDRLE
YHLEVLKHKGMIWQVGGTAQVDGKVVAEAE LKAMIAERE
>PROKKA_00483 Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-
acyltransferase
MSKIAKTAIISP KAEIGKGVEIGEF CVIGDHIK LNDGVKLHNNVTLQGHTFIGKNTEIFP
FAVLGTQPQDLKYKGEYSELIIGEDNLIREFCMINPGTEGGIKKTIIGDKNLLMAYVHVA
HDCVIGSHCILANGVTLAGHIEIGDYVNIGGLTAIHQFVRIAKGCM IAGKSALGKDVPPY
CTVEGNRAFIRGLNRHRMRQLLESKDIDFIYALYKRLFRPIPSLRESAKLELEE HANNPF
VKEICSFIESSRGVAYKSSEYSSEEKQEE
>PROKKA_00484 ATP-dependent Clp protease ATP-binding subunit ClpX
MNETLYCSFCKKPESRDPKKRRIIFASNLNKDVCVCEYCIDVMHGELHKYDRMDSLLALK
RDLRLRMESNAYEEEFLLSRIPAPKELKAVLDNYVIGQEQA KKVFSVAVYNHYKRLSFKE
KLKKQDNQDSNVELEHLEEVELSKSNILLIGPTGSGKTLMAQTLAKHLDIPIAISDATSL
TEAGYVGEDVENILTRLLQASDWNVQKAQKGI VFIDEIDKISRLSENRSITRDVSGEGVQ
QALLKIVEGSLVNI PPKGGRKHPEGNF IQIDTSDILFICAGAFDGLAEI IKKRTTQNVLG
FTQEKMSKKEQEA IHLVQTHDLVTYGLIPELIGRLPVLSTLDSISLEAMVDILQKPKNA
LIKQYQQLFKMDEVDLIFEEEAIKEIAKLALERKTGARGLRAI IEDFCLDIMFDLPKLKG
SEVRITKDCVLKQAEPLIIAKTHSKILP

>PROKKA_00485 Rod shape-determining protein MreB
MIFSKLIGLFSDHIAIDLGTANTIVLVKGQGIINEPSIVAVRMGLFDSKAYDILAVGSE
AKEMLGKTPNSIRAIRPMKDGVIADYDITAKMIRYFIEKAHKRKTWIRPRIMVCVPYGLT
SVERNNAVKESALSAGAREVFLIEEPMAAAIGAGLPVKEPQGS LIVDIGGGTTEIGVISLG
GLVISKSIRVAGDKLDQSIVEYIRKKFNLLIGERTGEEIKIEIGCAIKLDPITMEVSGR
DQVSGLLHTIELSSDDVF EAIKDQVREISSALRSVLEEVKPD LARDIVQNGVLTGGGAL
IKGLDKYLSDMVKLPVYVGDEPLLAVAKGTGEAIQDLDLLSRVGFSE
>PROKKA_00486 rod shape-determining protein MreC
MRFYFKFLWLLGIFLIFYFLDFKGSSSYISDRIKNALMSAKNSLADNIQAYFFQAKNIKE
FQKERLILEALKLENADLKERLNSVYPLENPKMIYSPTFMTSFISLEDTHSVSLNPVNL
EENKIYGLVSHNQAIGIAVLEKGRNLNGFLNAHKRCAYSVMIGKNQVLGFIGTNRQELIV
DFIVPSAEINIGDQVLTSGLDGIFGAGVFVGEVSSIEDHYTYKSAVLKNAFLSDAKLLRH
VFLSDVKN
>PROKKA_00487 Type III restriction enzyme, res subunit
MAKKKDLTTDNEIFVAQKLAEEELNTNEINEPLERLDFKSFDSNKELLDYQQQALINAFR
MLVAYFRDFKENKKEFYAFYQKHYSFAHCDFAKKKLNPLLKSHFKVENHCVSFENFINRL
AFYMATGSGKTIVIIKLVELLSVAIRMGLIPKKNIMFFSANENLIKQFEKEIEKYNRGKD
YFKQIDFKSLKSVTHKDFHHPKGSFIEKIPLFYRADLMNDEESKENLLNYKDYWDNGE
NYVILDEAHKGNKSESKRQAI FSLLSLKGFLNF SATFTTEESDLITAVYNLSVGEWVKLG
YGKESVLLKKNLNAFKELKDLNDREKEIALLKALLLLGMQKRYQTEGYFYDPLMLVFTH
SVNVKNSDAEIFFKTLARVIENDDGSDFLKAKEDLLEELKNPEFLFSDDKDKGYKVVKVFK
ESLSKSMDFKGLKEEVFYANNGHIEVIINPKNNQEIAFKLNTSDKVFCLIRIGDITEWICE
KLKSVKVVSKNLSFKEESYFSQIDKSSINILVGSRTFETGWDSTRPSVILFLNIGLDDDA
KKLVKQSFGRGVRIESVKNQRQLAYLDDIDKAVKKALKPNAMLETLFVIPTNYASLEAI
LKFKQKESENKGENRGSWREIKLEKTPIKHALFVPCYRKEQTSILKLPEASAFKMSAKNFK
DLKEYFNLMSSEKHFILKHEIYDPKDYAQLKEMIQAHFKKVSTWHYKDLDCMISEIKGKL
YPNQKVPKDEFNALDSEKIVHFKRIKVKADKKEELVKTIQEVKEYAPL DKETLIKKIAQG
EIDPYDTEKHKQDKTFKVGDDAELLKLKEHYTPLIKAKNCDWLKHVVVKSESDFLEEL
LKITETLQENYDFWAFSKIDEHLNLFIPYIDNSTGERKFFPDFIFWLEKGD TQIIICFID
PKGIKISDYQHKADAYKLFKDKVFSSKDDPNLKIKVVLKFYGNKDGVGERYRDDWIKEGE
LKDFFLTLIA
>PROKKA_00488 Modification methylase MboII
MIWQKKKGGSQDSENF AKEHEYILCYQKEKFTIIDTEIDHDIQDFNKTINGKQAKILKLE
KWGAGALRSDAPSLYYSIKDPGNDFYPIAPNGEEGRWRKKPENLDSEHIFWQENSKGRL
IPYEVIIYYDEIKNAKKVIKTRTIFTEYGTTEATKEILALFNGTKLFDTPKPEALISRIL
EIA TQENDLVLDFFAGSGTTC AVAHKLKRKYIGIEMGEHFDSVILPRLKKVIGGFKSGAA
KEFDGGGVVKVYALESYEEILRKIKYEDNDKPLAYDEQYSDLVECKNESYTLNVEALEKM
GVDIKETLENLWGVGVEFFNEKVVVKFGKNDKEVEILKALKEALI W
>PROKKA_00489 DNA methylase
MKTNEAQFYEVLENLFIGVKIEDKPESLLDPNARVLKNGMINLMKAKSQYYHNKKQELKK
LIDSKCQDNNDLKEELFDKLYSFFKRYLSANGGIYFNDTPLYDSLYTKSGYEKCSLK KDT
ALFYKTKDLYYVKSETIYKDFCFELEGILFNFDASLLESKKYNEKVDLVFTLKD TDAKTN
TLNFSVTLSSKGTQTKISEILKGC FNQGVKLDEEV LKKAFGKFKKQGSMDYFIHNAQGF
LKEQLDLYLFEYLFKEMTAFDLKRLNEINIIKEVALQVIVLVSEFENELCKIWNKPRFVL
NSHFIVSLDKLKAKNYDLNKITNHKNYPKQVKEWQDLNLEITDDLLENEFLPLDTIYFKD
LEEEIKSQFNEDEINGTLIKSENYQALNSLKNRYKEAIDCIYIDPPYNTQNNEFVYADNF
KRSSWLSMMENRLELAHALLSDKG VVVFVSIDDNEQAYCKALMDEVFN GGGVITL
>PROKKA_00490 Transcriptional regulatory protein QseB
MQKKIFLLEDDYLLSES VKEFLEHLGYEVFCAFNGKEAYERLSVERFNLLLLDVQVPEMN
SLELFKRIKNDFLISTPVIFITALQDNATLKNAFNLGASDYLKKPFDLDELEARIKRFFN
DDPIEIMPNI FYHQNCLSVRGKKEILPPKTAQLLEYFLEHKGQIISSQALENNLWEQAID
DSTLR TYIKVLRKLLGKDCIETHKGVGYRFNPL
>PROKKA_00491 sensor protein BasS/PmrB
MLVISVLAFNYEKNEKIKMIRMDMDKMASKIASEVIALHMQTHADYQNALNALISRYKDA
SIALFDNKKRVLYSNIPESANLIK NHKEAGFFSFRGEYYLLSDETFSHLG VAKMLFKNSK
PLHFSSLYRNIVLVFVIAFLCVIGVSVFLGR LFLKPIRNEITRIDHFLKNTTHELNTPMS
TLVLSLXLEDNQQHRIKIAIQRMSFLYRSFSYLV MQDIERESFVLLDLKALIIKENTL
FSEMIDYHKLEFKSDLVGVELKAKEQDFLSLYSNLLMNAIKYSVVNGYIHIELTHAFLKV

KNLGYEIPKDKIKELSVRYARFNSSVLGYGIGLDLVKKVCEKYKMRLEIHSEPSLKGSFY
ENSFCIHFGG

>PROKKA_00492 Bifunctional NAD(P)H-hydrate repair enzyme Nnr
MLSVYEKVNALDKRAIEELFSEDILMENAAMALERAVLQNASLGAKVILCGSGDNGGD
GYALARRLIGRFRVLVFEMKRAKSPMCQLQKERAKKAGVVIKTYEENALNQNECDVLID
CVVGSNFKGELEPFLNFESLSQKARFKIACDIPSGIDSKGRVDKGAFRADTTISMGAIKS
CLLSDKAKDYVGELKVGHLGVFNQIYEIPTDTFLEKSDLKLPLRDKKNAHKGDYGHAV
LLGKHSGAGLLSALSALSFGSGVSVQALECEITSNNKPLELVFCENFPNFLSAFALGMG
LENIPKDFNKWLELAPCVLDAGVFYHKEVLQALEKEVILTPHPKEFLSLLKLVGINISML
ELLDNKLEIARDFSQKYPKVVLKLGANTLIAHQGRTFINNLSVALAKAGSGDVLAGLI
LSLLSQNYTPLDAAINASLAHALAGLEFKNNYALTPLDLIEKIKRL

>PROKKA_00493 Replicative DNA helicase
MDHLKHLQQLQNIERIVLSGIVLANHKIEEVHSVLEPSDFYPPNGLFFEIALKLHEEDC
PIDENFIRQKMPKDKQIKEEDLVAIFAASPIDNIEAYVEEIKNASIKRKLFGGLANTIREQ
ALESQKSSDILGAVEREVYALLNGSTIEGFRSIEKVESAMDLITENQRKGSLEVTGIP
TGFVQLDNYTSGFNKGS�VIIGARPSMGKTSLMNMVLSALNDDRGVAVFSLEMSAEQLA
LRALSDLTSINMHDLESGRLLDDQWENLAKCYDHLSCCKLFFYDKSYVRIEQIRLQLRKL
KSQHKELGIAFIDYQLMSGSKATKERHEQIAEISRELKTLARELEIPIIALVQLNRSLE
NRDDKRPIILSDIKDSGGIEQDADIVLFLYRGYIYQMRAEDNKIDKLKEGKVVEEAQELHL
KVNEERRIHKQNGSIEEAEIIVAKNRNGATGTVYTRFNAPFTRYEDMPIDSHLEEGQETK
VDYDIVTT

>PROKKA_00494 Competence protein
MKDKTFQGAPELLSTPKEYLLCGVFLSLLLTNLYLEYLNQKLDPSKPTSLNAQILLQY
PKTKDQKTYFVLKLQSKGMIFYSTIKEPLKNLQYRHAQFFGKIKPCSFLESRLSCFFQTY
SFSLTRKQDFKSYWRHFIDSASSALVGNLYRALFIGDSLNDLRDRANALGINHLLAIS
GFHLGILSVSVYFLFSLYFTPLQKRYFPYRNAFYDIGVLVWVFLLGYLILLDFLPSFFRA
FLMGLLGLACFFGVRLSFKLLILACCIAIALLPKLLFSVGFLLSVCGVWYIFLFLKHT
QAFFKTSSFLARSFQAISLSALVFLNMLIVAHAFPMFSPYQLFSIPLGLIFIVFFPLSL
FLHAVGLGSLLDNVLSMPLTIPTISIPSPLWLLGAHLFLTILSARFFKVYLSMNVL SAGF
FLYCCYQYIIMPSLIVG

>PROKKA_00495 4-hydroxybenzoate octaprenyltransferase
MLKKITHKIKALSELVALEHTIFSSMFLMAMVISSYQKNQTLFFDLETLILCFLALLGA
RNFAMGFNRLVDRDIDKDNPRTKNRPSVDGRISVKGMVIFSALNALLFVGVSYFINPLAF
KLSLPFLIVLGGYSYFKRFSSLAHFIVGLALGLAPIAGSVAVLGDIPLWNVFLALGVMLW
VAGFDLLYSLQDMEFDERGLFSIPSQLGKWCNLNLSRLSHLVALICWLCFVKCYHGGLF
AYLGLGVSAIILLYEQILVARDYKNIPKAFFVSNGYLGVVFFISIVLDVGFKHA

>PROKKA_00496 hypothetical protein
MHELVLRSQALGFETRVLQCDLSFSYERFISKTRSLAVLEEFDWLNSGFDFSRLNVEND
TLELLKALYFKLEKLESLLKENLLELEQNNRIIALGHGLICLKKQSLIAPQTYYGRCVL
EGKILAFFGVARDKDFLEITRMHALDIKRYDSFIVDSERKGLKL

>PROKKA_00497 hypothetical protein
MLSSNDLFMVVLGAILLVLVCLVGYLYLKEKEFYHKMRLEKTLDESYQENYIYSKRLKE
LEGRLEGLSLEKSAKEDSSLRTTSLHLYSQLQEIQKSMDKERDYEEKIILENKFKDMG
HYAASDEINEKQVLKMYQEGYSVDSISKEFKVSKGEVEFILNMAGLKW

>PROKKA_00498 Phosphatidylserine decarboxylase proenzyme
MVALSNALSRVFGSVAGYEFPSFIQKGINALYVKIFKIDLSEFEPLENYKSLNALFTRSL
KKERPFDAKAPNICIAPCDALITKCAFLDNDSALQIKGMPYKAHELVGGEINPLSPSFFYVN
FYLSPKDYHHYHAPCDLEILEARYFAGKLLPVNKPSSLHKNKDLFVGNERVALVAKDIQGN
RLYFVAVGALNVGKMRFNFDKNIQTNAKARFTQTYSYNPPIKVKKGDNLGNFEMGSTIVL
FIQNTAFKDWKEKSVKFGESIGEFHAN

>PROKKA_00499 Quinolinate synthase A
MPTDNDLKTSIVELLCDLDALLVAHFYQKDEIVELAHYAGDSLELAKIASQSDKNLIVFC
GVHFMGESVKALAFNKQVIMPKLSCCSMARMIDSHYYDRSVHLLKECGVKEFYPTIYINS
NAEVKAKVAKDGGVCTSRNASKIFNHALKQNKKIFFLPDKCLGENLALENGLKSAILGA
NSKEEIKNADVVCYNGFCSVHQLFKLEDIEFYRQKYPDILIAVHPECEPSVVSNAFSGS
TSQIIEFVEKLSPNQKVAIGTESNLVNRLKAKRHHQNTFILSSTLALCPTMNETTLKDLF
EVLKAYKNHRAYNAIELKDEVAHLAKLALTKMMELS

>PROKKA_00500 putative nicotinate-nucleotide pyrophosphorylase [carboxylating]

MEIKTFLERALKEDLGHGDLFERVLEKDFKATAFVRAKQEGVFSGEKYALELLEMTGIEC
VQTIKDKERFKPKDILMEIKGDFSMLLKVERTLLNLLQHSSGIATLTSRFVKALNSHEVR
LLDTRKTRPLLRIFEKYSVLNGGASNHRLGLDDALMLKDLTHLKHVKDLKSFLTHARKNLP
FTAKIEIECESFEEAKNAMNAGADIVMCDNLSVLETKEIAAYRDAHYPFVLEASGNISL
ESINAYAKSGVDAISVGALIHQATFIDMHMKMA

>PROKKA_00501 N-6 DNA Methylase

MLKEYLEŠIKDLTPEKNELTHRLFLHNLDDKLKDHFNKEFKIEHEPERKQGSQPDFRVSY
QGLNIGYIENKRVGTNLRKIVESEKSDQILKYLELNPMLTDYLNFMWVGKDENNAPLI
KKEISVASLDELSSKSLKPKPKPQTECDLIELFKSFFNHEATPIANAKDFATHLSPRTKYL
KNALIKYQEKVQVSSI FNNFKEYLYEELS FEDFS DALAQTLTYSLFLAKLNHPFEKINLD
NVRSSI PKNFAVIREMADFLKKLDEIKEIQWLLNEILSSINHVDMSILKDLNDDKDPYL
HFYETFLSAYDPKLREKKGVYYPDSVVKFIINALDSLLKTHFKDAPLGLKSALDNENIK
LLDFATGTGTFLLEAFRKALET RKTSDGGISTKEEKYQNLKQFYGFYLIAPYAI AHLN
LSQAFKEEFKKPLKENDALQIILTNTLIQPS EIVAYRGLSPIFEKELSNAQEIKKDEKIL
IITGNPPYSGASSNEGLFEWEVKATYGIEPEFQTIEIEKNVKLTDKIQTLTKNIQTQKES
SGKDALKALKSLHSKYKLQNEKNPKWLLDDYVKFMRFAQNKIESLGHLGFGFISNNAFLD
NPTFRGLRCSLLECYDELYILNLHG NARKKEETPQGAKDENVFNIKQGV SINL FVKAQA
TKQKIYYDYVYGERAEKYAFLAQHDLDRIEWLELT PREPFYLLLPQKTPLLDEYEQGF SV
QDMFQISSVGIVTGKDKIFIANNTESLKEQVLKYCNEFNEQCIKDIHYRPFDIRKVYYDT
KKLERARENTFKHMLPPPPPTNPKTPNQ TCKNVALNTPRQLKNNDKSWTQCFISSNINDQ
GLSSGGNGAGVNYPLYQFRDPNYTENFTPEFRSFIDKHYSHSFEPLEVLGYIYALLYSPN
YRKRYEDFLKADYPKILFTKNKDLFRVLSLLGIELIGLHVLNKE SLNYSFEKLDKAIGES
CYKEAHDRNPIIKKPLHNEPEQR LYINHSA YFRVSKEIYDYRIGGGIAF

>PROKKA_00502 Putative agmatine deiminase

MKRMLAEFEKIQAILMAFPHEFGDWAYCIKEARESFLNIIQTIAKHAQVLVCVHTSDTIG
YEMLKNLPGEVIARIDTNDTWARDFGVISVENQGVLECLDFGFNGWGLKYPSNLDNQVNS
KLKSLGFLKHPLKTPYILEGGSIESDGAGSVLTTPNKRYRILQKSQIYPNDPQRLEILC
LQNASIPPLFA

>PROKKA_00503 Modification methylase DpnIIB

MIQIHADAFAEIIKDFHQQDLKVD AIIITDPYPYNISVKNNFSTLKS AKRQGIDFG EW DKNF
KLEWIKHYAPLVNPN GCMVIFCSYRFISYIAD FLEENG FVVKDFIQWVKNNPMPRNIHR
RYVQDTEFALWAVKKKAKWVF NKPKNKEYLRPLILKSPVVS GLEKTKHPTQKSLALMEKI
ISIHTNPNDIVLDPFMGSGTTGLACKNLERKFIGIESEKEYFQTAQKRLNLF

>PROKKA_00504 Modification methylase BspRI

MSYKILDLF CGAGGFSAGLERLEEFDALIGLDCDKQALITFENNHNKNAIGVCGDITQAAI
KEKVIKLAQTLGINMIIGGPPCQGFSNKGKNLGLKDPNRLFLEYIEIVKAIKPEIFIIE
NVKNLISCAKGYFLEEIKERLNALGYQLSYQILNAKDYGVPQNRERAFIVGASRFSFDFN
LLEPSQSVNIQDAISDLAYLCSNEGAFESDYLNPIQSSYQALMRKNSPKLYNHQATNHSQ
AALEKLKLINKEQGKECLPKNLHGKQQFKSTWGRLNWNKISPTIDTRFDTPSNGTNSHPE
LHRSITPREAARIQSFSDDYIFYGNKTSVCKQIGNAVPPLLALALGKAILKSARNDTNPS
R

>PROKKA_00505 hypothetical protein

MNFQAKVFVSCFLNAVD PPNLGVLLSRFQIKNGCIYGVYSYKASKSVHGYEESKVPVLKA
LNTLSMHKIWFNQESVTKIRGTFAFILENDLHLDENSFYKLLNSLIDNDFNRSMSMT
PNQRLFLSGFFESRGSIDTQRNFLTSDYFFHSPLEFKKFHYLIDFFNIPSEALNFNFREL
QPEYVQGINQRNAQFRIYLDWYLHHIGLFNPYKARIAHHVFKTTLVGDGIYYKLSYPPTT
EYRGNSFIERAHFYLSVYQQDLDDKSIEKLREQLGWIQKSEEFKRDSKIINFYRISTPN
VCSACYGDYDIKERSFISLPLYKITQNPDSYYTEIHHVISLGKDQELDVLANLAKLCPAC
HRALKKGSSEEEFQKRLIRNILNRNKNLEFAQLRFETDDFPTLIDRIYESLK

>PROKKA_00506 Modification methylase DpnIIB

MIEYSNNNNPRKKIYASEQVGKRVQDIWEFKDPQYPSYPT EKNAQLLDLIKTSSNKDSI
VLDCFCGSGTTLKS AFLLRKFIGIDNSDLAIQACKNKLETITKDLFISQNFYDFLVF

>PROKKA_00507 DNA methylase

MREMQNLLIQAENAIALLFLLNDKNLKGKIDLIYIDPPFATNNHFTITNGRATTISNSKN
GDIAYSDDKVVGMDFIEFLKQRLVLLKELLSEQGSYVHTDCKIGHYVKVMLDEIFGIQNF
RNEITRIKCNPKNFKRMGYGNIKDMILFYSGKNPIFNPKIPYTPQDLEKRFPKIDKDK
RRYTTVPIHAPGEVESGECSHLKVCYLQKGGIGALILPHLSVGTKKV

>PROKKA_00508 Tsp45I type II restriction enzyme

MVNYWTKLIIEYANQRSYLLDDLFQVYPTIPEGLREIDSKIWSNVEYHFKRKDNLALITEL
LNLDLFPIKDSY MAYLKRDKSALERNPRTINRICGRLYEMGLNKIFEECSEPKETNRQIG
PMFKDWNNKSLGVEPVDLNDFIANENDAILKASDNVMAEFAKSHLNYHHHKGLDFVARF
NKKYIIIGEAKFLTDFGGHQAQFNDAISTIETPNIAIKVAILDGVLYIESNNKMRKLLD
TTYRNYHIMSALVLRFLYQI

>PROKKA_00509 hypothetical protein

MLFNQTLTYISLFSGAGVGCYGLLEEGFECVATNEILDSILKPLNKN

>PROKKA_00510 Sodium/proline symporter

MGHVVLNTPIVTMFIVYSLMLLYIGFYFYKQNETTEDYFLGDRSMGPFVISALSAGASDMS
GWLLMGLPGALYVGGLINSHIAIGLSLALINWVFAKRLRIYTSVIANSITISDYFETR
FSDDKHILRLISAFVILIFFIFYISSGLVSGAKLFEATFGIQYTYALSIGTLIIVSYTFL
GGYKAVCWTDLIQGLLMSALIVVPIVMIIHLGGIGEGIKIIREIKPENLSFLQGSSVVA
IISSLAWGLGYFGQPHILVRFMSIRSIRDVPKATTIGISWMVISLIGACVMGLLGVAIVH
KFDLSLQDPEKIFIVMSQLLFNPWITGILLSAILAAMSTASSQLLVSSSTIAEDFYATI
FNKDAPQKLVVISRLSVLGVACIAFFISTDRNASILSIVSYAWAGFGASFGSVILFSLF
WSRMTRIGAIAGMLSGASTVILYDKFGKSFLDIYEIVPGFIVASVAIVVFSLSVRAGT
KEAFETMLKEIESLRH

>PROKKA_00511 1-pyrroline-5-carboxylate dehydrogenase

MQKIIDDSLELAKKLQDSISNHLSDQEKAFHFSKMQLLNPNPENKVMLIELMDRSFRCLDN
KARFEMIEHVLDKYKSREIFSPFEKLLLMGFLSFGKMLPDMSVPFFVNKIRSDTKAMVLD
QEENKLKERILKRKNEKIILNVNFIGEEVLGEEEEANARFEKYSQALKSNYIQYISIKITT
IFSQINILDFEYSKKEIVKRLDALYALALEEEKQGMFKFINLDMEEFRDLELTVESFME
SIAKFDLNAGIVLQAYIPDSYEYLKKLHAFSKERVVLKGLKPIKIRFVKGANMESEETIAS
MKDWALPTFSSKQDTSNYNKMLDFVLEGDNYKYIHIGAASHNIFEIAYVYTRIHANLDP
VVLEHFSFEMLEGMSLQASQELKEMHKLIILYAPVCDEAHFNNAIAYLVRRLDENTSSDNF
MKAFFNLKVGTSWKDQEQRFNLNLKGIAALDNATHRTQDRNAKQSGHTTYPNHSFKNES
DTDFILKANREWAKKVREKMHNAPILELYPEIDGRFEDPNLTPLEVFDKIHKKIASVHL
ADKEAILKALEVAKSDKSHFSQKSFTTEIHALLSQTQIFRERRGDLVGISALEVGKTFAE
TDAEVSEAIIDFLEFYPSYSLRVLQEQNPCKTFTPKGVGVVIAPWNFPVGVISVGTIAAPLAA
GNRVIYKPSLSSTVTGYKLCECFWDAGVPRDALIYLPKSGSDISEHLLKDESIQFAILTG
GEDTAYKMLKANPTLALSATGGKNATIVSKMADRDAQAIKNVIHSAFSNSGQKCSATSL
VLEKEVYEDENFKKTLIDATLSLSVGDPDFDKNKIGTLADKPNEKVIKAIDELKSYENYE
IPASFVNDNPYLMKPSIKYGTKKGDFTHQTELFTPILSVMEAKDLDEAIEIVNSTGYGLT
SAESLDEREWEYERIEAGNIYINKPTTGAIVLRQPFGGVKKSAVGFGGRKVGIFNYIT
QFVNIHQEEEDENALKNPLSETLEGLTQKGYDEHTELKHAIFMAKSYAYHYKHEFSQAK
DYVKIRGEDNLFSYTKVKSVMGYRITEKDTLSMDLGVLAACLISQIPLTSLIENERANKDL
TFFLECLKALQANAPIVYESLQKFSEKLNFTNVRVYLKSDLDLLHEQASALGMVLATAKP
CLNGRFELLYYHLERSVSISYHRYGNLGSRVLRQPTCHKSCCAEK

>PROKKA_00512 hypothetical protein

MKTIKNGIMIGTLGALLLSGCSSFDTQRFACLPKDHSLKDASTKKEAQYTPKGFDPYSS
NLNHWDSTF

>PROKKA_00513 hypothetical protein

MKELEQSQQVLKKGNDLLENGENQKLADKNKALTTEKDNLTDKENLTTELSTAKSQAECT
SQKLNELERRHAPYQKLEKLYEVFLEVKDRNLNFGFVATTHSAMDLIASVLSDSKYYLES
YNKASQELSDKRNDKGEKLAELFDLLFEYIKDSKFERLKEPSAYDHSCKKLYPEQNTSQK
MQRVVLIGYTYDKKHHHIIILSWIWDHKMGTFIEKCFGFSQIKEELEARIGELEDENAELE
NRNDKLFAEKENLATKISGLENSNDQL

>PROKKA_00514 TPR/MLP1/MLP2-like protein

MKTENDKLNQVIVLNKEHGSLKQERVRLQGEHGFLEKSCANLEKENQHLTDKLLKLESA
QKSLENTNAQLRQSLSENSNAQLVQAKEKIAEEKTELEREIA RLKSLESMEAKSDLDLHNR
RLASANQDLKRQNRKLEENIALKERVNGLNEQLTLQPQKPQ

>PROKKA_00515 hypothetical protein

MVNPLQSLRLPLGHPLVEKLCDFRLKDGVKFNEKIELNFKKEVSEEEKIKFKQALRALHA
IVNNEVSSRYLSDDNQKFIENTLAQAEEKITNELVEKALEIVSYSDVDVDFEAFKEMMLNVD
NIAVGLKSYSQSQLLDLNGGHWDLEVPSLSKESVTFRFDNLDSNGKEENFYARSSLKDLN
KQGVVAIDFGTKSTTAAYMDENGKYRLLSIGGDVDIESLEKYENPTIVEFRDKEKFLKAY
NALNHRPFTEKNDIQVAHEAQKGLSSAQGNDLYRFFSQLKQWAGADEKRNFRDFKEDFSL
ESFTNCTDFNPIEIIYAYYIGRCINMQNGVFLKYFLSYPIKYEKHQAEKIRESFEEKGLKK

SLPRHVFDDDEKTAKNFKVELRASEPSAYAISALKSYGFDKFAKLDKPIYYGVDFDGGGTT
DFDFGKWEKSANPKFAYKMTFSSGGDKYLGGENLLELLAFEAYGQNFQTLKEKDIVIAK
PNYDRIDTQRFSGFMQNSNTARLNLQMIASGLRPFLENLDANIIEAIEENEKFEIENFEK
DFKTMFLDRNGVETECDLKVDCKELLNLLKDKINESVANFFAGFSKMAENIDDECRAFH
IFLGGNASKSVLVKQAFENAKEEQLKAYQQKTSKDDFTFILYEPLGTEASDKQILELTGE
DVSKIPPYLLKPTCKTGVAFGLLSRDRAGIERPSIDSNPVFKYDLGIEREGKFHTRISR
DSLKPNEYQIFQTKEEWGGFDGLEIRYSDKPLANTNTLSIEDTQLVFIALEEHEEVDVKV
CCVDSQSIKVGLFKDQGQLIYESEAERL

>PROKKA_00516 hypothetical protein

MMTKNAYAFVVEESVTVFKHTKDEGLMPIFEGFVPLKEVFLKSFKERCNVEFLENLDLL
FLYDYPFSSEVFSCLKDLKNSIWDRKLVVALVEALEGFKDLNLSLQIEYKRSNSLGNGIQ
KLLTNADLGSNHKAIVIDSMKTYHQSQQEKYKRERGETLEVRPTTPPSYGGGSIRISDDK
KPSDSNEENF

>PROKKA_00517 hypothetical protein

MSKVQMDTEEVGRGVGHLERFKELLNEEVNSLSNHFNLDWSQDARRDKFSEVVDNLRST
FNEFDEAAQEQAIAWLKERIRVLEEDY

>PROKKA_00518 hypothetical protein

MAEWKTDTEEVKEIVRKREFKESLQEKRCLGFIKDLDSYALKIIVERRKIEMQLQKAIE
KLKKAIKERSGFWGFLGELGRDFGNAVGSVIPPAKLGAEECEKGLNLMEDNIEKWEKNVR
LLEQALEIYISIQTKASVELVEGAWESVKKSLNFYTDKHQEFIKRLNQASDAIDNEYNFPT
PGVLMEYDFERPPIVYSPKKSVDERLKDLDREDFASLYADWKDKINAFSHRDRAKTSKE
RELENLEDLMGASSYDENPNDELDRMASSKEQEFESLLEDLMPSVLGVPSYDESLTLAKN
RCVKNCCKALQDFAEKIKEAPNDSSAINEAFNHLETELERATESLNQKIDPILERNENYT
QKALEYREFLESRKEDFIIDEQNYPYPDEVRFNALRLAEFDSVFSIAIAPLEDLDKTVCAHH
ALKALQSALKDNDLDFDAAELEQIAKGFIIPRGYLWHFDANVLGNVALVREELLLGVKHTK
GYSLWTEFLQKQN

>PROKKA_00519 SMI1 / KNR4 family protein

MDRISTETELNWEFVEPLNENALNGLEERLKIGLSDAFKDFIKRSNYGFSQWRYFMVGNK
SYTFKHVLNFNLESFIDFMQSLKEWLEPEEIVFANDGYGGYYLWNTTTDVLFLDTDNG
SKKALLNLKMFLKKLESRG

>PROKKA_00520 hypothetical protein

MFSHKVFLEGCTNELRRICDYFVEEAMQDDLQKQKSEALEEMLKIVHDLENLEQDTQYE
MRKIDEQLEEAQRLEKQIHMNFHSPSEIDRLMREAKEHEREANKRYDEYLKSKDKE

>PROKKA_00521 DNA translocase SftA

MIDVNGLLKELDALDKVVPKKEPESFLKPIVLEIEEYQKNVRQIQVQFTDAPKFNEEGA
YPQFLSCGLLEIKGKNGANMEFLLPKVYFPFPKSLYIEHEKDGQFLREMLMRLLSAPLV
QLEVILVDALSLGGIFNLARLLNKNDFIYQQRILTESKEIEEALKHLYEYLKVNLOEK
LAGFRDFAHYNENEKDPLPLKALFLSGVDALSQNALYYLEKIMRFGSKNGVLSFVNLESE
KNNQSAEDLKRYAEFFKDRTSFECFKYLNVEVINDQGIKSQHMKDFADKIKAYYKQKKEV
KRELKDLQORDEKFWTESSQFKVSVPVGWDINHKKVRFEIGEVDQNHTLICGRSGSGKSNFL
HVLIQNLAFFYAPNEVQLFLLDYKEGEVFNAYTDPNILEHARLVSVASSVGYGMSFSLSWL
CKEMQKRAELFKQFNVKDLSDYRKHGEMPRLIVVIDEFQVLFSDNSSKGKESVEQSLNTL
LKKGRSYGVHLILATQTMRGTDINRSLMAQIANRIALPMDAEDSNSILNNDAAACELVRP
EGIFNNNGGHQKYHTKMSIPKAPDDFKPFIKKIHRDFNQRLAPIEHKIYNGEKALEMPN
TLKANEMRLHLGKEADYEQKDLIVEFENNESYLLVVSQDLNARIALMKLFAQNFKTANKE
LLFYNAEKRLVRELDELKKTTPHHAHARPSRERFRHCHP

>PROKKA_00522 hypothetical protein

MQGPLGSVLDTAMNPNSVLVIDNLNEAKELHDKIGVEKLRSFLEKATDNEQYCIIFAHD
KQIQANYDSVKLKELLNNHFKQRLAFRCNEENLNAIKNNLPPPSALNNLNALFVELSKDS
YTEFRPFSL

>PROKKA_00523 Urease accessory protein UreH

MNTYAQESKLRLKTKIGADGRCVIEDNFFTPPFKLMAPFYPKDDLAEIMLLAVSPGLMKG
DAQDVQLNIGPNCKLRITSQSFEKIHNTEDGFASRDMHIVVGENAFLDFAPFPLIPFENA
HFKGNTTISLHSSSQLLYSEIIVAGRVARNELFKFNRLHTKISILQDEKPIYYDNTILD
KTTDLNNMCMFDGYTHYLNVLVNCPIELSGARGLIEESEGVDGAVSEIASSHLCLKALA
KGSEPLLHLREKIA RLTTQTITQKI

>PROKKA_00524 Urease accessory protein UreG

MVKIGVCGPVGSGKTALIEALTRHMSKDYDMAVITNDIYTKEDAEFMCKNSVMPRERIIG

VETGGCPHTAIREDA SMNLEAVEEMHGRFPNLELLLLIESGGDNLSATFNPELADFTIFVI
DVAEGDKIPRKGPGITRSDLLVINKIDLAPYVGADLKVMERDSKKMRGEKPFIFTNIRA
KEGLDDVIAWIKRNALLED

>PROKKA_00525 Urease accessory protein UreF

MKSTESVGMPPKTPKTDNNAHVDNEFLILQVNDVFPVIGSYTHSFGLETYIQKKVSNK
ESALEYLNKANLSSQFLYTEMLSLKLTYES TLQQDLEKILGIEEII RLSTSPMELRLANQK
LGNRFIKTLQAMNELDMGEFFNAYAQQTKDP THATSYGVFAASLDMELKKALRHYLYAQT
SNMVINC VKSVPLSQNDGQKILLSLQSPFNQLIEKTLELDESHLCAASVQNDIKAMQHE
LSRLYMS

>PROKKA_00526 Urease accessory protein UreE

MIIERLIGNLRDLNPLDFSVDHVDLEWFFETRKKIARFKTRQGKDIAIRLKDAPKLGLSQG
DILFKEEKEIIAVNILDSEVIHIQAKSVAEVAKICYEIGNRHAALYYGESQFEFKTPFEK
PTLALLEKLG VQNRVLSSKLD SKERLTVSMPHSEPNFKVSLASDFKVVMK

>PROKKA_00527 Acid-activated urea channel

MLGLVLLYVGIVLISNGICGLTKVDPKSTAVMNFVVGGLSVVCNVVVITYSALHPTAPVE
GAEDIAQVSHHLTSFYGPATGLLFGFTYLYAAINHTFGLDWRPYSWYSLFVAINTVPAAI
LSHYSDMLDDHKVLGITEGDWWAIIWLAWGVLWLTAFIENILKIPLGKFTPWLAIIEGIL
TAWIPAWLLFIQHWV

>PROKKA_00528 Urease subunit beta

MKKISRKEYVSMYGP TTGDKVRLGDTDLIAEVEHDYTIYGEELKFGGGKTLREGMSQSNN
PSKEELDLIITNALIVDYTG IYKADIGIKDGKIAGIGKGGNKDMQDGVKNLSVGPATEA
LAGEGLIVTAGGIDTHIHFIS PQQIPTAFASGVTTMIGGGTGPADGTNATTITPGRNLK
FMLRAAEEYS MNFGFLAKGNVSN DASLADQIEAGAIGFKIHEDWGTTPSAINHALDVADK
YDVQVAIHTDTLNEAGCVEDTMAAIAGR TMHTFHTEGAGGGHAPDI IKVAGEHNILPAST
NPTIPFTVNTEAEHMDMLMVCHHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSITS
SDSQAMGRVGEVITRTWQTADKNKKEFGRLKEEKGDNDNFRIKRYLSKYTINPAIAHG
EYVGSVEVGKVADLVLWSPAFFGVKPNMI IKGGFIALSQMGDANASIP TPQP VYYREMFA
HHGKAKYDANITFVSQAAYDKGIKEELGLERQVLPVKNCRNITTKDMQFNDTTAHIEVNS
ETYHVFVDGKEVTSKPANKVSLAQLFSIF

>PROKKA_00529 Urease subunit alpha

MKLTPKELDKMLHYAGELAKKRKEKGIKLNVEAVALISAHIMEEARAGKKTAAELMQE
GRTLLKPDDVMDGVASMIHEVGIEAMFPDGTKLVTVHTPIEANGKLVPGELFLKNEDITI
NEGKKA VSVKVNVD RPVQIGSHFHFFEVNRC LDFDREKTFGKRLDIASGTAVRFEPGE
EKSVELIDIGGNRRIFGFNALVDRQADNESKKIALHRAKERGFHGAKSDDNYVKTIKE

>PROKKA_00531 Lipoprotein signal peptidase

MLKTTKKSLLVFIFGVFFLIFGVDQAIKYAILEGFRYESLVIDIVLVFNKGVAFSLLSFLE
GGLKYLQILLILGLFIFLMRQKELFKNHAIEFGMVFGAGVSNVLD RFRHGGVVDYVYYHY
GFDFAIFNFADVMIDVG VGVLLLRQFFFKQKQNKIKA

>PROKKA_00532 Phosphoglucosamine mutase

MKIFGTDGVRGKAGVKLT PMFVMRLGIAAGLYFKKHSQTNKILIGKDTRKSGYMVENALV
SALT SIGYNVIQIGPMPTPAIAFLTEDMRCDAGIMISASHNPFEDNGIKFFNSYGYKLKE
EEKAIEE I FHD EELLHSSYKVGESVGS AKRIDDVIGRYIAHLKHSFPKHLNLQSLRIVL
DTANGAAYKVAPVVFSELGADVLVINDEPNGCNINDQCGALHPNQLSQEVKKYRADLGFA
FDGDADRLVVVDNLGNIVHGDKLLGVLGVYQKSKNALSSQAIVATNM SNLALKEYLKSQD
LELKHCAIGDKFVSECMQLNKANFGGEQSGHII FSDYAKTGDGLVCALQVSALVLESKLV
SSVALNPFELYPQSLVNLNVQKKPPLESLKGYSALLKELDQLEIRHLIRYSGTENKLRL
LEAKDEKLLESKMQLKEFFEGHLC

>PROKKA_00533 30S ribosomal protein S20

MANHKS AEKRIRQTIKRTERNRFYKTKIKNI KAVREAVAVNDVAKAQERLKIANKELHK
FVSKGILKKNTASRKVSRLNASVKKIALA

>PROKKA_00534 Peptide chain release factor 1

MSILA EKLSSILKRYDEL TALLSSAEVISDIKKLTELSKEQSSIEEISIASKEYLSVLEG
IKENKELLEDKELSELAKEELKILEIQKSDLETAIKQLLIPKDPND DKNIYLELRAGTGG
DEAGIFVGD LFKAYCRYADLKKWKVEIVSSSENSVGGYKEIIALIKGKGVYSRLKFEAGT
HRVQRPETESQGRIHTSAITVAIMPEVDDVEVSINPSDLKIEVF FRAGGHGGQC VNTTDS
AVRITHLPTNISVSMQDEKSQHKNKDKALKILKARLYEKQIEEQQLANAKDRKEQVGS
RSEIRTYNYPQNRLSEHRINLTLYSLEEIMLSGSLDEVINPLIAHAQSQFE

>PROKKA_00535 Helicobacter outer membrane protein

MKKVFLGMALAFSVSMAEKSGAFLGGGFQYSNLENQNTTRTPGANNNTPIDTSMFGSNQT
APVAQAQATYTPSVINTNNYQMYGVDAMAGYKWWFFGKTKRFGFRSYGYYSYNHANLSFV
GSQLGIMEGASQVNNFTYGVGFDALYNFYESKEGYNTAGLFVGFGLGGDSFIVQGESYLK
SQMQICNNTAGCSASMNTSYFQMPVEFGFRSNFSKHSGLIEVGFKLPLFTNQFYKERGVGDG
SVDVIFYKRNFYSIFYNYMINF

>PROKKA_00536 hypothetical protein

MKAIKILFIMTSLSLNAISVNRALFDLKDSQLKGELTPKIVDFGGYKSSTKEWGASALNYI
NAANGDAKKFSTLVEKMRFNNGILGNFRAHARLKQALKLQKNLKYCLKIIARDSFYSYRT
GIYIPLGISLKDQKTAQKMLADLSVVGAYLKKQQENEKAQSPYYRSNNYNSYYSPPYGM
YGMGMYGMYGMYGMYDFYDFYDGMYGFPNMMFMMQVQDYLMLENYMYALDQEEILDHDAS
TDQLDTPTDDDRDDKDDKSLQQANLMSFYRDPKFSKGIQTNRLNSALVNLNDSRMLKDNS
LFHTKAMPTKSVDAITSQAKELNHLVGQIKEMKQDGASPNKIDSMVKNKAMEVRDKLDNNL
NQLDNDLKDQKGLSSEQQQAQVDKALDSVQQLSHSSDVVGNYLDGSLKIDGDDRDDLNDAM
NNPMQQPVQQTPTSNNMANHANDSKDQGGNALINPNSTTNTDDTHTDDTHTDTNTTNDAS
TTDTPTDDKDASGLNNTGDMNNTDTGNTDTGNTDTGNTDDMSNMNNGNDDAGNANDDMSN
GNDMGDDMNANDMNDMDMGNGNDDMGDMGDMNDMDMGDMGDMGDMGDMGN

>PROKKA_00537 Methyl-accepting chemotaxis protein CtpL

MKSTRIGSKIVMMVCAVVIVISAVMGVIINHKKVESVLQHQAPELLQKKAQLVSFKIQGIM
KRIFIGANTLEKFLSDENSAINDTLKRRMLSEFLANPHVLLVSAIYTNNNERVITAMNM
DSKIAYPNTTLNENMANQIRSLKDITHSDPYKEVNGDKIYGMDITLPLIGKNQNIIGAL
NFFLNIDAFYTDVIGQKKSNTFLMGKNGRLLINPNREIQDKILSAINPDKRVAKAVEYYN
NNEAGTLSYHSLSGDTETFLAIQPFDFFEKGKENGNGHWRWAIGKYVNKSLVFKEATNTKF
IIITTLILGLVLVLAFLVFIIISNLITKRISRVNNTLNDFNLLNPNKNNHAI SLTPPSAY
DEIGQMQASINENILKTQESIQADNKAIQNSIEVTNYVENGDFTEIACVPKNKDLQALR
NTINSIIQYFRNQIGANIESLNNALEHYKNLDFTHHIQDPKANMEKALNTLGQEISSMLK
ASLGFANALNHESKDLKTCVDNLTKTAHKQERSLKNTTQSLEEITNIIITIDSKSQEMIS
QGEDIKSVVDMIRDIADQTNLLALNAAIEAARAGEHGRGFVAVVADEVKLAERTQKSLSE
IEANINILVQSIADNAESIKMQNKGVENIHNSINALQQDVQDNLTIANYSLQVSTKIDGI
SQDILEDVSKKKF

>PROKKA_00538 30S ribosomal protein S9

MRKIYATGKRKTAIAKVWLTPGKGGLSINEQSLNQWLGGHEAIAKMKVMQPLLLTKQEQSV
DIKAVVFGGGYSAQAEALRHGISKALNAYDIAFRAILKPKGLLTRDSRVVERKKYGRKA
RRSPQFSKR

>PROKKA_00539 50S ribosomal protein L13

MTKTAKVNDIVRDWVVLDAKDKVFGRLITEIAVLLRGKHRPFYTPNVDCGDFVVVINANK
VKFSGMKLEDKEYFTHSGYFGSTKSKTLQEMLEKAPEKLYHLAVRGMLPKTKLGKAMIKK
LKVYRDDKHPHTAQTSKKDAK

>PROKKA_00540 hypothetical protein

MQKEQETREIAKKAVKIVFFLGLVVLMMINLYMLINQINASAQMSQQIKKIEERLNQE
RK

>PROKKA_00541 Malate:quinone oxidoreductase

MSMEFDAVIGGGVSGCATFYTLSEYSSLKRVAIVEKCSKLAQISSAKANSQTIHDGSI
ETNYTPEKAKKVRLSAYKTRQYALNKLQNEVIFETQKMAIGVGDEECEFMKKRYESFKE
IFVGLEEFDKQKIKELEPNVILGANGIDRHENIIGHGYRKDWSTMNFAKLSENFVEEALK
LKPNNQVFLNFKVKKIEKRNDTYAVISEDAAEEVYAKFVLVNAGSYALPLAQSMGYGLDLG
CLPVAGSFYFVPDLLRGKVYTVQNPKLFFAAVHGDPDAVIKKGKTRIGPTALTMPKLERNK
CWLKGISLELLKMDLNKDVFKIAFDLMSDKEIRNYVFKNMVFELPVIGKRKFLKDAQKII
PSLSLEDLEYAHGFGEVRPQVLDRTKRKLELGEKKICTHKGITFNMTPSPGATSCLQNAL
VDSQEIAAYLGESFELERFYKDLSPEELEN

>PROKKA_00542 hypothetical protein

MRYFLVIFLFLFVGCTKKDFTLKDLSLPQEASSYLASSQNGSHNNQSIDPQALRENKES
YLKAWYSPWLDKMKVKSNNKEVFWILKEMNKSTGYGEDLKPNKAFNDALIKSMDIERYP
AKIKAVVVRDSVRAVPTNKPFFYLSPKGYPFDRYQNSLIFQGTPLITHFNTDKTYAHIQ
SSFVYGWIKVSDLAYMHDKDIELLTHLKDYVMPKDKIPLYTDYGDFTYNARVGELFALI
PQSQKTPQKPQKKELKAYGFLRDAKGYATLQSVILEEKDFVFPKAFNSENMAFYDITML
GQKYGWGGLGNRDCSAFTRDSFANFGILLPRNSYAQSRYANNYVDLSFMKAKEKEDYIL
KNATPFGTLIYLGKHHIMLYLGAHNHQAIVAHSIWSVQTQKHFKTL SHKIGGVVITSLWLA
EEHNGAFSKKKLLIDRVLGMSDLKDFVNKTSSPLNAN

>PROKKA_00543 RNA polymerase sigma factor RpoD
MKKKANEKKAQKRAKQEAKAEATQENKQENKAKESNKAKEGKIKENKTESKVKETAKEL
VPVKKLSFNEALEELFANSLSDCVSYESI IQISAKVPTLAQVKKIKELCQKYQKKLVSSS
EYAKKLNAIDKIKKTEEKQKVLNEELEDGYDFLKEKDFLEWSRSDSPVRMYLREMGDIKL
LSKDEEIELSKQIRLGEDIILDAICSVPYLIDFIYAYKDALINRERRVKELFRSFDDDDDE
NSVSDSKKDDDDSEDEENEERKKVVSEKDKRVEKQVESFKALDKAKKEWLKALEAPIDE
KEDELVRLLTLAYKRQTLKDRLYDLEPTSKLINELVKTMETTLKSGDGFEKELKRLEYKL
PLFNDTLIANHHKILANITNMTKEDIIAQVPEATMVSVMYMDLKKLFLTKEASEEGFDLAP
NKLKEILEQIKRGKLISDRANKMAKSNLRLVVSIAKRFTSRGLPFLDLIQEGNIGLMKA
VDKFEHEKGFKFSTYATWWIKQAISRAIADQARTIRIPIHMIDTINRINKVMRKHIQENG
KEPDLEVVAEEVGLSLDKVKNVIKVTKEPISLETVPVGNDDDGKFGDFVEDKNIVSSIDHI
MREDLKAQIESVLDQLNEREKAVIRMRFGLLDDESDRTLEEIGKELNVTRERVRQIESSA
IKKLRSPOQYGRILRNYLRI

>PROKKA_00544 Aminodeoxyfutasolase nucleosidase
MVQKIGILGAMREEITPILELFGVDFEEIPLGGNVFHKGVYHNKEIIVAYSIGKGVHSTL
TTTSMILAFGVQKVLFSVAGSLVKDLKINDLLVATQLVQHDVDSLAFDHPLGFIPESAI
FIETSESLNALAKKVANEQHIALKEGVASGDQFVHSKERKEFLVSEFKASAVEMEGASV
AFVCQKFGVPCCVLRSISDNDADDEANMSFADFLEKSAQTSKFLKSMVDKL

>PROKKA_00545 Malonyl CoA-acyl carrier protein transacylase
MQYALLFPQGSGSCVGMGKSFYESHTLAKELFERASNALKVDMKKTLEENELLKESAYT
QPAIYLVSYIAYQLLNKQANGGLKPVFALGHSLGEVSAVSLSGALDFEKALKLTHQRGKM
MQEACANKDASMMVVLGVSEESLLSLCQRTKNVWCANFNNGGMQVVLAVGKDDLKALEPTL
KEMGAKRVVFLEMSVASHCPFLEPMIFKFQELLEKSLKDKFHFEIISNATNEAYHNKSKA
VELLSLQLTQPVRYQDCVKSNNDRVDVFFELGCGSVLKGLNKRSLSNKPTISVGDNKGLDE
AIEFLEEYV

>PROKKA_00547 hypothetical protein
MTYRELGGKVVLEQAEPLSVKEIFEKACEMGLDKECKDGKILPHSLGSQLGEHDISNDKK
QFYVTNNKKRPFYRWLKSREREFPPQKTPDSKEEDDGQSECSDTAKKQKNSFDERDLHLK
LVKFLDEDPNLKLLCKTIRHEEKRKSEGGECKWNPDIWGVYFPYNKYFPYNKYKEETLK
FLHHTGQKRHLKLSFELKKELSLSNLKESYFQAVSNSTWANEGYLVVFGIKDKDKDKDKD
KVLDELKRLNQSFQIGVILESEISNSKILLPAKEREIDIPTLNMLIEQSPEDFKPFMEN
INKQIEAGFDKQIDMNDFFKELLGDEAIQKHIEKHIEDKGIAEKNESALSFSFPHF

>PROKKA_00548 Modification methylase DpnIIB
MKAPCLSRPIFKSTDELFTLYQGDCNEVLPPQFENAFDLIFADPPYFLSNDGLSIQSGKIV
SVNKGWDWKEDGVNGIDEFNYQWINNAKALKNTGSLISGTYHNI FSLGRLLQKLDKFI
LNLITWQKTNPFPNFCRYLTHSAEQIIWARKSHKHKHVFNYEVLKKINNDKQMRDVWSF
PAIAPWEKVNKGKHTQKPLALLVRLLLMASDENSLIGDPFSGSSTTGIAANLLKRQFIGI
EKESEFIKISINRKLELDARFKEIRSKIKDLNFQ

>PROKKA_00549 Glycosyl transferase family 11
MILAANKSVFVHIRRGDYVGIGCQLGIDYQKKALEYMAKRVPNMELFVFCEDLEFTQNL
LGYPFMDMTTRNKKEEAYWDMLLMQSCQHGIANSTYSWWAAYLINNPEKIIIGPKHWLF
GHENILCKEWWKIESHFVKSQKYN

>PROKKA_00550 hypothetical protein
MAFKVVQICGGLGNQMFQYAFKSLQKHSNTPVLLDITSFDWSDRKMQLFLPIDLPYAS
EKEIAIAKMQHLPLKLRDALKYIGFDRVSQEI VFEYEPKLLKPSRLTYFFGYFQDPYF
AISPLIKQTFXLPPPPKMEI IKKKRKNNTASLL

>PROKKA_00551 hypothetical protein
MPKPKKNTLPCLSLSVKMSYFMRLLIKWRTRSLSHKMMTLIQILSILALASKASEDLEEQL
KKIKAYIHRITLNAKIASDVYNRVLILVNEYCTNEELFDKESVKISDILLIQDIQLYALVDE
MLKEDKYQVQHTILKGIIKRKYDEAYSLNSEDRLLEYQERLLEPSCASFSNKKFK

>PROKKA_00552 Putative 2-hydroxyacid dehydrogenase
MKTFFKGGVLDKASVGLKALEVLKEAADFDYFYEITSPNQVVERCKDAEIMVLNKVVITKE
VLSQLPKLKLICITATGTDNVDIKSAKDLGIEVKNVSTYESTESVAQHTLACALSLLGRIN
DYDRYCKSGEYSQSDIFTHISGMKGLIKGGQWGVIGLGAIGKRVAKLAQAFGAKVVYYS
PKDKKEEYERLSLEELLKTSIGIISIHAFLNSTRDLIALKELQSLKEGAILINVGRGGIV
NEKDLALTLETKDLYASDVVFVEKPFKDHAFNPKIQNKLLLTPIAWAYSDSLKTILIE
KTENIQDFLASQK

>PROKKA_00553 hypothetical protein

MKKIVLVAIALLLISGCASYKITPEHVTSYNNGIQVMTSTQAKSKVQLEIAQSKLKGLNES
PLVLVYAAQVIEGNPVVFSRK AISVSINQTNLPVLSLRQVMKSSFD FEGILQSFNISVPT
APIDNINMITPPMFYYGQGSFLAYNGMMYGGMGMYGPGFGMMMMDDVEEQEIMQESRQAL
KILAINYLKNNTLNVEGKARGGFVVVDTKNLKTPGVVVVKVFEDEIHTFKIDISKM

>PROKKA_00554 Threonine synthase

MPFVPTRSLKEKKIDFIEAILNPNA PKGGLYTLERFETLEWQDCLSMSYSELVEHVFELL
NLEIPKNLLASALKRYENFDNPKNPAPIFALNERL FVQELYHGPSLAFKDMALQPLASLF
SNLAVGKNEKYLMLVSTSGDTG PATLES LAGIPNVFVCLYPKDGTSLVQKLQMTQSAS
NLKVFGISGDFDDAQNALKNLLKDDDFNEALKARQLKLSVANSVNFGR IAFQIVYHIWGF
LELYKKGAINSKEKITLAIPSGNFGNALGAFYAKKMGLNIAKIKVVTNSNDVLREFIETG
RYDLTHRSLKQTYSPAMDILKSSNVERVLFSLFGFERTLELMQALEEEKFYALKPKELAL
LQEHFSCASCSD EACLKTIQEVYAEHQYLIDPHTATALNASLKTNEKTLVSATASYEKFP
KTTLLALNEQKKNDNDKAALET LKNSYNTPD SQRLDDL FERG IKHQEVLKLNEIKSSILL
WLESLH

>PROKKA_00555 Methyl-accepting chemotaxis protein PctB

MSKGLSIGNKIILCVALIVIVCVSILGVSLSNRVKEILKESALHSMQDSLHFKVKEVQGV
LENTYTSMGIVKEMLPKDKTREIKIHL LKNFILANSHVAGVSMFFKDREDLRLTLLRDND
TIKLMENPSLGNPLAQKAMKNKEISKSLPYRKMPNGAEVYGVDILLPLNGNTQEVVG
ALMVFISIDSFSNEITKNRSDLFLIGVKGVLLSANKSLQDKSITEIYKSVPKATNEVMA
ILENGSKATLEYLDPFSHKENFLAVETFKMLGKTESKDNLNWMIALIIEKDKVYEQVGSV
RFVVVVASVIMVLALIIAITLLMRAIVSNRLEVVSSTLSHFFKLLNNQAHSSDIK LIEAK
SNDELGRMQTAINKNILQTQKIMQEDRQAVQDTIKVVSDVKAGNFAVRINAE PASPDLKE
LRDALNGIMDYLQESVGTHMPSIFKIFESYSGLD FRDRIQNASGRVELVTNALGQEIQKM
LETSSNFAKDLANDSANLKECVQNL EKASNSQHKSLMETSKTIENITTSIQGVSSQSEAM
IEQGQDIKSIVEIIRDIADQTNLLALNAAIEAARAGEHGRGF AVVADEV RKLAERTQKSL
SEIEANINILVQSISDTSSEIKSQVKEVEEINASIEALRSVIEGNLKIASDSLEISQEID
KVSNDILEDVNKKQF

>PROKKA_00556 hypothetical protein

MLIHICCSVDNLYFLKKAKEAFAGEKIVGFFYNPNIH PYSEYLLRLEDVKRTC EMLGIEL
LEGDYELEKFLDKAKGKELLGEKSERCFECF DLRLLEASALKAFELGEEKFTTTLLTSPKK
DPNQLIAKGQHIAQRHNLEFVVFRNDNFEHFKSELDLNLQALARENELYRQNYCGCQFAL
KIQKESQNRSPFELYSPLKRQILPASIEERTQVFR TLDMAKKDANKPFLAQKTIATYRLL
NGGVWLSKNSNPLNCCILARSKSKAKVRINDLRWVFSQRLSALVGYSQRDETFLFTLEGL
NTLMAKNYDNLKELNLNPLNYEEELSLRALVSGSESINPIIVLEERIEKTLFVEIKSIFQ
EEKVFYLL

>PROKKA_00557 Helicobacter outer membrane protein

MKTLFSVYLFLSLNPLFLEAKEITWSQFLENFKNKNEDDKPKPLTLEKNNEKQKILDKNQ
QILKRALEKSLKFFFI FGYNYSQAAYSTTNQNLTLTANSIGFNTATGLEHFLRNHPKVGF
RIFSVYNYFHSVSLSQPQILMVQNYGGALDFSWIFVDKNIYRFRSYLGIALEQGVLLVDT
IKPGAITTIIPRTKKTFFQAPLRF GFIVDFIGYLSLQLGIEMPLVRNVFYTYNNHQERFK
PRFNANLSLIVSF

>PROKKA_00558 PGL/p-HBAD biosynthesis glycosyltransferase/MT3031

MLKVSVITACFNSEKTIEDTILSVLHQTYKNIEYIIIDGASADGTLGIIQKYNRIACVV
SEKDEGIYDAMNKG IKRSSGDIIALLNSDDFYKDEFVVEKVVHEFENKNCDSVYADLVFV
KPDCCLEKVVRYYESGEFSPKTL LYGVPAPHTL FVKKAIYERYGLYKTDYKISADFEMII
RLFVVQKISFSYLKEVLVIMRTGGVSASGFKSLLL RNKENIRACQENGIQANVFSMLLKY
PRKIMGLFKRESKGGLKRND

>PROKKA_00559 Methyl-accepting chemotaxis protein 4

MFSSMFASLGTRIMLVLAALLGLGGLF IG FVKVMQKDVLAQLMEHLETGQYKKREKTLA
YMTKIIIEQGIHEYYKDFDNATARKMALDYFKRINDDKGMIYMVVVDKNGVVL FDPVNPKT
VGQSGLDAQSVDG VYYVRGYLEAAKKGGDYTYKMPKYDGGVPEKKFAYSHYDEV SQMVI
ATTSYYTDIDTENKAIKEGVNKVFDENTTKLFLWILTATIALVVLTLIYAKLRIVKRIDE
LVLKINAFSRGDKDLRAKIDVGDRNDEISQVGRGVNLFVENARLIMEEIKGISTSNKTS
DKLVQIAQETQKSMKNFSTTLNSVKNKATDVASMMNASIEQSQGLRKRLIETQGLVKESK
DAIGDLFSQITESAHTEEELSSKVEQLSRNADDVKSILDIINDIADQTNLLALNAAIEAA
RAGEHGRGF AVVADEV RNLAGRTQKSLAEINSTIMVIVQEINAVSSQMNLNSQKMERLSD
MSKSVQETYEKMSSNLSSVSDSNQTMDDYAKSGHQIEVMVSDFAEVEKVASKTLADSSD
ILNIATHVSGTTMNLDKQVNLFKT

>PROKKA_00560 Trifunctional nucleotide phosphoesterase protein YfkN precursor
MRKLILVIFLTLTSLISAKEVKIVFLETSDIHGRLFSYDYAVGEQKPNNGLTRIATLIKK
QRAENKNVVLIDSGDLLQGNSTELFNDEPVHPLVLAENDLKFDIRVLGNHEFNFSKDFLE
KNIKWFNGGVNANIIKTADNKPFPVKPYIIKKIDGVRVAVVGYVVAHIPTWEAATPEHFA
GLKFLDAEEALKKTLELKGKYDILIGTFHLGREDEKGGDGIPDLAKKFPQFDIIFAGHE
HAVYNTKIGKVHTIEPGAYGAYLAKGVVIFDTKTKKKIVTTENLPTKGVPEDEELAKKYE
YVDKKSKEYANEVVGVEVTKTFIDRPDFITGGEKITTMPTAALQETPVIELINKVQKYAK
ADVSAALFNFGANLKKGPFKRKDVAYIYKFANTLIGVNITGENLLKYMESYRFYNQLQ
PGDLTISFNENIRGYNFDMFSGVKYQVDVTKPAGQRIINPTINNKPIDPKAIYKLAINNY
RFGTLSTTLKLVT DANRYNSYDALQDNGQIRDLI IKYITEEKGGKVTPELEGNWEIIHY
DFKNPLLEKLREKLKEGSIKIPTS KDGRTLN VKSIKES EVK
>PROKKA_00561 S-ribosylhomocysteine lyase
MKTPKMNVESFNL DHTKV KAPYVRVADRKKGVNGDLIVKYDVRFKQPNQDHMDMPSLHSL
EHLVAEII RNHANYVVDWSPMGCQTGFYLTVLNHDNYTEVLEVLEKTMQDVLKATEVPAS
NEKQCGWAANHTLEGAQDLARAF LDKRDEWSEVGV
>PROKKA_00562 Cystathionine gamma-synthase
MRMQTKLIHGGISED TTTGAVSVPIYQTSTYRQDAIGRHKGYEYSRSGNPTRFALEELIA
DLEGGVKGF AFASGLAGIHAVFSLLQSGDHVLLGDDVYGGTFRLFNKVLVKNGLSCTIID
TSDLSQIKKAIKPN TKALYLETPSNPLKITDLAQ CASVAKDHGLLTIVDNTFATPYYQN
PLLLGADIVVHSGTKYLG GHS DVVAGLVTTNNEALAQEIAFFQNAIGGVLGPQDSWLLQR
GIKT LGLRMEAHQKNAFCVAEFLEKHPKVERVYYPGLPTHPNYELAKAQMRGFSGMLSFT
LKN DSEAVAFVESLKL FILGESLGGVESLVGIPAFMTHACIPKEQREAGIRDGLVRLSV
GIEHEQD LLEDLEQAF AKIG
>PROKKA_00563 Cysteine synthase
MIITTMQDAIGRTPVFKFTNKDYPIPI NSAIYAKLEHLNPGGSIKDRLGQYLIEEGFKTG
KITSKTTII EPTAGNTGIALALVAIKHHLKTIFV VPEKFSAEKQQIMRALGALVINTPTS
EGISGAIKKSKELAESIPDSYLPLQFENPDNPAAYYHTLAPEIVQELGTNLTSFVAGIGS
GGTFAGMARYLKERIPAIRLIGVEPEGSILNGGEPGPHEIEGIGVEFISPF FANLDIDGF
ETISDEEGFSYTRKLAKKNGLLVGSSSGAA FVAALKEAQR LPEGSQVLTIFPDVADRYLS
KGIYL
>PROKKA_00564 hypothetical protein
MKTIEWNEEQRKA FQDLLREFAALIDAKAQEEKQTGRTPKIPKYGSCQRRLNKFLAPWGY
ACKISLGS GNLSNEPSIAFCRQDILGEGFVNGEKPTPKKGFYIWFAYYWSNDFEFYLCI
GRSIEKDGEKECQKCPAYDKIV IENACYQKLYDDLES DLENITNYFLRFANEFNQIPTAY
FELEPSSTSH
>PROKKA_00565 Chaperone protein DnaK
MGKVIGIDLGTTNSAMAVYEGNEAKIIANKEGKNTTPSIVAF TDKGEILVGESAKRQAVT
NPEKTIYSIKRIMGLMFNEDKAKEAEKRLPYKIVDRNGACAIEISGKVYTPQEISAKILM
KLKEDAESYLGESVTEAVITVPAYFNDSQRKATKEAGTIAGLNVLRIINEPTS AALAYGL
DKKESEKIMVYDLGGGTFDVTVLETGDNVVEVLATGGDAFLGGDDFDNRVIDFLAAEFKS
ETGIEIKNDVMALQRLKEAAENAKKELSSAMETEINLPFITADATGPKHLVKKLTRAKFE
SLTEDLMEETISKIESVIKDAGLTKNEISEVVMVGGSTRIPKVQERVKAFINKDLNKS VN
PDEVVAVGASIQGGVLKGDVKDVL LLDVTPLSLGIETLGGVMTKVIDRGTTIPAKKSQVF
STAEDNQPAVSIMVLQGERELARDNKS LGKFDLQGIAPAPRGVPQIEVTFDIDANGILTV
SAQDKNTGKSQEI KISGSSGLSDSEIEKMVKDAELHKEEDARKKEVIEARNHADSLAHQT
QKSLDEHKTNLNENDANEIQNAINALKDCVKNDNATKTELEDKTKALAQAAQKLGEAMAN
KNNAEQPKKKDDDDVIDAEVE
>PROKKA_00566 heat shock protein GrpE
MKDEHNQEH D HLSQKEPEFCEKACKEQQGEEKQEASEKEGEIKEDFELKYKEMHEKYLRV
HADFENVKKRLERDKSMAL EYAYEKIALDLLPVIDALLGAYKSAVEVDKESALT KGLELT
MEKLHEVLARHGIEGIECLEEFDPNFHNAIMQVKSEEKENGKIVQVLQQGYKYKGRVLRP
AMVSIKND
>PROKKA_00567 Heat-inducible transcription repressor HrcA
MVIDEIFQIMMLRRIKVGSDLNKKESLLDAFVKTYLQILEPISSKRLKELANLKISCATI
RNYFQIFSKEGMLYQAHSSGARLP TFKAFENYWRKSLHFEVLKVNEKRLKSASENFGLFV
LLKKPSLERLERVIECEKRFLILDFLAFSCALGYSVKMEKFLELVGKSVKEVRLIAASV
NALSLARQLERLEYSSAQMTRFNL MGLKTLN SPLFFDILEGKVLERFKKGLHFVEPDCM
LVIRPIEFQNKRMQLLCVGKLECDYEGFFQTISEEE

>PROKKA_00568 hypothetical protein

MNTHTRGIDSNLIHSLQSSISLSMFRKGFFGLYQGSISARIGANQFVINKRNAVFDQLNEN
TLLVLHDKIDYRWKEASQDSPIHASVYREFLDKFIAYARPPYSLAYSLRHNRLLPDYD
GYRSLGEEISIFNPKDYDSWQERADTEILRQLQESKKYFVFIKGCIFAYHRELSKLMEV
FDLIENSCKVLRRLGDLMDYCYNDPRLSV

>PROKKA_00569 hypothetical protein

MTINTHYTPNFTQLQTLNNINNDATIKDRNQAEQDLQSSIQDGFSQDNTQNAPDFDRLQ
ALNAINNDGEIKDKSQVEKSLIDGGNTPEIYSSLDTYA

>PROKKA_00570 hypothetical protein

MMDIYQKNLQALFKKDPLLFAKLKAIKENKKYEVFLGNDSANFNLLDKETNTPLFEKSPL
DSSLELYKNNIYMLYPYLYYFGLGNGVFYRLLLGNLNLKRLVVIEPEIEIIFIVLNLLD
FSTEILENRLILLHASFCNYNMIASLFDMDKKSRLYARMYDLKLFNAYYERYSSQMIEIN
QHFTRALEHGAI SVGNDADKDALIGIKQHAANLPEVIKSPSLVDFVSALKNRDTAIIVSTG
PSLNKQLPLLKEIAPYATLFCIDASFPILAKAGIKPDIVLSLERVDLTAKFYEETPLDFQ
EGVIFALT SIVHKRLIQAIIKKGVKQFSFRPFGYTNLFDLHQYGYVGIGMSAANMAYELVV
HSRFKRCVFIGQDLSFSQSGNSHASGAIYGDREIKPKKDKDKIFIEKYGGNGEVETTLLVW
KLFLEFFEKDI FNTPYKLEVINATEGGARIKGTKEMPFKEVCEKIDKSKPKPPINLIYPT
QSEQAKNLKIARQKCEEI IKYANEKKTQVEEVFLKVAPFLEKVEKLHEEKLEELDFKEL
ENLSAEIDNIKELFDDKRFNSYFMDAIQSYIFHQELHIAEIVCKKTSNEDELRAKQLEYI
YAHKYWLFSLAGGIDCVIEAIKMALKEW

>PROKKA_00571 Flagellin B

MSFRINTNIAALTSHAVGVQNNRDLSSSLEKLSSGLRINKAADDSSGMAIADSLRSQSAN
LGQAIRNANDAIGMVQTADKAMDEQIKILDITKTKAVQAAQDGQTLESRRALQSDIQRL
EELDNIANTTSFNGQQMLSGSFSNKEFQIGAYSNTTVKASIGSTSSDKIGHVRMETSSFS
GEGMLASAAAQNLTEVGLNFKQVNGVNDYKIETVRISTSAGTGIGALSEI INRFSNTLGV
RASYNVMATGGTPVQSGTVRELTINGVEIGTVNDVHKNADGRLTNAINSVKDRTGVEAS
LDIQGRINLHSIDGRAISVHAASASQVFGGNGFAGISGTQHAVIGRLTLTRTDARDIIV
SGVNFSHVGFHSAQGVAEYTVNLRAVRGIFDANVASAAGANANGAQAETNSQGIGAGVTS
LKGAMIVMDMADSARTQLDKIRSDMGSVQMELVTTINNISVTQVNVKAAESQIRDVDFAE
ESANFSKYNILAQSGSFAMAQANAVQQNVLRLLQ

>PROKKA_00572 DNA topoisomerase 1

MKHLIIVESPAKAKTIKNFLDKNYEVIASKGHVRDL SKFALGIKIDETGFTPNYVVDKDH
KELVKQIIELS K KASITYIATDEDEGEAIGYHVACLIGGKLENYPRIVFHEITQNAILN
ALKTPRQIDMSKVNAQQARRFLDRIVGFKLSSLIASKITKGLSAGRVQSAALKLVIDRER
EIKAFKPLTYFTLDALFDPHLEAQLISYKGNKLKVQELIDGKKAQEIKNELEKESYIISS
IVKKS KKSPTPPFMTSTLQQSASSLLGFSPKTMSIAQKLYEGVATPQGVMGVITYMRT
DSL NIAKEALEEARNKILKDYGKDYLPKAKVYSSKNKNAQEAHEAIRPTSIILEPNALK
DYLKPEELKLYALIYKRFLASQM QDALFESQSVVACEKGEFKASGRKLLFDGYKILGN
DDKDKLLPNLKENDPIKLEKLESNAHVTEPPARYSEASLIKVLES LGIGRPSTYAPTISL
LQNRDYIKVEKKQISALES AFKVIEILEKHFEIIVDSKFSASLEELD NIAQNKADYQQV
LKDFYYPFMDKIEAGKKNIIISQKVHEKTGQSCP KCGGELVKKNSRYGEFIACNNYPCKY
VKQTENAHDGANQELCEKCGGEMVQKFSRNGAFLACNNYPECKNTKSLKNTPNAKEIIEG
VKCPECGGDIALKRSR KGSFYGCNNYPKCNFLSNHKPINKRCEKCHYLMSERIYRKKKAH
ECIKCKERVFLEEDNG

>PROKKA_00573 molybdenum cofactor biosynthesis protein A

MAKENLPVVFGPVL SRRFGKSLGVDLSPSKKQCNYNCIYCELGKAKPIERMEEVIKVETL
INAIQNALNNLTTPIDVLTITANGEPTLYPHLLELIQNIKPF LKGVKTLILSNGSLFYEP
KVQQALKEFDIVKFS L DAIDLKA FERVDPKPYSKDINKILEGILSFSQIYQGQLVAEVLLI
KGVNDSANNLKRIADFLKKINIARVDLSTIDRPSSF KAPKLSEDELLKCSLFFEGLCVSL
PKRSTAQAAQAKKLISCGIDELLALISRRPLSAEEAPLILDPNAFRHLETLLNHKQITIK
KVGSLIFYCAF

>PROKKA_00574 hypothetical protein

MNVCFADDQSKKEVAETQKEAENARDRANKSGIELEQEKQKTSNIETNNQIKVEQEKQKT
SNIEINNQIKVEQERQKANNTQKDLVKEQKDLVKEQKDLVKEQKDFIKYAEQNCQENHSQ
FFIKKLGKGGIAIEIEAECKTPKPTKTNTQTPIQPKHLPNSKQPHSQRGSKVQELIAYLQ
KELESLPYSQKAIKQVDFYRPSSIAYLELDPRDFKVTEEWQKENLKIRSKAQAKMLEMR
NLKPDPQAHL SASQSLLFVQKIFADVNKEIKVVANTEKKVEKAGYGYSKRM

>PROKKA_00575 hypothetical protein

MPVIRVLVMLATMMMKLVTAKEKKVFKNVGISIMGIAFWEAIKDSIKKQIKKSDWICGN
VKTADDYLKTHPNSWFNSAIGVATITTMLYECVFC
>PROKKA_00576 hypothetical protein
MVIKSVKAGKTNKVGKNTETANTKANKETHFKQANAITNIIRSIGGFFTKIVKKVRELVK
KHPEKSSAALVVLTHAACKRAKELDDKVQDKSKQAEKENQINWWKYSGLTIATSLLLAAC
SVGDTDKQIELEQEKQKTSNIETNNQIKVEQEKQKTSNIEINNQIKVEQEKQKTSNIETN
NQIKVEQEQQKTEQEKQKTNNTQKDLINKAEQNCQENHSQFFIKKLGKGGIAIEIEAEC
KTPKPTKTNTQPIQPKHLPNSKQPHSQRGSKVQELIAYLQKELESPLPYSQKAIQVDFY
RPSSIAYLELDPRDFKVTEEWQKENLKIRSKAQAKMLEMRNLKPDPAHLSASQSLLFVQ
KIFADVNIKIKVAVANTEKKVEKAGYGYSKRM
>PROKKA_00577 Phosphoenolpyruvate synthase
MRYIKFFKELNNKDVNLVGGKNASIGEMFQELVPIGKVPDGFATSEAYWYLLEQGGAK
QKIIELLENVDATEIDVLKIRSKQIRELIFGTPFPSPDLRDEIFQAYEILSQQYNMKEADV
AVRSSATAEDLPDASFAGQQDTYLNKIGKTELIHYIKSCLASLFTDRAISYRASRGFDHL
KVALSVGVQKMRADKGSAGVMFSIDTETGFKDAVFITSAWGLGENVVGGTINPDEFYVF
KPTLEQNKRPPIKRQLGNKTQKMVYAPRGSEHPTRNIKTTKKEWQSFSLSDEDVLILAKY
AIEIEKHYSKEAKQYRPMIDIEWAKDGESGEIFIVQARPETVQSQKTKEESQVFEKFKFKN
PNEKKEIILQGRAIGSKIGSGKVRIINDLEHMNSFKEGEILVTDNTDPDWEPCMKKASAV
ITNRGGRTCHAAIVAREIGVPAIVGVSGATDNLYTGMEITVSCAEGEEGYVYAGIYEHEI
ERVELSNMQETQTKIYINIGNPEKAFSFSQLPNHGVGLARMEMIILNQIKAHPLALVDLH
HKKSVEKNEIENLMAGYANPKDFFVKKIAEGIGMISAAFYPKPVIVRTSDFKSNEYMRM
LGGSSYEPNEENPMLGYRGASRYSESYNFAFSWECEALALVREEMGLTNMKVMIPFLRT
IEEGKKVLEILRKNLLESGKNGLEIYIMCELPVNVILADDFLSLFDGFSIGSNDLTQLTL
GVDRDSELVSHVFDERNEAMLKMFKAIEACKRHNKYCGICGQAPSDYPEVTEFLVKEGI
TSISLNPDSVIPTWNAVAKLEKELKEHGLT
>PROKKA_00578 Threonine--tRNA ligase 2
MSAELIAVYKDEQIIDLESKVLGLSDGIKALKGTEPIYFDDSPLALEVIRHSCAHLAQ
SLKALYPDAKFFVGPVVEEGFYDFKTSSKISEEDLPKIEAKMKEFAKSKLAITKETLTR
EQALERFKGDELKHAVMSKISGDAFGVYQQGEFEDLCKGPHLPNTRFLNHFKLTKLAGAY
LGGDENNEMLIRIYGIAFATKEGLKDYLFQIEEAKKRDHRKLGVELGLFSFDDEIGAGLP
LWLPGGARLRKRIEDLLSQALLLRGYEPVKGP EILKSDVWKISGHYDNYKENMYFTTIDE
QEYGIKPMNCVGHKIVYQSALHSYRDLPLRFYEGVVHRHEKSGVLHGLLRVREFTQDDA
HIFCSFEQIQSEVSAILDFTHKIMKAFNFSYEMELSTRPTKSIGDDEVWEKATSALKEAL
KEHRIDYKIDEGGAFYGP KIDIKITDALRRKWQCGTIQVDMNLPERFKLAFTNEHNHAE
QPVMIHRAILGSFERFIAILSEHFGGNFPFFVAPTQIALIPINEEHHVFALKLKEELKKR
DIFVEVLDKNDLSLKKVRLAEKQKIPMILVLGNEEVETEILSIRDREKQDQYKMPLKEFL
NMVESKMQEVSF
>PROKKA_00579 Translation initiation factor IF-3
MSRNEVLLNGDINFKEVRCVGDNGEVYGISSKEALNIAQNLGLDLVLISASAKPPVCKV
MDYNKFRYQNEKKIKEAKKKQKQIEIKEIKLSTQIAQNDINYNKVKHAREFIESNKHVKFK
VVLKGRESQNSKAGLDVLFVQTMMDLANPEKEPKTEGRFVSWMFVPKAKEALKNEKKT
KENNPPFNRLNLMKGENHAKNED
>PROKKA_00580 50S ribosomal protein L35
MPKMKTNRGASKRFKVKKNLIKRGSAFKSHILTKKSPKRKANLNAPKHVHHTNAHSVMSL
LCRA
>PROKKA_00581 50S ribosomal protein L20
MRVKTGVVRRRRHKKVLKLARGFYSGRRKHFRKAKEQLERSMYAFRDRKQKKREFRSLW
VVRINAACRMHNTSYSRFMHALKVANIELDRKVLADMAMNDMQAFKSVLESVKEHL
>PROKKA_00582 Helicobacter outer membrane protein
MKKSIVVGAISLAMTSLLSAETPKQEKAIKTSPTKKGERNAAFIGIDYQLGMLSTTAQNC
SHGNCNGNQSGAYGSNTPNMPTASNPTGGLTHGALGTRGYKGLSNQQYAINFGFVVG YK
HFFKKSPQFGMRYYGFFDFASSYKYTYNDYGMRDARKGSQSFMFGYGAGTDVLFNPAI
FNRENLFHGFFLGVAIGGTSWGPTNYYFKDLADEYRGSFHPNSNFQVLVNGGIRLGTKHQG
FEIGLKIQTIRNNYYTASADNVPEGTTYRFTFHRPYAFYWRYIVSF
>PROKKA_00583 hypothetical protein
MKKLAFSLLLFTGTFLGLFLNASDFKSMDDKQLLEQAGKVAPSEVPEFRAEVNKR LAMVN
EEDRKNYKADFKKAMDKNLASLSQEDRNKRKKEILEAIANKKKMTMTMKEYREEGLDLHDC
ACEGPFHDHES

>PROKKA_00584 hypothetical protein

MKRILFFLAATTFLLRAEADSATINTTVDPNVMFSESSTGNVKKDRKRVLKSMVNLEKER
VKNFNQYSETKMSKGDLSAFGAFFKGSLESCVEQKVCYIEHKDGKVSFVNDREKFYKHV
LKDLGTELSLPLFNWLYKGSDFGALHEQFGDMYDGIKYLISMVRVSQKEKARKLDAITL
KKMEDQAEKDTKAAFQKRSSGELESHTDSPEFISSSKRTQNASNSDLDDSTTNANTLKETA
SKEPEASSKKEKKPKKKRRLSKKEKQQQALQQEFQISDSSKSEK

>PROKKA_00585 L-serine dehydratase 1

MASFSILSIFKIGVGPSSSHTIGPMEAGARFCGLLKGILEQVVRVQITLHGSLALTGKGH
LSDEAVLIGLHGIHANELEVTTKKALLHEALENKVLKLANQRNIQFDYSKNLIFDNKPLA
RHQNALILKAFNAKNEVLKEETYYSVGGGFVYTEKELDNLSEESGNESVAYDFSSAKELL
ELCQKHQKNIAEIVRLRENALKNRPDATMVKIYHAMLECYDNGANSKEKYLPGSLRVTRL
APSIKTRLEKHPTSGKDPLALIDYISLYARAIAEENASGGKVVTAPTNGACAVVPSVLLY
AKNHLFENLSQKAINDFLLTSAAIGYLYKKNASLSGAEAGCQAEIGVASSMAAGGLAHL
QATTQQVLIASEIAMEHHLGLTCDPVGGLVQIPCIERNVLGAIKAISASKLALEDEYKPK
VSLDEVIATMYATGKDMNEKYKETSLSGLLAKTLEC

>PROKKA_00586 Serine transporter

MVQEKAVIRDPKKLNADFRLRMVSLFGTAVGAGILFLPIRAGGHGIWAIVVMSAIIIFPLT
YLGHRALEYFIGSKDKEDITMVVRSHFGTQWGFLITLLYFLAIYPICLVYGVGITNVFDH
FFTNQLHLAPFHRGLLAVALVSLMMLVMVFENATIVTRICNALVYPLCLILLLFSLYLIPY
WQGANLFVVPSFKEFVLAIWLTLPVLVFSFDHSPIIISTFTQNVKEYGAFKEYKLNQIEL
GTSMLLLGFVMFFVFSCVMCLNADDFVKAREQNIPILSYLANTLNPLINYAGPVVAFLA
IFSSFFGHYYGAKEGLEGIIIQSLKLKASKTLSISVTIFLWLTITLVAYINPNILDFIE
NLGGPIIALILFVMPMIAFYSSSLKRFRNFKVDIFVFVFGSLTALS VFLGLF

>PROKKA_00587 Phospho-2-dehydro-3-deoxyheptonate aldolase

MSNTTWSPASWHSFKIEQHPTYKDKQELERVKKELRSYPPLVFAGEARNLQERLAQVIDN
KAFLQGGDCAESFSQFSANRIKDMFKVMMQMAIVLTFAGSIPIVKVGRIAGQFAKPRSN
ATEILDNEEVLSYRGDIINGISKKEREKPERMLKAYHQS VATLNLIRAF AQGGLADLEQ
VHRFNLDVFKNNDFGQKYQQIADRITQALGFM RACGVEIEKTPILREVEFYTSHEALLH
YEEPLVRKDSL TNQFYDCSAHMLWIGERTRDPKGAHVEFLRGVCNPIGVKIGPNASVSEV
LELCDVLNPRNIKGRNLNLRVMGSKMIKERLPKLLQGVLEEKRHILWSIDPMHGNTVKTS
LGVKTRAFDSVLDEVKSFFEIHRAEGLASGVHLEMTGENVTECIGGSQAITEEGLSCHY
YTQCDPRLNATQALELAFLIADMLKKQRS

>PROKKA_00588 hypothetical protein

MRISLLAVILALLFVACHETTKQILQNEADSTPSEKTIWQPEQK

>PROKKA_00589 Putative peroxiredoxin bcp

MEKLEVGLAPDFRLKNNDGVEISLKDLLHKKVVLYFYPKDNTPGCTLEAKDFSALFSEF
EKKNVAVVGVSPDNAQSHQKFISQCSLNVILLCDEDKKAANLYKAYGKRMLYGKEHLGII
RSTFIINTQGVLEKCFYNVKAAGHAQKVLESL

>PROKKA_00590 Lactate utilization protein C

MSKELILKRIKEARAKHAIQGANPIYRNIKVEFEDLVEEYKHFQVLNKAEVIESAKENL
EQAILKALENFKSKKVLHSTDNLNFEAFKDFTLQPYDKEIEAMREELFEIDTALLHGVC
GISSLGMIGAVSSHASPRLLSLITLNCIILLKKE SIVRNLSEGMQALKNQSQNGVLPTNM
LLIGGPSRTADIELKTVFGVHGPQKVAILLY

>PROKKA_00591 Lactate utilization protein B

MEKYHSDQEEYEEIITDQLGDMQLRENLR SAMDTLRANRKNLIK NRYSEWENLRELGKEVK
LKILSRLDEYLELF EKNAAQNGFKIHYAKDGDEANEI IYNLAKEKNIKRILKQKSMASEE
IGLNHYLKEKG IQAQETDLGELIIQLINEHPVHIVPAIHKNRKQIGKIFEEKLNAA YEE
EPEKLNAIARKHMRKEFESFKMGISGVNFAIANEGAIWLVENEGNGRMSTTACDVHVAIC
GIEKLVESFDAAI LNNLLAPSAVGVPITCYQNIITGPRKEGDL DGPKEAHIILLDNRS
NILADEKYRALSCIRCGTCLNHCPVYDKIGGHAYLSTYPGPIGVV VSPQLFGLNNYGH
PNLCSLCGRCTEVCPEIPLAELIRD LRS DKGEGRGVIKAKSTQHSGMEKFSMKMFAK
MASDGAKWRFQLKMAQFFSPLGKLLAPILPLVKEWASVRTLPNMDTSLHAKVQHLEGVIY
E

>PROKKA_00592 Lactate utilization protein A

MKVNFFATCLGAAYSNASLNAIKLLRKENLEVVFKKDQTCCGQPSYNSGYEETKKVVL
YNIKLYSNNDYP IILPSGSGTGMMRHDYLELFEGHAEFNMVKDFCSR VYELSEFLDKKLQ
VKYEDKGEPLKITWHSNCHALRVAKVIDSAKNLIRQLKNVELIELEKEEECCGFGGTFSV
KEPEISAVMVKEKIKDIESRHVDVIVSADAGCLMNISTAMQKMGS LTKPMHFYDFLASRL

GL

>PROKKA_00593 L-lactate permease

MEFYQVYDPLGHIWLSALVALSPIALFFISLIVFKLKGYSAGFLSLLLSIIIALFVYKMP
VQMVSAFFFYGLYGLWPPIAWIVIAAIFLYNLSVKSGYFEILKESILTTPDHRILVILI
GFCFGSFLEGAIGFGGPVAITAAILVGLGLNPLYAAGLCLIANAPVAFGAVGIPITAMA
SVVGIPLELEISQMVGRVLPFISIGIPFFIVFLMDGFRGIRETFPAVAVTGFSFAIAQFLS
SNYLGFPQLPDIISALVSLISTTLFLKFWQPKRIFTSNGKEPTINTEKHHICKVVVAWMPF
VLLTITITIIWTQPWFKALFKEGGALAFSSFAFEFDSISQKIFKTVPIVTEATNFPVVFKL
PLILTTGTSIFLAALLSVFLLRVKISDAIGVFGATLKEMRLPILTIGVVLAFAYVANYS
MSATLALALADTGHVFTFFSPVVGWLGVFLTGSDTSSNLLFGSLQMLIATQLGLPEVLFL
AANTSGGVVGKMISPSQIAIACAAGVLVGKESELFRTTVKYSIALAIIMGIVFTLIAYVF
PFIIPVTPT

>PROKKA_00594 L-lactate permease

MSEFHQVYDPLGSIWLSALVALLPILLFFLSLMVFKLKGYTAAFLSVALSAVIAVLVYKM
PVSMVGSSFLYGLYGLWPPIAWIIIAAIFLYKLSVKSGYFEILKESVQSITLDHRILVIL
IGFCFGSFLEGAIGFGGPPIAITAAILVGLGLSPLYAAGLCLIANAPVAFGAVGIPISAM
ASAVGVPAILISAMTGKILFFVSLLVPPFFIVFLMDGFGKIKETFPAVFIAAFSFAQAQFL
SSNYLGPELPGIISALVSLVATALFLKFWQPKAIFRSDGKAASFTKSNHHICKVYVAWSP
FVILVLVIVLWIQPFKALFEKDGLLAFSNFYFEFNNISNHIFKSPPFVEANQSVSFPVV
FKFLLINTVGTSIFLAALVSMVLVRVIDALSVFGETLKEMRYPILTIGLVLSFAYVSN
YSGISSTLALALHTGLAFTFFSPLIGWVGVLGTGSDTSSNLLFGSLQQQLTAQRLHLPEV
LTLTANTVGGTLGKMISPSQIAIACAAGVLGKESDLFKFTVKYSLIFVAIMGVVISAIA
YLIPEVVPPIK

>PROKKA_00595 A/G-specific adenine glycosylase

METLHNALLKWEYEEFGRKGLPFRNLKGVNAPYEVYISEVMSQQTQISTVIERFYSPFLKA
FPTLKDLANAPLEEVLWRLGLGYLLRAKNLKKSAEICVKEHNSQLPNDYQSLLKLPGIG
TYTANAILCFGFREKSACVDANIKRALLRLFGLDPNNTAKDLQIKANDFLNLNESFNHNQ
ALIDLALICSPKPKCAICPLNPYCLGKNNEKHTLKKKQEIYQEEERYLGVVIQNNQIAL
EKIEQKLYFGMHHPNLKENLEYKLPFLGAIKHSHTKFKLNLNLYLATTKDLKNPIRFYS
LKDLETLPISMTLKIILNFKQKNLFGS

>PROKKA_00596 Citrate carrier

MIKQTLIIILAPFFFIATLLYFLGAPDGLNPNWLYFCIFIGMIIIGLVLEPVPPGLVALSAL
VLCVALKIGASDKVESANKAISWGLSRYANKTVWLVFVAFILGLGYEKSLLGKRIALLLI
RFLGQTPGLGLGYAIGLSELCLAPFIPSNARSGGILYPIVSSIPPLMGSTPNNNPKIGA
YLMWVALASTCITSSMFLTALAPNPLAMEIAAKMGVNEISWFSWFLAFLPCGVVLILLVP
LLAYKTCKPTLKGSKVSLWAKKELEGMRGFSLEILMLSLTLALLGWIFGKSLGLHAS
ATALIVMVLMAFCKIVSYEDIKKNKSAFNIFLLGSLLTAMAGGLKNVGFLNFIGNAAKNF
LEHAHLDPPIAVLFIYALFYLSHYFFASITAHVSALFALFVGIGSHIQGVNLQELSLFLM
FSLGIMGILTPYGTGPSTIYYGSGYIPSKDFWKWGFIFGVVYLIVFLSVCTPWVKFIAYR
WL

>PROKKA_00597 hypothetical protein

MQENVPLSYDYSISKFLYAMVGFGIIGMLIGIVLAFELSFNPNLYIAGEYGIFGRRLRPL
HTNAVIYGFITLGGIOWASWYYIGQVRLKITYHQHPFLKIVGLLHFWLWILLILGVISLFA
GLTQSKEYAELMWPLDIIIVVAVWLWGVNMFSGMSVRRENTIYVSLWYYIATYVGIAVMY
IFNNLSVPTYFVADMGSVWHSISMYSGSNDALIQWWGHNVAFAVFTSGVIGTIYYFLPK
ESGQPIFSYKLTLSFWSLMFVYIWAGGHLLIYSTVPDWVQTLSSVFSVVLILPSWGTAI
NMLLTMRGQWHQLKESPLIKFLVLASTFYMLSTLEGSIQAIKSVNALAHFTDWIIGHVHD
GVLGWVGFTLIASMYHMTPLRFKREIYSGRLVDFQFWIMTLGIVLYFSSMWIAGITQGM
WRDQYGNLTYQFIDTVKVLIPYYNIRGVGGLMYFIGFIIIFAYNIFMTITAGKKLEREP
NYATPMSR

>PROKKA_00598 Cytochrome C oxidase, mono-heme subunit/FixO

MFSFLEKNPFFFTLAFIFVFAIAGLVEILPNFFKSARPIEGLRPYTVLETAGRQIYIQEG
CYHCHSQLIRPFQAEVDYRGAYSLSGEYAYDRPFLWGSKRIGPDLHRXGDIYRTTDWHEKH
MLDPKSVVPHSIMPAYKHLFIKKSDFDTAYAEALTQKKVFGVPYDTENGVKLGSVVEEAKK
AYLEEAKKITADMKDKRVLDAIQRGEVLEIVALIAYLNSLGNLRINANQNAK

>PROKKA_00599 Cbb3-type cytochrome oxidase component FixQ

MDLESLRGFAYAFFTILFTFLYAYIFSMYRKQKKGVDYERYGYLALNDALEDELIEPR
HKKVHDNGIKEN

>PROKKA_00600 Cbb3-type cytochrome c oxidase subunit CcoP
MDFLNDHINVFLGLIAALVILVLTIIYESSSLIKEMRDSKSQGELVENGLIDGIGEFANNV
PVGWIASFMCTIVWAFWYFFFYPLNSFSQIGQYNEEVKAHNQKFEAKWKHLGQKELVDM
GQGIFLVHCSQCHGITAEGHLGSAQNLVLRWGKEEGIMDTIKHGSKGMDYLAGEMPAMELD
EKDAKAIASYVMAELSSVKTKNPQLIDKGKELFESMGCTGCHGNDGKGLQENQVFAADL
TAYGTENFLRNILTHGKKNIGHMPSFKYKNFSDLQVKALAEFIQSLKPLED
>PROKKA_00601 hypothetical protein
MKFLNGLAGNLLIVVILLCAVFFTLKAIHIQKEQATNYYRYKDINALETNTQNRANYE
LVNQGSKK
>PROKKA_00602 hypothetical protein
MKFTTLEKILALMVVATILMTIVISFVFNLFST
>PROKKA_00603 hypothetical protein
MRILWLIVVFCCLGADDYVLNNSKGRLEKSVAFVEGVSKELYLKTGVRFVIDMTDFEK
NPIALAAKKERQNYQEGFLKQLKPPFVVFYHDAQKIELVANPKDLLDTDKIFFEKIAP
LLPTNPKEYTPQRISAMLINGYSVAVDALAQKYRVNITQNFNAPKGVTFVKVAVYILLLT
LLGAFLGLYFFKKS
>PROKKA_00604 hypothetical protein
MKEKNFWPLGIMSVLILGLGIVVFLVVFALKNSPKNDLVYFKGHNEVDLNFNAMLKTYEN
FKSNYRFSVGLKPLIKSPKTPILPYFSKGTHGDKKQENLLKNALILEKSNTLYVQLQPL
KPALDAPNIQVYLAFYSPSPQPRLLGTDCRSACEPLKFDLLES DKMGYKILFKFVFKN
KEELILEQLAFFK
>PROKKA_00605 hypothetical protein
MGFLKIFKHDAALGQVGNVVGNFLITLTILAVCFSSQSAEETTMTLSYTLFFVLGAFL
AAISVGAIKNLNALFSKRGVLSFSLPVSLESLLLPKILLPMVFFIFSLFWFVASVRLGY
LFNAQSSVLFILHTALKTFVLKPTKTGLIALFLGLVLMKFLFVLSVLNAARIKKARFL
GLLFILVGVVLELAFNSLLPLMSSSLSSINEGFYFLLQQLQENKYYLLWGVDFLKILL
YGMIRYLLMRKLELD
>PROKKA_00606 1,4-dihydroxy-6-naphtoate synthase
MISVAHSPDADDIFMYAIFGWIDCPIKNKTFHNIALDIETLNQEALKNTYDVSAISFG
LYPKIANDYALLPTATSFNGYGPKLVKKKGVKLKDFRVALSGEHTTNALLFKIYYKHA
RIAYMNFLDIEKAVLEEKVHAGVLIHENILDFHNELEVEKELWDVWKELIKVDLPLPLGG
MAIRRSIPLYRAILIKKALIKAVEVALKHQNLSEMLLERSLIRVNKERLQTYLSLYANE
TSTRLSEIQILAIDKLFELGYQHGFYASLLKTKDCLLTDEYLQYRFS
>PROKKA_00607 recombinase A
MAIDEDKQKAISLAIKQIDKVFGKALVRLGDKQVEKIDAISTGSLGLDLALGIGGPVK
RIEIIYGPESGKTTLSLHIIAECQKNGGVCAFIDAEHALDVHYAKRLGVDTENLLVSQ
DTGEQALEILETITRSGGIDLVVVDSVAALTPKAEIDGDMGDQHVGLQARLM SHALRKIT
GVLHKMNTTLIFINQIRMKIGMMGYGSPETTTGGNALKFYASVRIDIRRIAALKQNEQHI
GNRAKAKVVKNKVAPPFREAEFDIMFGEGISKEGEIIDYGVKLDIVDKSGAWLSYQDKKL
GQGRENAKALLKEDKALADEITLKIKESIGSNEEIMPLPDEPLEEME
>PROKKA_00608 Enolase
MLTIKDIHALEVMDSRGNPTIQASVILSDNTKASAIVPSGASTGKREALELRDNDKTRFL
GKGVLRA CENVNSVIKHHLIGLEAINQAFVDERLRALDGT PNYANL GANAVLGVS MLAR
ASAKALNPLYRYLGGANALTLPVPM LNIINGGTHANN SIDFQ EYMIMPLGFESFKEALR
ASAEVYHTLKKLLDGKNQLTSVGDEGGFAPNFNNNVEPLEVISQAIEKAGYKLGEI IALA
LDVASSELVDEHFNYHLKGENKILDSHELVA YYYKELVAKYPIVSIEDGLSEDDWEGWAF
SKELGRQIQLVGDDLFVTNASLLQKGIEKNIANAILIKPNQIGT ISETLETIRLAKHHAY
QCVMSHRSGESED SFIADFAVALNTGEIKTGSTARSERIAKYNRLLEIEHELKGGIYIGK
ELFKHG
>PROKKA_00609 hypothetical protein
MASGLFENDGIKDNKARDFFYSHSSLIVFFLLLLGFGYYLGKLLFGGSSLEVYDLDRDKH
ERLQQEITELQSKNVRLQKRLFELKELRPRD
>PROKKA_00610 hypothetical protein
MLKKMVGLVVVLSVLLARDNPFEP EINSKNLQGGFNGIYDSYFKEIHVDLPTSARILKQI
TLTYQDIDGSIH SKVVGIDKSIDWHYPLKLSQHTLNQA AF EKRYQIQDFDFLMANNTMIL
RSPYKILRSFVLVNPYRIVLDTQKGPLDIYQNRDLNQKFFSQIKVGTHKDYRITLILDG
KYRYLLEEKNGAYELKLK
>PROKKA_00611 Shikimate kinase

MQHLVLIGFMGSGKSSLAQELGLALKLEVLDTDMIISERVGLSVRGIFEELGEDNFRMF
KNLIDELKTLKTPHVISTGGGIVVHDNFKGLGTTFYLKIDFETLIKRLNQKEREKRPLLN
DLTQAKELFEKRQALYEKNASFIIDARGGLSNSLKQVLQFIA

>PROKKA_00612 hypothetical protein

MLSRDIVQYSKIRTELYAYLTYLFSHNIRNHLPEITLDYLNKQIRKMHAEIKMAKNFFVL
DAKGMILILKPSQLKEQGHEGILEHDLTEGIELESHASFSDKYFFYQAVSEKRCILTD
PSKKGNHLVVSASYPVYDQNNDLAFVVCLQIPLRVAIEISSPSKYFRTFSEGSVMYFMI
SIMLTLVSLLLFVKCISSFWTAIVNFSSFDIKEVFHPIVLLTLALATFDLVKAIFEEVL
GKNSGDNHHAHRTMIRFLGSIIIALAIEALMLVFKFSVSEPKITYAVYLAVGVAVLLI
SLAIYVKFAYSVLPKRER

>PROKKA_00613 Glycosyl transferase family 8

MSIIPIVIAFDNHYAMPAGVSLYMLACAKTEHPQSQNDSEKLFYKIHCLVDNLSLENQ
HKLKETLAPFSAFMSVDFLDISTPNLYTTSIEPSAIDKIHEAFLQLNIYAKTRFSKMVMC
RLFLASLFPQYDKIIMFDADTLFLNDVSESSFIPLDSYFGAAKDFSSPKSLKHFQTERE
REPRQKFSLHEHYLKEKDMKIIICENHYNVGFILVNLKLWRADKLEERLLNLTHQKGQCVF
CPEQDLLTLACYQKVLQLPYIYNAPFMANQKRFIPDKKEIVMLHFYFTGKPWVLPFTSY
SKEWHETLLKTPFYAEYSVKFLKQMTFELSLKDKQKTFEFLAPLLNKKTLLEYVFFRLKR
IFKRLKENF

>PROKKA_00614 Putative beta-lactamase HcpD precursor

MIKSWTKKWFLILFLMASCFGHLVATTGEKYFKMANQALKRGDYHKAVAFYKRSCNLRVG
VGCTSLGSMYEDGDGVDQSIPKAVFYRRGCNLRNYLACASLGSMYEDGDGVQKDLPKAL
YYYRRGCHLKGGVSCGSLGFMYFNGTGVKQNYAKALSLSKYACSLNYGISCNFVGMYRN
AKGVQKDLKKALANFKRGCHLKDGASCVSLGYMHEAGMNVKQNGEQALNLYKKGCFLKEG
SGCHNVAVMYTGTGTPKDLDAISYYKKGCTLGFSGSKVLEEVIKKSDDLQDDAQND
TQDDTQ

>PROKKA_00615 putative transcriptional regulatory protein

MGRAFEYRRAAKEKRWDMKSKVFPKLAKAITLAAKGGSEPDTNAKLRTAILNAKAQNMP
KDNIDAAIKRASSKEGNLSEITYEGKANFGVLIIMECMTDNPTRTIANLKSIFYNKTQGAS
IVPNGSLEFMFNRKSVFECLKNEVKNLKLSDLEDFALIDYGLEEELEEVEDKIIIRGDYN
SFKLLNEGFEESLKLPIKASLQRIATTPIELNDEQMELTEKLLDRIEDDDDVVALYTNE

>PROKKA_00616 Delta-aminolevulinic acid dehydratase

MFKRLRRLRSSENLRAMIRETHLNINDFIAPLFVIESDSCIKNEISSMPGVYQMSIKPLL
KECEELVGLGIKAVLLFGIPKHKDATGSHALNKDHIVAKATREIKKRFKDLIVIADLCFC
EYTDHGHCIGILENASVSNKDTLEILNLQGLILAESGVDILAPSNMMDGNVLSLRKALDKA
GYFHTPIMSISTKFASSYGPFRDVANSAPSGDRKSYQMDYANQKEALLESLEDEKQGA
DILMVKPALAYLDIVKEIRDHTLLPLALYNVSGEYAMLKLAQKHNLINYESVLLETMTCF
KRAGADMIISYHAKEVANLLQRN

>PROKKA_00617 Sensor protein PfeS

MRFSIFFKVVALFMITLFSFGAFAYYFVSSQISHENYQNEMRHYQFVTTINEILNNYSY
RAIEDYLYKIGFRETITENLEKVLAKRRHQHHRNIGYAEVFKFSDMVFILLKKDEHFVL
YKDLHSVSYRNYFLAITVGLLLILFLFLFVLQSLPLRELRSQVKRFAQGDKSVSCKSKQ
KDEIGDLANEFDNCILKINAMNESRVFLRSIMHELRTPTITGKILSSMLKEELSCKRFS
SIFDHLNMLIEQFARIEQLASKNYGSNKEKFLMSDLIDKIEKMLLIDENKESPIHVFSSN
YIIEADFELFSIALKNMVDNAIKYSDDKQVFLDFIGNNLVVSNSKSKPLKEDFEKYLQPYF
KSSNPSQAHGFGLGMYIKNALAMGLNLSYHYGNRICFTIHDCVFNSFYDLEEDNEEL
PPPPPKI

>PROKKA_00618 Response regulator ArlR

MIEVLMIEDDIELAEFLSEFLLQHGIHVNTNYDEPYTGISTANTQNYDLLLLDLTLPLNDG
LEVCRISKQKHIPIISSARSVEDKIKALDYGADDYLPKPYDPKELLARIQSLLRRSH
KKEEVSEPGDANIFRVDKDSREVYMHKKLDLTRAIEYIISLLISKKGIVFSRESIAIES
ESINPESSNKSIDVIGRLRSKIEKNPKQPQYIISVRGIGYKLEY

>PROKKA_00619 hypothetical protein

MAGRMELKNIISETLNEIEKMAKTIDDGFNATQKTPSFFKTPSHLQNTPN SKNANTPLEP
KNAAKIETQEKITEENTEEKEEEVPEIIIEEIVQENPTQALIPNERVFLKGLLERTLVLF
KGMQALEEKEALKRLDLVARFLQYQLSVLEKRLESLERENTE

>PROKKA_00620 hypothetical protein

MPLETITLAHIYEEQGFFEEALQIYTNILKKTPDHAEALKQIKRLEKIQKNGAPFKHDAI
LERHYLNFIKGDYLSVENLEKWLVEWN

>PROKKA_00621 putative protease YhbU precursor

MNQVELLS P A G N L K K L K I A L N Y G A D A V Y G G V S H F S L R N R A G K E F T L E T F K E G I D Y A H A L N
K K V Y A T I N G F P F N S Q L K L L E E H I D K M A E L E P D A F I I A A P G V V K L A L K I A P H I P I H L S T Q A
N V L N L L D A Q V F Y D L G V K R I V C A R E L S L N D A I E I K K A L P D L E L E I F V H G S M C F A F S G R C L I
S A L Q K G R V P N R G S C A N D C R F D Y E Y Y V K N P D N G V M M R L V E E E G V G T H I F N A K D L N L S N H I A
E I L S S N A I S A L K I E G R T K S S Y Y A A Q T T R I Y R L A V D D F Y H N T L K P S F Y A S E L N T L K N R G F T
D G Y L M R R P F E R L D T Q N H Q T A I S E G D F Q V N G E I T E D G R F F A C K F T T T T N T A Y E I I A P K N A A
I T P I V N E I G K I Y T F E K R S Y L V L Y K I L L E N N T E L E T I H S G N V N L V R L P A P L P A F S F L R T Q V
R V

>PROKKA_00622 hypothetical protein

M T Q E E L D A L M N G G D L E N L E A L E A K E E T N E E P K E K E E T K E S K E S S E K M T V K K E D A E K Y G K
I S P N E W P P P P P T E E H K V V H Q L D D V T R D S E V K A T Q I F D Q L D L I G A S A E K I A K M V K K I Q E P L
Q K H Q E I F D N L H G H F P N V E S F K T A L N E Q Q E I L N A L K S I E E E A A N C S D S S M Q A M D I M Q F Q D I
H R Q K I E R V V N V M R A L S Q Y M N S L F E G K I D D S K R V S S A T F I T G D D D K D L A S A D D I E A L I A S F
G A K

>PROKKA_00623 Peptide chain release factor 2

M D N Y T Y S E L L K S L Q N K C D N I A L I I K P E K I K Q E L G R I E K E Q E D P N F W Q D V L K A R D T N K E K V
R L N R L L E T Y Q K T K N S L D E S V E L F E L A Q N D N D E V T L S L L Y E E A P I L E H S V Q K V E I E I M L S G
E N D A S N A I I T I Q P G A G G T E S Q D W A S I L Y R M Y L R W A E R R G F K S E I L D Y Q D G E E A G I K G V A F
I I K G E N A Y G Y L K N E S G V H R L V R I S P F D A N A K R H T S F A S V Q I S P E L D D D I D I E I D E K D V R Y
D Y Y R A S G A G G Q H V N K T E S A V R I T H F P T G I V V Q C Q N D R S Q H K N K A S A L K M L K S K L Y E L E L E
K Q Q S S A K N E E K S E I G W G H Q I R S Y V L A P Y Q Q V K D A R S N T A Y S N V E A I L D G D I D A I L E G V L I
A K A

>PROKKA_00624 Molybdopterin molybdenumtransferase

M I S F K E A L K I H S S I P L K P L E I E V I S L F E S I G R I L A E D I I C T H A L P K F N Q S A M D G Y G F K M Q
D L G Q K T Q V I Q H I F A G D D V S A L E V R E N E C V K I M T G A M V P K G I E T I I P I E C M L E S H E N S A L A
P K D F K I N A N I R Q K G E N A S L N S V L V P K N T R L N Y G H I A L I A S Q G L K E I K A F R K L K I A L F S S G
D E L V P L G Q N A L E C Q V Y D V N S V G I F N M L K N Y N T H F L G V L K D D K N L Q L K I L E S Q D Y D V I L S S
A G V S V G D K D F F K D A L K E K N A L F Y Y E K V N L K P G K P V T L A R L N Q S I I I G L P G N P L S C L L V L R
V L I L P L L E R L S L N K D F K L K P F K A Q I N A P L K L K G E R V H L I L G N Y S N H Q F I P Y N N N R Y D S G A
I Q A L A R V D S I A L I D E G V G L V Q G E I E I L R F E N

>PROKKA_00625 Flagellar biosynthetic protein FliR

M L D F I Q E L S T P H V R D F F L L F L R V S G V L S F F P F F E N H L V P L S V R G A L S L Y V S A I F Y P T L E F
S N A A Y T P E G F I I A C L C E L F L G V C A S I F L Q I V F A S L V F A T D S I S F S M G L T M A S A Y D P I S G S
Q K P I V G Q A L L L L A I L I L L D L S F H H Q I I L F V D H S L K A V P L G Q F V F E P A L A K N I V K A F S H L F
V I G F S M A F P I L C L V L L S D I I F G M I M K T H P Q F N L L A I G F P V K I A I G F V G I I L I A S A I M G R F
K E E I S L A F S A I S K I F

>PROKKA_00626 hypothetical protein

M V L F L S I F K K S F N D F L S A R M L L I N L G P I L L S L A F F G A V F Y Y N G G S I V G Y C Q T L L P Q S L N D
Y S H S Q G F F A G V F A W V F K A L V Y F L I F W I V I L L S L V I N I F A S I F Y T P L V V S Y L H Q K Y Y P H V V
L E E F G S I L F S I K Y F L K S L A F M L L F L A V L T P F Y F I P F I G V F G V F F S I I P H F L F F K N T M S W D
I A S M I F N Y Q S Y Q N L L K Q H R L K H Y R F S F F C Y L F S L I P F F N F F A T L L Q T L M L T H Y F F I L K E K
E C

>PROKKA_00627 Putative peptidyl-prolyl cis-trans isomerase Cbf2 precursor

M K K N I L N L A L V G A L S A S F L M A K P A H N A N N A A H N T K K T T D S S A G V L A T V D G R P I T K S D F D M
I K Q R N P N F D F D K L K E K E K E A L I D Q A I R T A L V E N E A K T E K L D S T P E F K A M M E A V K K Q A L V E
F W A K K Q A E E V K K I Q I P E K E M Q D F Y N A N K D Q L F V K Q E A H A R H I L V K T E D E A K R I I S E I D K Q
P K T K K E A K F I E L A N R D T I D P N S K N A Q N G G D L G K F Q K N Q M A L D F S K A A F A L T P G N Y T K T P V
K T E F G Y H I I Y L I S K D S P V T Y T Y E Q A K P T I K G M L Q E K L F Q E R M N Q R I E E L R K H A K I V I N K

>PROKKA_00628 Fructose-bisphosphate aldolase

M L V K G N E I L L K A H K E G Y G V G A F N F V N F E M L N A I F E A G N E E N S P L F I Q A S E G A I K Y M G I D M
A V G M V K T M C E R Y P H I P V A L H L D H G T T F E S C E K A V K A G F T S V M I D A S H H A F E E N L E L T S K V
V K M A H N A G V S V E A E L G R L M G I E D N I S V D E K D A V L V N P K E A E Q F V K E S Q V D Y L A P A I G T S H
G A F K F K G E P K L D F E R L Q E V K R L T N I P L V L H G A S A I P D D V R K S Y L D A G G D L K G S K G V P F E F
L Q E S I K G G I N K V N T D T D L R I A F I A E V R R V A N E D K S Q F D L R K F F S P A Q L A L K N V V K E R M K L
L G S A N K I

>PROKKA_00629 Elongation factor P

M A I G M S E L K K G L K I E L G G V P Y R I V E Y Q H V K P G K G A A F V R A K I K S F L D G K V I E K T F H A G D K

CEEPNLVEKTMQYLYHDGDTYQFMDIESYEQIALNDSQVGEASKWMLDGMQVQVLLHNDK
AISVDVPQVVALKIVETAPNFKGDTSSASKKPATLETGAVVQVPFHVLEGEVIKVNTE
EYLEKVK

>PROKKA_00630 Pseudaminic acid synthase

MLECPKIVAELSANHNQDLNLAKESLYAIKESGADFKVLQTYTPSCMTLNSKEDPFIIQG
TLWDKENLYELYQKASTPLEWHAELFELARKLDLGIFSSPFSSKALELLESLNCPMYKIA
SFEIVDLDLIEKAARTQKPIILSSGIATHTELQDAISLCKEVGNFDITLLKCVSAYPSQI
EDANLLSMVKLGETFGVKFGLSDHTIGSLCPILATTLGASMIEKHFILNKSLQTPDSAFS
MDFNEFKSMVEAIKQSVLALGEEEPKINPKTLEKRRFFMRSLFVIKDIQKGEALTSDNIK
ALRPNLGLHPKFYKEILGQKASKFLKANTPLSADDIERSL

>PROKKA_00631 Lipoprotein-releasing system ATP-binding protein LolD

MIKAINISHAFKPLYNGVNLRIKPKESLAILGVSGSGKSTLLSHLATMLKPDSTISLL
EHQDIYALNSKKLLELRRLKVGIIIFQSHYLFKGFSALENLQVASILAKQEINHSLLEQLG
IAHTLKQGVGELSGGQQQLRSIARVLSKKPKIIIADEPTGNLDTTSANQVISMLQNYIAE
NEGALVLATHDEHLAFTCSQVYRLEKEVLIKEK

>PROKKA_00632 apolipoprotein N-acyltransferase

MFVSSVYGNVAVLDAYAIENPYISIIILTSLLAPLSMLAFLKTPRNSAFALGFFVVGALLFYW
CALSFYSDFTYLLPLIIILIALVYGVLFYLLLYFENPYFRLLSFLGSSFIHPFGFDWL
PDSFFSYSVFRVDKLSLGLVFLACIFLSAQNLKKYRIIGVLLLLGALDFNGFKTSDLKKV
GNIELVSTKTPQDVKFDSNYLNDIENNILKEIKLAQSKQKTLIVFPETAYPIALENSPFK
AKLEDLSGNIAIILIGTLRTQGYSLYNSSFLFSKESVQIADKVILAPFGETMPLPKFLQKP
LEKLFFGESAYLYRNASNFSDFTLNDFTFRPLICYEGTSKAAYSNSPSKVFIILMSNNAWF
SPSIEPTLQRTLLKYYARRYDKIILHSANFSTSIIILNPSLLGDIFSGNDHD

>PROKKA_00633 Colicin V production protein

MNYIDLALLVVVAVFGIRGFYHGFVSEIAGTLGIVLGVYLASRYSVAVGNLFSEHLYDLR
NETMTNLIGFLLVLASIWVFFLAFGVLLGKVLVFSGLGIIDKALGFIFSCLKTFLVLSFI
LYALSKMEVMKDANAYLQEKSTFFSTMKSVASKIMRLDGVKHVEQNLKDNLEEMSDEVKN
KESFNKNKQSFNKAMDKGMESELKEKAKDLPKNMLDPKANQTPNPPTPSNKEPL

>PROKKA_00634 Lysine--tRNA ligase

MFSNQYIQQRIHKANSLREEGKNPYQNGLKRSLTNAAFLEKYAYVKGLEEPKDKEKCESI
VGRVKLLRLMGKACFIKIEDESAILQAYVSQNELNDEFKSLKKHLEVGDIVLVKGFPFAT
KTGELSVALEFPHILSKTIVPLPEKFHGLSDIELRYRQRYLDLIVNPSVKDVFKKRS
SSVRKFFEMKGFLEVETPMMHPIPGGANARPFITYHNALEVERYLRIAPELYLKRLIVGG
FEAVFEINRNRNEGMDHSHNPEFTMIEFYWAYHTYEDLIELSKRLFDYLLKTLNLD
SKI IYNDMEVDFNQTSVISYLDALLETIGGISKDILEKEDRLAYLLEQGIKVEPNLT
YGKLLAEAFDHFVEHQLINPTFVTQYPIEISPLARRNDSNPNIADRFELFIAGKEIANG
FSELNDPLDQLERFKNQVAEKEKGDEEAQYMDDEDYVWALAHGMPPTAGQGIGIDRLV
MLLTGAKSIKDVLFPAMRPVKNDFNIESGE

>PROKKA_00635 Serine hydroxymethyltransferase

MAYFLEQTDSEIFELIFEYKQNEHLEMIASENYTFPSVMEAMGSILTNKYAEGYPNKR
YGGGCEVVDKIESLAIERAKKLFNCQFANVQAHSGSQANNAVYHALLKPYDKILGMDLSC
GGHLTHGAKVSLTGKHYQSFSYGVNLDGYIDYEEALKIAQIVKPEIIVCGFSAYPREIDF
KKFREIADEVGALLLGDIAHVAGLVVANEHAHPFPHCHVVSSTTHKTLRGPRGGLILTND
EEIAAKIDKAIFPGTQGGPLMHAIAAKAVGFKENLKPEFKAYAKLVKSNMQVLAKALKEK
NHKLVSOGGTSNHLMLDFLDKPYSGKDADIALGNAGITVNKNTIPGETRSPFVTSGIRIG
SAALSARGMGAKEFEIIGNKISDILNDINNVSLLQHLVKEELKAMANQFPVYHQPIF

>PROKKA_00636 hypothetical protein

MTEMELKLIKIDTSHYFEKKPGLGERMDYAGRCYKNKFQRVNAMLTSLSLIQHLKREIEI
AHNLILRNDKVENIVFDYNGRNPFRFYHKAQLLLREEGFMNFTAYNTKTPGHLHLYVHK
GHTELGGERLVKTLMSKLAQGLPKWKVFPSPNEWPKFENILALPYEVFAKERGSSWAKHL

>PROKKA_00637 cell division protein ZipA

MSEKERLNEVILEEENNGGGTKKVFLIVAIIAIILAVLLMVFWKSTRVAPKETFLQTD
SGMQKIGNTKDEKKDDEFESLNLDPSKQEDKLDKVADNVKKQENDAFNMPIQTDQTQ
TEMKTTEETQEAKKELKAVEHTSTQKESQAVAKKETPHKKPKATPKDKQAHKDKHAAKEPKAKKE
AHKEVPKKANSKTTLTGHYLQVGVFVHTPNKAFLQAFNQFPHKIEDRGATKRYLIGPYK
SKQEALMHADEVSKMTKPVVIAR

>PROKKA_00638 hypothetical protein

MVVAKNEYNKKLYDIIIDGQQRTTTFMLLHVLANKQNEKDKQETRKYLYQKGELKLEVAP

QNSFFKALLEAAEKGNI SHCEKDADTEGKQNLFEVLKAILDKVSKLSEEEANERLEALL
EMVLMWFEEPNPGRAIRTFQSVNDRGVPLLLLDKLSLLIYYSNIFCDGKRRLDQFINDH
FGEIFKIFAKIKSNHISSVGGSDGDI FRYHAGSQRFDGIEFLGHYEASTDKTYEKLKD
ELKEIKKSKLSFQSYVSDLKNFYQAFDLDLSEIDTNPTTFKVMLINRINPHFFNSLIR
LKINNELDDETLKLFKTDIVLFKTGRNMRSAAYNLISAYLKKGKEGLKSEMIAQCRNDI
GLASLKLVNNAFNLSCFHYVFFEKNQEMGLADLKKLIPGKQFSQEKEHIIPINLLELDN
EIEIQKLGFEDEKDLNENYIDTYGNLISLESPLNRNASDKDLYEKDEIYKSSEVPFNRRFN
VKGFNKKVLIERNNAMREWLDITFFKDFAAH

>PROKKA_00639 hypothetical protein

MKTTIEEIFQKEGYNIPNYQRDYAWKDKNFRDL

>PROKKA_00640 hypothetical protein

MLEKLIERVLFATRLLAPLCIAMSLVLVVLGYVFMKELWHMLSHLDTISETDLVLSALG
LVDLLFMAGLVLMVLLASYESFVSKLDKVDASEITWLKHTDFNALKLKVSL SIVAI SAIF
LLKRYMSLEDVLSIPKDTPLSHNPIFWQVVIHLVFVCSALLAAVTNNIAFSQNKGH

>PROKKA_00641 Major cardiolipin synthase ClsA

MKIFLVLLSVFFFNGCFGLVYKTPISSPPI SYDPYTTTIGSLYAKNLKENPNHSAAILLE
DGFDALLHRVGLIRMSQKSIDMQTYIYKNDLSSQVIAKELLNAANRGVKVRILLDDNGLD
SDFSIMLLNFHKNIEVKIFNPYYIRNKGLRYFEMLADYERIKKRMHNKLFIVDNFAVII
GGRNIGDNYFDNDLDTNFDLDAFFGGVASKAKESFERYWRFHRSIPVSLLRTHKRLKN
NAKEIAKLHEKIPISAEDKNQFEKKVNDFIDRFQKYQYPIIYGNAI FLADSPKKIDTPLY
SPIKIAFEKALKNAKDSVFIASSYFIPGKKMMKIFKNQISKGIELNILTNSLSSTDAIVV
YGAWERYRNQLVRMGANVYEIRNDFFNRIKGRFSTKHSLSHGKTIVFDDNLTLLGSFNID
PR SAYINTESA VLFDNPSFAKRVRLSLKDHAQQAWHLV VYRHRVIWEATEEGILIHEKTS
PDTSFFLR LIKEWSKVLPEREL

>PROKKA_00642 Fumarate reductase iron-sulfur subunit

MSDNERTIVIRVLKFD PQSAVSKPHFKEYQLKETPSMTLFIALNLIREHQDPDLSFDFVC
RAGICGSCAMMVNGRPLACKTLTSSFESGVITLMPMPSFTLIKDL SVNTGDWFLDMTKR
VESWAHSKEEVDITKPEKRIEPDEAQEVFELDR CIECGCCIASCGTKLMRPNFIGAAGMN
RAMRF MIDSHDERSDDDFYELVGDDDG VFGCMSLIACHDTC PKELPLQSSIATLRNRMKL
VGKSR

>PROKKA_00643 Fumarate reductase flavoprotein subunit

MKITYCDALIIGGGLAGLRASIACKQKGLNTIVLSLVPVRRSHSAAAQGGMQASLANAKK
SEGDNEDLHFLD TVKGS DWGCDQQVARMFVTTAPKAIRELASWGV PWTRIKK GDRPAVVN
GEHVTITERDDRHYILSRDFGGTKKWRTCTADATGHTMLYAVANEALH HKVDIQDRKD
MLAFIHHNNKCYGAVVRDLITGEISAYVSKGTL LATGGYGRVYKHTTNAVICDGAGAASA
LETGVAKLGNMEAVQFHPTALVPSGILMTEGCRGDGGVLRDKFGRRFMPAYEPEKKELAS
RDVVSRRILEHIQKGYGAKSPYGDHVLWDIAILGRNHVEKNLRDVRDIAMTFAGIDPADS
EEQTKDNMQGMPANEPEY GQAMAKQKGWIPKPMQHYSMGGVRTNP KGETHLKGLFCAGE
AACWDLHG FNLGGNSVSEAVVAGMIIGDYFASHCLEAQIEINTQKVEAFIKESQDYMHF
LLHNEGKEDVYEIRERMKEVMDEKVGVFREGKRLEEALKE LQELYARSKNICVKNKVLHN
NPELEDAYRTKKMLKLALCITQGALLRTERSGAHTRIDYPKRDDEKWLNR TLASWPSAEQ
DMPTIEYEELDVMKMEISPDFRGYGGKGNFIPHPKKEERDAEILKTILELEKLGKDRIEV
QH ALMPFELQEKYKARNMRLEDEEVRARGEHLYSFNVHDLLDQHNANLKGEHHE

>PROKKA_00644 Fumarate reductase cytochrome b subunit

MQQEEIIEGYGASKGLKKSGIYAKLDFLQSATGLILALFMIAHMLVSSILISDEAMYK
VAKFFEGSLFLKAGEPAIVSVVAAGVILILVAHAFLALRKFPIN YRQYKVFKTHKHLMKH
GDTSLWFIQALTGFAMFFLASIHLFVMLTEPESIGPHGSSYRFVTQNF WLLYIFLLFAVE
LHGSIGLYRLAIKWGWFKNVS IQGLRKVKWAMSVFFIVLGLCTY GAYIKKGLENKDNGIK
TMQEAIEADGKFHKE

>PROKKA_00645 Triosephosphate isomerase

MTKIAMANFKSAMP IFKSHAYLKELEKTLKPQHFD RVFVFPDFLGLLPNSFLHFTLGAQN
AYPKDCGAFTGEITSKHLEELKIHTLLIGH SERRALLKESPSFLKEKFDFFKDNFKIVY
CIGEDLT TREKGFKAVKEFLSEQLETIDL NYSNLIVAYEPIWAIGTKKSASLEDIYLTHG
FLKQILNQKTPLLYGGSVNIQNAKEILGIDSVDGLLIGSASLELENFKTIISFL

>PROKKA_00646 Enoyl-[acyl-carrier-protein] reductase [NADH] FabI

MGFLKGKKGLIVGVANNKSIAYGIAQSCFNQ GATLAFTYLNESLEKRV RPIAQELNSPYV
YELDVSKEEHFKVLYNSVKDLGSLDFIVHSVAFAPKEALEGS LLETSSAFNTAMEISV
YSLIELTNTLKP LLNGASVLTLSYLGSTKYMAHYNVMGLAKAALES AVRYLAVDLGKHN

IRVNALSAGPIRTLASSGIADFRMILKWNEINAPLRKNVSLEEVEGNAGMYLLSSLSSGVSGEVHFDVAGYHVMGMGAVEEKDNKATLLWDLQNNKGY

>PROKKA_00647 UDP-3-O-acylglucosamine N-acyltransferase
MKLSELLSAYSIETEFSNDFEVHALAELDKATPNDISYIDQARYLKLLKDSKAGAVFIRK
KESSKVPKRMQALIVDNPHLAFAKASHAFKIPFFKNPQSVSEPKHFVKVTIMPVNMIGEG
VEIGENSLIYPGVVIADGVKIGKNCVLYPRVILYQNTILEDNVTIHAGSVIGGDGFGYAH
TALGEHVKIEHVIGIVRIQKNVEIGANTAIDRAVFGETLIKEGVKIDNLVQIGHNCVLGEH
SIVVSQVGLSGSTTTGRNVVFGGQVVGIGGHLHVGEFTQIGGKSAVGKDLPPNTNFAGAIP
AMEIHEWHHFLAHLRTNFRKQKQTSLLQKAKGFFKS

>PROKKA_00648 S-adenosylmethionine synthase
MKDSFLFTSESVTEGHPDKMADQISDAVLDYIIERDQKAKVACETLVSNFGCMITGELKT
SVYAPMQEIAREVVKIGYTDALYGFDYRSAAVLNGVGEQSPDINQGVREDGEIGAGDQ
GLMFGYACKETETLMPLPIHLAHQLAFALAQKRKDNTLPFLRPDGKSQVSVRYENNKPV
VDTIVISTQHSPEVSQKHLKEAVIEEIVYKVLPEYLHDNIKFFINPTGKFVIGGPQGD
GLTGRKIIVDTYGGFCPHGGGAFSGKDPKSVDRSAAYAARYVAKNLVASGVCDKATVQLA
YAIGVIEPVSIYVNTHTSKHSSAELEKCVKSVFKLTPKGIIESLDLLRPIYSLTSAYGH
FGRELEEFTWEKTNKVEEIKAFFKR

>PROKKA_00649 Nucleoside diphosphate kinase
MKQRTLSIIKPDALKKKVVGKIIDRFESNGLEVAMKRLHLSVKDAENFYAIHKERPFFK
DLIEFMVSGPVVVMVLEGEDAVAKNRDLMGATDPKLAQKGTIRADFAESIDANAVHGSDS
LENARNEIAFFFAARDL

>PROKKA_00650 hypothetical protein
MQFEMRKIAFNAPKAFSLEHEGVVLEGEVVRVGAKLFRLEACLKGEIMLVCDTSGKEFKK
SLDESLVLHISDGLWDTQSQSLDFDNLDVIESFNGFIDLSEILRSEVESIKLDYHYAD

>PROKKA_00651 50S ribosomal protein L32
MAVPDRRVSKTRAAKRRTHYSVKLAKPVKAKDGTWKLPHHINKFTKEY

>PROKKA_00652 Phosphate acyltransferase
MKIVIDLMGADHGVLPPIIEGVSRALENKSFSAVLVGDKDKATPFIKELASKVEMIHTQD
YIKMEEAATEAIKRKESSIIYLGMDILKNGADALISAGHSGATMGLATLRLGRIKGV
ICTLMPSVGVKRPVLLDAGANTDCKPEYLIDFALMGYEYAKSVLHYDSPKVGLLSNGEED
IKGNMLVKETHKMLKAYDFFYGNVEGSDIFKGVVDVVVCDGFMGNVVLKTTEGVASAIG
IFKDEIKSSFSKMGALMLKNAFGILKQKTDYAEYGGAPLLGVNKSVIISHGKSNARAIE
CAIYQAISAVESQVCLRITKAFESLKPSVSQSDQQDA

>PROKKA_00653 3-oxoacyl-[acyl-carrier-protein] synthase 3
MEFYASLKSIAMHVPSEKNAEFQQFLDTSDEWIEKRTGIKERRFANDEEKSSDLGVIA
AKQAIERAHLTPKDIDLVVVATLSPDFLAMPSTACVLSAKLGIENKPAFDISAACGFIY
LLSVAKAYVESGMCENVLIVGAECTSSVLDFKDRGTCILFGDGAGACVIGRTKRLKESIL
DVQISANGNFSNYLYTPRTLKPTPFNAKEEALPEFLRMKGNEVFKLAVKTLLKDVEMILE
KNALKPEDVRLFIHQANFRIIQAVREHLDFKDEQVVLTVHKYGNNTSAASIPMAMCEAYE
EGRLLKKGDLMLLLDAFGGGLTWGSALVYFGG

>PROKKA_00654 hypothetical protein
MKKIIIFLFLMVLGGLKSQSSYCSDFCEGTPDSRIPPMGFHFSFVHSVKYYLQDPQERD
LKKCHQAFDSTLKVNFITKSFKKDKHAQMALEQAQKGT

>PROKKA_00655 hypothetical protein
MFKKMCLSLLMISGVCVGAKDLDFKLDYRATGGKLMGKMTDSSLLSITSMNDEPVVIK
NLIVNRGNSCEATKKVEPKLGDKFKKEKLFDELKYSQQIFYRLNCKPNQLLEVKIITDKGE
YYHKFPQ

>PROKKA_00656 hypothetical protein
MKFYKRVLKLHHFRNLGRNLPALLNSSFKEKHGGLVILVGENNVGKSNVLEALKIFNDA
DVKLCSEKDYFKAHESEDAVLNLEEETILDHKTIGFSCVDLKIQTKEVSEGLKELSKILI
SYFPCFVFGSFINLIMS YGVLDSFLKSYKEKLLSSFSTRQANHLLFKKLIKHLSGNNEL
VKNFCQCIREII EYNAPNKEYKPNQFFIMGKGKQKQLSEIYSRLKRLNERKIEPKDVKS
VLKDMPPQKVKTLD EIFKTTDFTKFTPETEVKDIKEIDEKYPINENFKQQFRTRLNIGNL
KKKIKNSLKYLEKTREDFERKKERLIREIENDCKNQKVLKFNYDVLLDNIQQICKYIAS
HAVNDISKDIKSMCMQLYLKKIDLLVNSEIEQYRYNDFLESARKFLWEDIKTLDEKSGVH
LFPKNIGEIKDKFEANKEKVQSKNYFEFAEYCRECNPYTAFQNLNRNKVQFPLSGGLSYK
SYKLVPMTKEYKELKITDNDLKTALFTLFDYSSLSSEFNQSDWFFRNSLFRKMDFHPNTIW
NFFGSILNYKNFDQKNLKYEQALQIMFDKNNDLEIYSIYNKSFVIPKKYLQEIDQESLKE

IRQSEFFPNIEVKGECDNNVCQLEFFNNKSSLLFKVNFTEILENLAEILEYMQLKIDSL
ITKEFNRLLAIAEDSSQDSYQLKIHVRHNNKLSKEDKEYTAYEIKLEIHDCRKSDNQNEPI
ILSQQSTGFQWAFNFMFGFLYNWGSFSLNKNIIYVMDEPATHLSVPAKKEFRRFLKEYA
HKNHVTFVLATHDPFLVDTHLDEIRIVEKETEGSVIKNSFNYPNNAGKSDALDKIKR
SLGVGQHVFNHPQKHRIVFVEGITTDCYLSAFKLYLCYKEYKDNPIPTFLPISGLKNNP
NDMKETIQKLCELDNNHIVLTDDDRKDNNDPQRAKSEEFKNANEEMHDLIRILQLSKCDE
HFKQIENCFSVNDKKEYAKNKRMEALAMAFKTRLLYSEKDNVMSEETKKNFLKLFEWIKKE
CNNPTIKKEYIKFDYNTQPIL

>PROKKA_00657 Septum site-determining protein MinD

MLTQEDVLNALKTIIPNFEDIVSFGFVKNIALHDDQLGLLIEIPSSSEETSEILRGKI
SEVVQKIGVKALNLDIKTPPKPQAPKPTTKNLAKNIKHVVMISSGKGVGKSTTSVNLISI
ALANLNQKVGLLDADVGPNI PRMMGLQNADVIMDPSGKKLIPLKAFGVSVMSMGLLYDE
GQSLIWRGPMMLRAIEQMLSDIIWGDLDVLVVDMPPGTGDAQLTLAQAVPLSAGITVTTT
QIVSLDDAKRSLDMFKKLHIPIAGIVENMGSFVCEHCKKESEIFGSNSMSELLEAYHTQI
LAKLPLEPKVRLGGDRGEPDIVISHPNSVSAKIFEKMAQDLAFLERVEKEKLADNKDIQP
TQTHACSH

>PROKKA_00658 General stress protein A

MQEVIPVVVAFDNNYCIPAGVSLFSMLANAKRERERVKLFYKIHCLVEDLSTENIAKL
EETIAPFSTFSSIEFLDISNEKLEPRHNYRKLDPLIASEIKKLYLKLNAFSQKRFSKMIM
CRFFASFSLFPQYDKMIMFDVDTLFDVNDISESFFIPLETHYFGAVREKDLVAMDRNSAKDL
YELRQMHAKNIGVADAFPNLEEAQILFDNYFNAGFLALNLKSWREENLENQLIAFFILKN
EKLLFNDQDALCFVCRGRILELPYSYNAHPSFLDTPSFPSIKEARMLHFWGDKPWKLFVS
IGAKKWHEALIQTFFKDAYFNAPFLDHLFESLQNRDKEIHALNKILSFSDKWHSFESLLP
RLSSKLLMEFLLFKAKQKVKRLIKRVF

>PROKKA_00659 Putative outer membrane protein

MVLNQKVWLGLLALHGVFLNAFEYQISARVGSFSRIALNQSIINSKKGITYPTGSYVTTTG
ALQIDSSLLPKGIEDHKLGFVGGEIGSLAYDSTKFLMDEANPKAGFQPANWYYMGRWEG
YLMQHSQNWTREQKAQNARPYVLNLYLDYQYKIDIFGIKLGRYPSKALFLSGFNQGFV
YRWKKFRIVWFSTFGRALANEQYIRDFYAPVNYKQKINYGHNFNLIYENKYIRIAPFIW
FYPKNFNAPGFEITHDTKSYWKSLEWRIQTTFYAWFPLYSDYLSKDYRATLVGKKSASLF
VFQVRNFRSYRFGWSVYKNFNGNSAQLGWNGSPVDPFYDTKDDTPYEDAYSNFYNANSIT
INAFVGKSVKNLLVQLYGKLTYSRADAQSLGVTFKYNLKKHIYFMLMVNGYQITMHKGY
KVGFFTSGYNPDFAQTIQNRSYLMSSMSYRF

>PROKKA_00660 Chaperone protein HtpG

MSNQEYTFQTEINQLDLMIHSLYSNKEIFLRELISNASDALDKLNYLMLTDEKLKGLNT
TPSIYLSFDSQKKTLMIKDNGIGMDKNDLIEHLGTIAKSGTKSFLSALSGDKKKDSALIG
QFGVGIFYSAFMVASKIVVQTKKVTSQAYAWVSDGKGKFEISECIKEEQGTEITLFLKDE
DSHFASRWEIDGIVKKYSEHIPFPIFLTITDTKFEGEGDHKKEVKEEKCDQINQASALWK
MNKSELKDKDYKDFYQSFADHNSEPLSYIHKNVEGSLEYTTLFYIPSKAPFDLFRVDYKS
GVKLYVKRVFITDDDKELLPSYLRVFKGVIDSEDLPLNVSREILQONKILANIRSASVKK
ILSEIERLSKDEKNYHKFYEPFGKVLKEGLYGDFFENKEKLELLRFYSKDKEKLISLKEY
KENLKENQKSIYYLLGENLDLLKASPLLEKYAQKGYDVLLLSDEIDAFVMPGVNEYDKTP
FRDASHSESLKELGLEEIHDEVKDQFKDLMAFEENLKDEIRGVELSSHLTSAVALIGDE
QNAMMANFMRMQGSVPESKKTLELNPNHAILQKLLKCEDKEQLSAFIWLLYDGAKLLEK
GALKDAKSFNERLNIVLLKAL

>PROKKA_00661 Beta-lactamase HcpA precursor

MLGSVKKTLFGVLCIGALCLRGLMAEPDAKELVNLGIESAKKQDFAQAKTHFEKACELKN
GFGCVFLGAFYEEGKGVGKDLKKAIQFYTKGCELDNGYGCNLLGNLYYNGQGVSKDAKKA
SQYYSKACDLNHAEGCMVLGSLHHYGVGTHKDLRKALDLYEKACDLKDSPGCINAGYIYS
ITKNFKEAIVRYSKACELKDGRGCYNLGVMQYNAQGTAKDEKQAVENFKKGCKSSVKEAC
DALKELKIEL

>PROKKA_00662 Succinyl-diaminopimelate desuccinylase

MNALEITQKLISYPTITPKECGIFEYIKSLFPTFKTLECGENGKVLNLFYRIFNPPKDHA
EEKHAKENTKPLHFCFAGHIDVPPGNHWQSDPFPKPVKEGFLYGRGAQDMKGGVGAFLS
ASLNFNPKTPFLLSILLTSDEEGPGIFGTKLMLEKLKEKDLLPHMAIVAEPTECEKVLGNS
IKIGRRGSINGKLILKGIQGHVAYPQKQCNPIDTLASVLPLISGVHLDNGDECFDPSKLV
ITNLHAGLGANNVTPGSVEITFNARHSLKTTKESLKEYLEKVLKDLPHTELESSSSPFI
TASHSKLTSVLKENILKTCHTTPLLNTKGGTSDARFFSAHGIEVVEFGVINDRIHAVDER

VSLKELKLLEKVFLGVLEGLSEA

>PROKKA_00663 tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG

MVKESDILVIGGGHAGIEASLIAAKMGARVHLITMLIDTIGLASCNPAIGGLGKGHLTKE
VDVLGGAMGIITDNSGLQYRVLNASKGPAVRGXRAQIDMDTYRIFARNLVLNTPNLSVSQ
EMTESLILENDEVVGVTNINNTYRAKKVITTTGTFLKGVVHIGEHQNQNGRFGENASNS
LALNLRRELGFKVERLKTGTCTPRVAGNSIDFEGLEEFGDTNPPYFSYTKDFNPTQLSCF
ITYTNPITHQIIRDNFHRAPLFSGQIEGIGPRYCPSIEDKINRFSEKERHQFLFEPQTIH
KSEYYINGLSTSLPLDVQEKVIHSIKGLENALITRYGYAIEYDFIQPTELTHALETKKIK
GLYLAGQINGTTGYEEAAAQGLMAGINAVLALKNQAPFILKRNEAYIGVLIDDLVTKGTN
EPYRMFTSRAEYRLLLREDNTLFRLLGEHAYRLGLMEQDFYKELKKDQQAIQENLKRKEY
VLTPSKEVLKRLSELDENPINDKVDGVSLLDARDSFNAEKMRSSFFSLTPLNERVLEQIKI
ECKYNIYIEKQHENIAKMDSMLKVSIPKDFVFKGIPGLSLEAVEKLEKFRPKSLFEASEI
SGITPANLDVLHLYIHLRKN

>PROKKA_00664 Sodium-dependent dicarboxylate transporter SdcS

MENHSHANTHTDTRIDDKSTKIVRLGLIGGALIALIIYYALNAQMPHIVEEIPKLNLSN
YKAMPVVAGVAVLMGIWWMTEAIDLPTALLPLVLFVSFVSVDQFSSVSSSYASPIIFLFM
GGFILALSMQKWNLHTRIALSIILLVGTSPPRLILGFMIATGFLSMWVSNTATAVMMFXV
GMSVLQLVAKLVGKENASNAFYQKEEITKAHGGIMGNIVHKGKDIAQVIQEKTTIYRTNF
SICLMLGIAAYAASIGSLGTLTGTPPNALLAGYMKTSFNIEIDFAQWMVFGTPLAFIMLIL
AWLLLTYAIFPLKIKEIPGGKEVVKLKLNLGRLSQAEISVGVI FILASLWIFLDTILK
SWGIKIDKIDSVIAMGVSVLLFILPANNQGDRLIDWGVAKKLPWDVLLLFGGGLALSAQF
SKTGLSLWIGHLVSGFSLPILFIIVMVTLMVIFLTEITSNTATAAAFLPVIGGVAMGMG
YENHQSLLLTIPVALSATCAFMLPVATPPNAIAYGSGYVKITDMIKAGLWLNLVGVVLIS
AFSYFLVSLIFN

>PROKKA_00665 Phosphatidate cytidyltransferase

MKEGLFKEKSRITGFVLIIIVAGLILYADNLLLFWAVLGGIYAVGFSEALRLFQVKASFS
LYFILVLSWVAAYFNRPICALISAMVMASVIAYQKAHSEAILPFLYPGVGFFALFGV
YKDFGAVAIIWLLVVVASDVGAFFGGKLLGKTPFTPTSPNKTLEGALIGVVLASVLGSF
VGMGKLSGGFLMALLFSFLIALIAVFGDLYESYLKRRVGIKDSGKILPGHGGVLDRLDSM
LFGALGLHALLYFLEVWKETAVFLGD

>PROKKA_00666 1-deoxy-D-xylulose 5-phosphate reductoisomerase

MVVLGSTGSIGKNALKIAKKFGVEIEALSCGKNIALINEQIKVFKPKKVAILDPNLDLYL
EPLGAKVFVGLEGIDAMIEECVSNLVINAIVGVAGLKASFSLQRNKKLALANKESLVSA
GHLLDISQITPIDSEHFGWLWALLQNKTLPKPSLIISASGGAFRDTPLESIPIQNAQNALK
HPNWSMGSKITIDSASMVNKLFEILETYWLFGLSKIDALIERSSIVHALVEFEDNSIIA
HLASADMQLPISYAIDPKLASLSASIKPLDLYALSAIKFEPISMERYTLWRYKDLLLENP
KLGVLNLASNEVAMEKFLNQEIAFGGLIQTISQALESYAKMPFKLSSLEEVLAALDKEVRD
RFKNYR

>PROKKA_00667 putative kinase inhibitor protein

MKTFEVMIQTDSQGYLDAKFGGNAPKAFLNPNGLPTYSPKISWQKVEGAQSYALELIDHD
AQKVCGMFPVHVWVGNIHNVLEENASMMDKRITQGVNSLTQGFIRSPLNESEKQRSNLN
NSVYIGPMPPNGDHHYLIQVYALDIPKLALKAPFFLGDLHDKMRNHIIAIGRKEFLYKQF
VRK

>PROKKA_00668 hypothetical protein

MESRINRLSAKIDALLEQQKRVISLLETYSVYSTPQEASNKLFGKISQGMELAPEIAQE
DEEKRLYLQYLSHVDITKNKQDSQLKKDCLEFIQRFNIPKPIIITALYNLKGIKPTKKE
VAKQLQKLYVWERRYQQGGIDALKDRRGRPLKKP

>PROKKA_00669 Cysteine desulfurase

MLQRIYLDNNATTRIDPKVKEIMDPFLRDHYGNPSSLHQFGTETHPAIAEALDKLYKGIN
ARDIDDVIITSCATESNNWVLKGVFDECLKKGKNHIVTTVAEHPAVRSTCNFLES LGVE
VTYLPINEHGSITAEQVKEAITEKTALVSVMWANNETGLIFPIEEIGAICKEKGVLFHTD
AVQAIGKIPVDVLKANADFLSFSAHKFHGPKGIGGLYIRSGVGLTPLFHGGEHMNGRRSG
TLNVPYIVGMGEAMRLAVEHLDYEKEVVGKLRDKLEEALLKIPDVMVVGDRIDHRVPNTTL
VSVRGIEGEAMLWDLNRSNIAASTGSACASEDLEANPVMVAIGASKELAHTAIRLSLSRF
NTEAEIDKTIEVFSQAIRLRNISSY

>PROKKA_00670 NifU-like protein

MAKHDLVGSVLWDAYSKEVQRRMDNPHTLGVITEEQAKAKDAKLIVADYGAEACGDAVRL
YWLVDDESTDTIIDAKFKSFGCGTAIASSDMMVELCLNKRVDQAVKITNLDVERGLRDDPD

TPAVPGQKMHC SVMAYDVIKKAAGMYLGKNAEDFEEEEIIVCECARVSLGTIKEVIKLN DL
KSVEEITNYTKAGAFCKSCVRPGGHEKRDYLLVDILKEVREEMEA EKLKVTANKSQNGEL
AFREMTMVQKIKAVDKVIDENIRAMLMMDGGDLEILDIKESDDYIDVYIRYMGACDGCMS
ATTGTLFAIENALQELLDRSIRVLPI
>PROKKA_00671 hypothetical protein
MKT TENTDETHLRETKNKLGRKPKADANKKTRAVSLYFSDEQYQKLEKMANEEEEESVGSY
IKRYILKALRKIEQNGA
>PROKKA_00672 hypothetical protein
MAKKTSLFECQHCGFTSPKWL GKCVQCNAWESFIELNQAQKEVLN TLKNPIPKAQKSVSI
AAIEHEEVIKFSSTQSELDIVLGGGI AKGGLYL VGGSPGXGKSTLL LKVASGLAKNQQKV
LYVSGEESLSQIKMRATRLDCIEKELYLLNEINWPVIKANIESENYFACVIDSIQTLYSP
EISSAPGSISQVREITFELMRLAKTRDIAIFIIGHITKEGSIAGPRVLEHMVDSVLYFEG
DPSRELRLIRSFKNRF GPTSEIGLFEMKEQGLVSAKEASSLFFSQEDPMEGSAITITLEG
SRALILEIQALVSECSFGAPKRLANGFDTNRLNMLIALLEKKLEIPLNRHDFVINVSGGI
KISEPACDLAVIASILSSFKNRKIDNKTAFLGEVSLNGRILEAPNLNARLKEMENYGLFK
AILPKKPSQKTSIKCYEANVVGKIVEWM
>PROKKA_00673 Peptide methionine sulfoxide reductase msrA/msrB
MQANENMGSQH QKTDERVIYLAGGCFWGLEAYMERIYGVIDASSGYANGKTSSTNYEQLH
ESDHAESVKVVYDPKKISLDKLLRYYFKVIDPVS VNKQGN DVGRQYRTGIYYVNSADKEV
IDNALKELQKEVKGKIAIEVEPLKNYVRAEEYHQDY LKKHPGGYCHIDLKKADEVIVDSD
KYTKPSDEVLKKKLTKLQYEVTQNKHTEKPFENEYYNKEEEGIYVDIT TGEPLFSSADKY
DSGCGWPSFSKPI NKDVVKYEDDES LNRRKRIEVL SRIGKAHLGHVFNDGPKELGGLRYCI
NSAALRFIPLKDMEKEGYGEFIPYIKKGELKKYIQDKKTH
>PROKKA_00674 Sulfite exporter TauE/Safe
MEESTAFILALVGLFTGITAGFFGIGGGEIVVPSAIFAHFSYSHAVGISLMQMLFSSVVG
SIINYKKGLLDLREGSFAALGGLMGAILGSFILKIIDDKILMAV FVVVVCYTFIKYAFSS
NKKPKHFEE MHFDLHANNKTPEKKRAIPFVSMDRTHGVLMLAGFVTGIFSIPLGMGGGIL
MVPFLGYFLKYDSKKIVPLGLFFVVFASLSGVISLYNGRVLDNISVQAGVITGIGAFLGV
GIGIKLIALANEKVHKILLLLIYALSILATLHKLIMG
>PROKKA_00675 Helicobacter outer membrane protein
MSVGYQIGEAVQKVNTGALQNLADRYDNLSNLLNQYNYLNSLVAQASTPSAITGAIDNL
SSSAINLTSATTTSPAYQ AVALALNAAVGMWQVIALFIGCGPGPTNNQSYQSFGNTPALN
GTTTTTCNQAYGTGPNGILSIDEYQKLNQAYQIIQTALNQNQGGGMPALNDTTKTGVVNIQ
QTNYRTTAQNNIIEHYT KNGKEIPTSYSGGASFSP TIQLEYNNNAEYLLQQAATIMQVL
ITQKPHVQTSNGGKAWGLSSTPGNVVDIFGPSFNAINEMIKNAQT ALEKTQQLNANENAQ
ITQPNNFNPYTSKNKQFAQEMLNRAEAQA EILGLAKQVADNFHSIQGP IQGDLEECKAGS
AGVITNNTWGS GCAFVKETLNSLEQHTAYYGNQVNQDRALAQTILNFKEALNTLGND SKA
INSGISHLPNAKSLQNMTQELGKNPFRRIGVIDYQNNNGAMNGIGVQVGYKQFFGKKRNW
GLRYYGFFDYNHAYIKSNFFNSASDVW TYGVGMDALYNFINDKNTN FLGKNNKLSVGLFG
GFALAGTSWLNSQQVNL TMMNGIYNANVSTSNFQFLFNLGLRMNLARPKKKDSDHAAQH G
IELGFKIPTINTSYYSCHGR
>PROKKA_00676 Bicarbonate transporter Bica
MFEKIQKEWLSNIQKDLLSGFVVGLSVIPETAGFAIMVGLDVGVAFYTTFYMAFVLSLFG
ARKAMISAAAGSVALILVG VVKNYGLE YAGVATLMAGILQILLGYLKIGNLLRFIPQSVM
YGFVNALGILLLMEQFKFLENQNLGVFVLLVIGILIIYLFPLITKKIPSNLICILIVSAI
ALIFDTHAPNLGSIEQGVSGFHFIIPKNLDFKIMVELLPYALS LALVGTIESLLTAKTL
DVILKDGVS DKNKETKAQGLGNIVSGLLGGMTGCALVGQSIINAKSGAKTRLSTFFAGFS
LMVLILVFNEYVVKIPIVAVVAVMVMISFTTFNFQSIINIKKIKLYDTLNMLLVAVVLY
THNLAIGVVVGVLVNALWIKSEGIA
>PROKKA_00678 Helicobacter outer membrane protein
MKKTILLSLMVSSLLAENDGVFMSVG YQIGEAVQQVKNTGEIQKVS NAYENLNLLTRYN
ELKQTASNTDSSTAQAIDNLKESASRLKTT PNSANQAVSSALSSAVAMWQVIASNLASGT
LPTSEYNKIN AISQLLQNTLENKNNELTIANDYNQLLTQASTIIISTLQSQCPGIDGGNGK
PWGINASGNACAI FGNTFNAITSMIDSAKKA AAEARRTAPEGPNQPSAFTNADFNKNLNQ
VSSVINDTISY LKGDNLATIYNTLQKTPDSKGFQSLVSRSSYSYSINETQYSQFQTTTKE
FGHNPF RSVGLINSQSNNGAMNGVGVQLGYKQFFGKNKFFGIRYYAFFDYNHAYIKSNFF
NSASNVFTY GAGSDLLNFINGGSDKNRKVSFGIFGGIALAGTTWLNNQSANLKITNSAY
SAKINNTNFQFLFNTGLRLQGIHHGVELGVKIPTINTNYYSFMGAKLAFRRLYSVYFNYV

LAY

>PROKKA_00679 3-deoxy-manno-octulosonate cytidyltransferase
MIIIPARLKSSRFENKVLENIFGLPMVIRCAKNANLVDECVVACDDDESIMQTCQKFRIKA
VLTSKHHNSGTERCLEAARILGLKNDERVLNLQGDEPFLEKEVILALLEATKNAPFMATC
AKVIDEEQAKSPNLVKVVLDNQNNALYFSRSLIPFLRDFDAKRQTPLLGHIGIYGFNKE
ILEELCALPKPCVLEEIEKLEQLRALYYQKKIAVKIVQSQSVGIDTQEDLENALKIFSPNL
LER

>PROKKA_00680 disulfide isomerase/thiol-disulfide oxidase
MILRASVLSALLLVGLGAAPKHSVSANDKRMQDNLVSVIEKQTNKKVRILEIKPLKSSQD
LKMVVIEDPDTKYNIPLVSKDGNLVIGLSNIFFSNKSEDEVQLVAETNQKVQALNATQQN
SAKLNAIFNEIPADYAIELPSTNAANKDKILYIVSDPMCPCQKELTKLRDHLKENTVRM
VVVGWLGVNSAKKAALIQEEMAKARARGASVEDKISILEKIYSTQYDINAQKEPEDLRTK
VENVTKKIFESGVIKGVPFLYHYKA

>PROKKA_00681 hypothetical protein
MKKPYRKISDYAIVGGLSALVMVSIVGCKSNADDPKEQSSLSQSVQKGAFVILEEQKDK
SYKVVEEYPSSKTHIIVRDLQGNERVLSNEEIQKLIKEEEAKIDNGTSKLVQPNNGGSNE
GSGFGLGSAILGSAAGAILGSYIGNKLFNNPNYQQNAQRTYKSPQAYQRSQNSFSKSAPS
ASSMGGASKGQSGFFGSSRPTSSPAISSGTRGFNS

>PROKKA_00682 Putative acid--amine ligase YgiC
MQVIPLKPLDNKTLEEIGLDWHTNDDMSSYIADENVVVSQKEADAYYDACNELYDMFVET
AEEAIQNDRFFELDIPNALIPMIKQSFEEEVHWHIYGRFDLAGGLDGKPIKLLFENADTP
TMLYETALIQWALLKANGYDENKQFNNLYEALGENFKRMVTLGEDTSRFEEMYEGWKILF
SSVRGNIEEERTMRFLQDAAQSVGFETDFSIDEVEFNIEEGVFKNGLNYEFLFKLIPWE
NIAIDEPELALLMQGMENKNTIFLNPAYTILFQSKRFLKLLWDRYPKHPLLETSYEPL
ANKKQIKKVAFGREGANSEIFEASMQSLLKTDGVYSNHKPVYQEFYELNSHNGLYYQPNV
FFAYESCALGFRKGGLILDNFSKFVSHRLQ

>PROKKA_00683 hypothetical protein
MTAMMRYFHIIYATFFFFPLALLFAVSGFSLLFKVRQDTGATIKEWVLEKSLKKEERLDFL
KDFLKENHIAMPKKIEPREHRGALVIGTPLYEINLETKGDTKIKTIERGFLGALIMLHK
AKVGVVFKTLSGIFCVFLLLFYVSAFLMVAFKDTKRMFISVLIGSVVFFGAIYWSL

>PROKKA_00684 Putative beta-lactamase HcpC precursor
MSVKILKILVCGLFFWSLNLHLWGKQDNSFLGVAERAYKSGNYSKATSIFKKACNDGVSE
GCTQLGIIYENGQGTRIDYKKALEYKYTACQADDREGCFGLGGLYDEGLGTTQNYQEAID
AYAKACVLKHPESCYNLGIYDRKIKGNADQAVTYYQKSCNFDMAKGCYVLGVAYEKGFL
EVKQSNHKAVIYYLKACRLDDGQACRALGSLFENGDAGLDEDFEVAFDYLQKACGLNNSG
GCASLGSMYMLGRYVKKDPQKAFNYFKQACDMGSAVSCSRMGFMYSQGDAVPKDLRKALD
NYERGCDMGDEVGCFALAGMYNMDKENAIMIYDKGCKLGMKQACENLTKLRGY

>PROKKA_00685 hypothetical protein
MRLLIALLGLSLLSLNAKEADFISDWEYGLALYKNPRGVACAKCHGIKGEQQEITFYEYEG
EKKILYAPKINHLDFTKFDALSLGKGMPKYNLNLLEEIQAIYLYITSLEHKEEHKDSPK
P

>PROKKA_00686 Porphobilinogen deaminase
MGNLVIGSRGSELALWQANHIKERLKKECFIESEIQIVKTTGDKILDAPLNKIGGKGLFT
KELEELLLKGAIDLAVHSLKDVVPVFEKGLDLACITKRADARDTFLSVKFPDLMSLPKGA
KVGTTSLRRSMQIKLKRQDLDTESLRGNVQTRLKKLECGEFDAILAEAGLCRLNVQGAK
YRKAFSVEEMIPSMGQALGVEMLNKHKHFITLQKLNDEESAFCCHLEREFIKGLNGGCQ
IPVGVHANLMDDEVKIQAVLGLPNGKEVITKEKQGDKNKAFDLVQELLEEFLOSGAKEIL
EKAQLF

>PROKKA_00687 Proline--tRNA ligase
MLFSKLFAPTLKEPPKDAVLKSHKHLAQAGYIYQVSGSIYNFLPLAKKVLDKIENITHKR
MQEHGAQNILMSFVVLASLWEKSGRLDKYGKELLVFKDRKNNDVFLSPTLEENITEIAAN
FIKSYKQLPVHLYQIHTKFRDEIRPRFGLVRAREFIMKDGYSFHEDAESLDKEFLNTQSA
YKEILSDLGLDFRIVEADSGAIGGSKSKEFVVLTCEGEDTIVVCKNCDYAANIEIAKRSK
RPEPSNVPKAQLAKFPTPNTTSTQSVAEFFKTEPYFVLKALVKKVIHKDKETLACFFVRG
DDNLEEVKALNALNLGANALELREASKEDLNHAGLIAGFIGPYGLKKHVSIIIFDEDLK
ESDCLIVGANEKDFHAVGVDLKGFFENLVYADIVQKESDHCPNCQGALKYHKSLEVGHIF
KLGQGYAKSLKASFOLDKNGKEQFFEMGCYGIGISRLLSAILEQKSDDLGCVWTKNTAPFD
VVIVVSNWKDEAQKKLAFEVYERLLQKGVLDALLDDRARFGAKMRDFELIGERLALIVGK

QTLNKEFECVKRANLEKQTIKDIELEEKILEILASE
>PROKKA_00688 Glutamyl-tRNA reductase
MELETHLSKYFTLAFTHKSMSLETREKLAINSSSTTLKEFLQTIKTHCPNIKECMVLSTCN
RFEIYASLKHGANTNEQKNALLKILAQNKKMSVSDLEKCVLINTDESAVHHVFSVCSSLD
SLVVGETQITGQMKNAYKFAFEKFC SKDLTRLHFAFKCAAKVRNLTGISKQGVSISSV
AVKEALSIFEKEKIKDKKALVIGLGEMAQLVIKHLNKGFEALILGRNAAKFEFEFIKELE
DPKKVSFQNIENLNAYINEYELLFCATSSPHFIVQNGMLKETIFRRFWFDLAVPRNIEKP
VLDNIFLYSVDDLEPMVRENVENRQESRTKAYEIVGLATMEFYQWIQSLEVEPVIKDLRE
LARISAQKELQKALKKRYVPKEYESNIEKILHNAFNTFLHHPTIALKKNAQKEESDVLVG
AIKNLNFNLDKSNANHAQNLNLYKCEYYEE
>PROKKA_00689 Octaprenyl-diphosphate synthase
MQEKRLQAIQNKIASWIKIEIESGFIDALFSKIGPSKMLRSKMLLALLDEKTDAILLDKAF
NLCAIVEMIQTASLLHDDVIDKATMRRKLPSINALFGNFNAVMLGDVFYSKAFFELSKMG
EAIAQILSNAVLRSLRGEIEDVVFVGESFNSDKQKYWRILEDKTAHFIEASLKSMAILLNK
NAKMYADFGNLNFGMAFQIIDDLLDITQDAKTLGKPNFSDFKEGKTTLPYLLLYEKLNQHE
QGLLISYFKQDSHEIIEWTKEKFKQYGIIEETLKIAQVYSKKALEAIEGENNLIILEKLAQ
DVIYRTF
>PROKKA_00690 hypothetical protein
MNELIRYGLIFLFFLEAFGLDYGIDKTLELKKDEVFKAIVKDT SNEQTKEITLYWTLNAN
KGLVINMRFNHFPYQFILYTDHARNTYNLKVFEKFSNSTLSLVFKDFKEDKATLRFLA
LMPLVFSPEP
>PROKKA_00691 hypothetical protein
MDVTSPFSLSDQKEASMRDYSELEIFEGNPLDKWNDIIFHASKKLSKKELERLLELLALL
ETFIEKEDLEEKFESFAKALRMDEELQQKIENRKTDIVIQSMANILSGNE
>PROKKA_00692 DNA protection during starvation protein
MKTFEILKHLQADAIVLFMKVHNFHWNVKGTDFFNVHKATEEIIYEEFADMFDLLAERIAQ
LGHHPLVTLSEALKLTRVKEETKTSFHSKDIFKEILGDYKHLEKEFKELSNTAEKEGDKV
TVTYADDQLAKLQKSIWMLQAHLA
>PROKKA_00693 Sensor protein ZraS
MKKSKRLKRPYLKRSHLKHNSKASSFKGLLKEDNVISLENFKPKESDLENFSNKKDM
QELLGLLNQFILQSYKVEKEFKDYKALYEWVIEILPQAIWVNVNENGFFYKNSLANQSHE
VFNKAKLENFNTEIEHENKSYLVQQNSIQGKQIITATDISAQKRQERLASMGKISAHLAH
EIRNPVGSISLLASVLLKHANEKTKPIVVELQKALWRVERI IKATLLFSKGIQANRTKQS
LKTLES DLKEALNCYTYSKDIDFLFNFSDEEGFFDFDLMGIVLQNFYNAIDAIEALEES
EQGQVKIEAFIQNEFIVFTIIDNGKEVENKSALFEPFETTKLKGNGLGLALS LQVVKAEH
GSIALLENQEKTFEIRILNAS
>PROKKA_00694 hypothetical protein
MINNNKAMLEQYNVSKLASEEKLKTLAQNKNDKLLKEQTDSFEALLLKFMLDTAMKMDNP
LYPKAPGDEIYTSMYKDTLSKELSGNFGYSEMLFNFLKEQEKQKP
>PROKKA_00695 Flagellar P-ring protein precursor
MKRVFLWLIFVLAFAHKLLAEKIGDIASVVGVRDNQLIGYGLVIGLNGTGDKSGSKFTMQS
ISNMLESVNVKISADDIKSKNVAAVMITASLPPFARQGDKIDIHIS SIGDAKSIQGGTLV
MTPLNAVDGNIYALAQGAIVSGNSNNLLSAHIINGATIEREVS YDLFHKNAMTSLSKNPN
FKNAIQVQNTLNKVFGNKVAIALDPKTIQITRPERLSMVEFLALVQEIPINYS AKNKIIV
DEKSGTIVSGVDIMVHPVVVTSQDITLKITKEPLNDSKNTQDLN NMSLDTAHTLSSNG
KNITIAGVVKALQKIGVSAKGMVSILQALKKSGAISAEMEIL
>PROKKA_00696 DEAD-box ATP-dependent RNA helicase CshA
MELNQPLPTEIDDDAYHKPSFNDLGLKESVLKSVYEAGFTSPSPIQEKAIPAVLQGRDV
IAQAQTGTGKTAAAFALPIINNLKNNHTIEALVITPTRELAMQISDEIFKL GKHTRTKTV C
VYGGQSVKKQCEFIKKNPQVM IATPGRLLDHLKNERIHKFVPKVVLDESDEMLDMGFLD
DIEEIFDYLPSEAQILLFSATMPEPIKRLADKILENPIKIH IAPSNITNTDITQRFYVIN
EHAEAEAIMRLLD TQAPKKSIVFTRTKKEADELHQFLASKNYKSTALHGDM DQDRRAS I
MAFKKNDADVLVATDVASRGLDISGVSHVFNYHLPLNTESYIHRIGRTGRAGKKGMAITL
VTPLEYKELLRMQKEIDSEIELFEIPTINENQI IKTLDHAKVSEGIISLYEQLTEIFEPS
QLVLKLLSLQFETSKIGLNQQEIDAIQNPKEKTPKPSHKKTPQHERARSFKKGQHRDRHP
KTNHHSKKPKRR
>PROKKA_00697 Modulator of FtsH protease HflC
MPIDLNEHLKKKNPQRETPTPNTPNNGGRFIPPSNSFNSKKLSVLIVIVLLGVIAFLAKP

FEVISSGEIGIKITAGKYEPTPLQPGIHFFVPIIQDILIVDTRIRNINFSRTEDMGVAGK
NQGIFRNDAINVMDSRGLTVSIELTVQYRLNPQTTPQTIATYGLSWEQKIINPVVRDVVR
SVVGRYPADLPIKRNEIAALINSNGINKEVSKLPNTPVELSSIQLEIVLPAKIKEQIEK
VQIARQESERVKYEVEVERSKQEAQKQAAALAKGEADANRIKAQGVADAIVIEAKAKSQANLS
ISQSLSDKLLRLRQIEVQGGQFNEALKTNNAQIMLTPGGAVPNIWIDTKSKVKSSIAETK
EP

>PROKKA_00698 hypothetical protein

MASLAFIQAFLESFKGFLSQATLISVLIASVLILFCAILLLLALLLRNRVASIATTAFL
GAFLSMPFVLKVLLTQAIYPIETRILHANPLSYSNAFSLQVGKHNHKSFLNKCVRLEV
LKNPHNFVEEHAFKWFVKVSKSYEKTFKEKILPEESKVFSFFIDNYPHSTAPYHVSFLCL

>PROKKA_00699 Glutathione import ATP-binding protein GsiA

MLEIKNLNCVLNAHFSLQNNINISLNPSEVAIVGESGSGKSSIANLIMRLNPRFKPHNGE
VLFETTNLLKESEEFMQHLRGNI IAYIAQDPLSSLNPLHKIGKQMNEAYFLHHKNASKTL
LKEQVLNAMKQVQLDEKFLDRYPYELSGGQRQRCIAMGIINAPKLLICDEPTTALDAQI
QNQILDLLKQLSAEKNIALLFISHDLKAVKRLADRVYVLKKGEIVETNLTKELFNGPKHE
YSKLLIQASNLPKLNKARDETLLEVKDFSVYYLQKRFFRPSLKKPLIASVNFSLKAKEN
IGIIGESGSGKSSLALGLLKLALNSGEEKILGQSVGPLNSKAFKPYRKILQMVFDQPYAS
LNRLSIQSIILTEALRFAYPKASKEEWHHLAKLCLEEVCLNPESLNFYAYELSGGERQRI
AIARAIALKPKIILLDEPTSALDKSIQKSVLELLLDLQEKQDLSYLFISHDLVDVIKAFCD
KVLVISEGKVVMGAIEEVFDNPKHAYTKRLLESRL

>PROKKA_00700 Inner membrane ABC transporter permease protein YejE

MSSLFKMRILSFKKNKRAVFSLYLFIALLSLLAPLWVNDRPLFIYNDNKAYFPIFKNY
AEVEFGGDFFTPTDYNPDYVKNLTLLKDAFIIHALIPYSYDTIIMDLDSPAPTPPSFKHLL
GTDDQARDVLARLVYGYRVSLVFGILLTLFSVLIGVSLGAFQGGYGGMLDLVGQRLSEIW
SAIPMLFLLIVISSAFNSNFWIILFLVLLFSWMGLSQVVRTEFLKARNMDYTKAARALGV
NDLKIIIFYHVLPNALVATITYVPFLMAASISTLVSLDFLGFGMPIGSASLGELVNQGKDN
LTPHLAIVAFVAISLLSVLVFIGEGVRDAFNANMLK

>PROKKA_00702 Helicobacter outer membrane protein

MKNHSFKKTIALSLLASMSLCRAEEDGAFFVIDYQTSLARQELKNPGFTQAQELKQLIRD
GAVRLQTSAPLSYYLDILGNKTKILLSESLKNNVQPSQPNAQPPQNGPSNQALANLEQSL
GILGKLLDLSQQYASQGVKPLVVDVGKEQIGITDSMLLVAQNIVLALGQVDLNKIQQNN
GNGQLYENIMKVMLLGAGGTNGAYNGVSVGDIATGMQNFSSQTGLIGANSTVSELNALIK
SGISLDRETGLGFSFIEKNICSGASSCFSGNQLIYKKGLDRTINIINTVLGQFESSASSL
YKISYIPNLFSLKDYQSASMNGFGAKMGYKQFFTHKKNIGLRYYGFLDYGYANFGDTNLK
VGANLVTYGVGTDFLYNVYERSRRRERTTIGLFFGAQIAGQTWSTNVTNLLSGQRPDVKS
SSFQFLFDLGVRTNFAKTNFNKHLRLDQGIEFGVKIPVIAHQYFATQGSSASYSMRNFSFYV
GYSVGF

>PROKKA_00703 Helicobacter outer membrane protein

MKNANTKEIKNTRMKKGYSQYRVLKKGLLKTALLFSLPLSVALAEDDGFYMGVGYQIGGA
QQNINNKGSTLRNNVIDDFRQVGVMAGGNGLLALATNTTMDALLGIGNQIVNTNTTVGN
NNAELTQFKKILPQIEQRFETKNKAYSQALQVYLSNVLYNLVNNSNNGSNGVVPEYVG
IIKVLYGSQSEFSLAAESVALLNALTRVNLDSNSVFLKGLLAQMQLFNDTSSAKLGQIA
ENLKNNGGAGSMLQKDVKTISDRIATYQENLKQLGGMLKNYDEPYLPQFGPGTSSQHGVIN
GFGIQVGYKQFFGNKRNI GLRYAFFDYGFTQLGSLNSAVKANIFTYAGTDFLWNIFRR
VFSDQSLNVGVFGGIQIAGNTWSSSLRGQIENSFKEYPTPTNFQFLFNLGLRAHFASTMH
RRFLSASQSIQHGMFVGKIPAINQRYLKANGADVDRRLYAFYINYTIGF

>PROKKA_00704 Adenylosuccinate synthetase

MADVVGIGIQWGDEGKGKIVDRIAKDYDFVVRYQGGHNAGHTIVHKGVKHSLHLMPSGVLY
PKCKNI ISSAVVVSVKDLCEEISAFEDLENRLFISDRAHVILPYHAKKDAFKEQSQNIGT
TKKGIGPCYEDKMARS GIRMGDLLDNKILEEKLNAHFKAIEPFKEAYDLGENYEKDLRGY
FKTYAPKIRPFIKDTTSMLEANQKGEKILLEGAQGTLLDIDLGTYPFVTSSNTTSASAC
VSTGLNPKAINEVIGITKAYSTRVGNPFPSEDSTPMGDHLRTKGAEEFGTTTKRPRRCGW
LDLVALKYACALNGCTQLALMKLDVLDGIDAIVCVAYERKGERLEAFPSDLKDCAPIYQ
TFKGWEKSVGVRKLDLEPNAREYIRFIEKEVGVKIRLISTSPEREDTIFL

>PROKKA_00705 hypothetical protein

MKKFASVLVQLKTLALEKIEQKLESKRLELQQNEREVLDKQAQLSAFKNPELGMSLFLQ
TQQLKSALRMEIEYYQQEGENLNKDLKILEKDYLLANQELEKAKIIIEKEKQKEQKILEK
KEQALLDENAMILHWQKEGLHA

>PROKKA_00706 hypothetical protein
MRKILLMGLILQALFSEEAQELLQCSAIFESKKAELKDDLRLQSEKEQSLRILQTENAR
LLDEKNDLLNQEKEVEEKLKNLAAKEEAFKTLQMEEEKTP
>PROKKA_00707 MgtE intracellular N domain protein
MIEENEGILREIKQAKDSKIGETYSKMKDSKSALILENLTQNALEILMALKPQELGKIL
AKMDPKKAAALTELWQKPPRENKEIPKTPPTPPPIVGTPLKEPKPMIKDPNTKEPAGV
>PROKKA_00708 Metalloprotease MmpA
MFIVAVLMLAFLIFVHELGHFTIARICGVKVEVFSIGFGKKLCFFKLFGTQFALSILPLG
GYVKLKGMCKEENGNETADDSYAQKSPFQKLWILFGGAFFNFLFAILVYFFLALSCEKV
LLPVIGGLEKNALEAGLLKGDKILSINHEKIASFREIRSVVVRARGELVLEIERNHQILE
KRLTPKIVAVISESNDPNEIIKYKIIGIKPDMQKMGVVSYSLIQAFKQALS RFEEGVVLI
VDSLRLIMGSASVKELSGVIGIVGALSHANSLSMLLLFGAFLSINLGILNLLPIPALDG
AQMLGVVFNKIFHITLPAFMQNALWLAGVGFLVFMFLGLFNDITRL
>PROKKA_00709 Exodeoxyribonuclease 7 large subunit
MHVLSVSEINAQIKALLEATFLQVRVQGEVSNLTIHKVSGHAYFSLKDSQSVIKCVLFKG
NANRLKFALKEGQEVVVFVGVSIVPRGDYQINCFEIEPKDIGSLTALQLKEKLRLKG
YFDKENKLPKPHFPKRVAVITSQNSAAWADMKKIASKRWPMCELVCINTLMQGEQCVQSV
VESIAYADSFHDTKNAFDAIVVARGGSMEDLYSFNDEKIADALYLAKTFSMSAIGHESD
FLLSDLVADLRASPSNAMEILLPSSDEWLQRLDGFNVKLHRSFKILLHQKKAHLEHLAA
SLKRLSFENKHHNLTKLEKLKIALENKTLEFLRFKKTLLLEKISIQALTS PFLQTKTERL
NRLLENALKLAHANLKL PQFGAFVSKNHQAVELEALKRGDKIELSNEKARASATILSVDRV
>PROKKA_00710 Modification methylase RsrI
MHKV FIMEALECLKRIEKESIQTIIYIDPPYNTKSSNFYEYEDDHADYEKWIKEHLILAKAV
LKQSGCIFI SIDDNKMAEVKIIANEIFGTHNFLGT FITKQATRSNAKHINITHEYVLSYA
KNKAFAPGFKILRTL LPVYAKALKDLMRTIKNVFRQKGQAQQLVLKEQIKELSKKEHFN
FLKNYNLVDEKGEIYFAKDLSTPSSPRSVAIQEINLFLEPLKSRGWSSDEKLKELYQNR
LIFKNNRPYEKYLLKESQDNCLSVLDFYSRQGTKDLEKLGLKGLFKTPKPVALIKYLLLC
STPKDSIILDFFAGSGTTAQAVIEVNKDYCLNWSFYLCQKEEKIKNNPQAVSILKNKGYQ
NTISNIMLLRLEKIIKRSEYEILKTKSILF
>PROKKA_00711 hypothetical protein
MQFLNQSLGFFNKGCFEPIDRNFITESYQALKPIEEIQNKYNKHDNDSFLNELRDSMVAL
YLDYELINIQKHGLDAKRSSSDEFLEIKQVSFQSKTWSATFNDDTLEKAKVFCDIKTTLA
VGVWNNISNLLFIVYGKHPEIGLYLEQKVKECHNESRRSTQTIGVSKLIKEFDFKMKPID
LKEQELINLFLNLFKGFHFS
>PROKKA_00712 hypothetical protein
MGLADIILERFKDFMREQPEPYKFLQVFFYMQEKERFLNHKMNDYIKQNKSKEEASILARQ
GFVSTIGRVLEKNHRTFIKRFLY
>PROKKA_00713 hypothetical protein
MTNDKILRAKCINGELDRVKRALLVHFGGYSVLLDIILYQTNKDNVKILAILSVKNSFRE
RFTETPYWKLKLLQSPVTSHIKVFMITPDNDDEISFKDKPKKARIVMEHE
>PROKKA_00714 Modification methylase DpnIIB
MKPYFSLEKLDLYHG DASVLETFEKGFYDL CVTSPPYNLSIEYQGSNDFRAYDDYLNWCK
NWLKNCFYFWGKEQARLCLNVPLDTNKHGKQSLGADI IAIACEGWKYQNTIIWNESNISR
RTAWGSWLQASTPYAIAPVELIVVFYKNEYKRKKQTSTMSKEEFFLLYTNGLWNFSGESKK
RLKHPAPFPRELPRRCIQLFSFLED TIFDFPSGSGTTILEANALGRFSVGLEIEKEYCEL
SKKRILESLSLV
>PROKKA_00715 Chaperone protein ClpB
MNLFEKMTDQLHEALDSALALALHHKNAEVTPLHMLFVMLNNSQGILIQALQKMPVDIEA
LKLSVQSELNKF AKVSQISKQNIQLNQALIQSLENAQGLMAKRGDSFIATDVYLLANMSL
FESVLKPYLDTKELQKTLESRLRGRTIQDKNDDSNLESLEKFGIDLTQKALDNKLDPVIG
RDEEII RMQILIRKTKNNPILLGEPGVGKTAVVEGLAQIRVNKEVPKTLNKRVI ALDL
SLLVAGAKYRGFEERLKKVIEEVKKSANVILFIDEIHTIVGAGASEGMDAANILK PAL
ARGELHTIGATT LKEYRKYFEKDMALQRRFQPI LLNEPSINEALQILRGLKETLETHNI
TINDSALIASAKLSSRYITDRFLPDKAIDLIDEGAAQLKMOMESEPAKLSSVKRSIQRL
MEKQALEMEKKESNAKRMQEILKELSDLKEEKIQLEAQFENEKEVFKEISR LKMEMESLK
KEAERFKRNGDYQQAGEIEYSKIPENKKKEEELQHKWEAMQONGALLQNALTENNIAEIV
SQWTHIPVQKMLQSEKNRVLNIESELQKRVVGQEKAIKAIKAIKRNKAGLSDSNKPIGS
FLFLGPTGVGKTESAKALAQFLFDS DKNLIRIDMSEYLEKHSISRLIGAAPGYVGYEEGG

QLTEAVRRKPYSVLLDEVEKAHPDVFNLLLQVLDEGHLTDSKGV RVDFKNTILILTSNV
 ASGALLEEDLSEAEKQKAIKESLRQFFKPEFLNRLDEIISFNALDSHAVANIVEILFENI
 QKKALGRGINITLDEKAKELIAEAGFDRFYGARPLKRALYEMVEDKLAELILEDKIKEND
 SVAFFVENNEIVPKIK
 >PROKKA_00716 Thiol:disulfide interchange protein DsbD
 MFDNTLVNLFDTAPLLTSLLAGILTFLSPCVLPLIPAYMSYISQISLEDIKDGKAKRVSV
 FLKSLMFVVGFSLVFLGVGMSMAKLIHSFSFSWVNYIAGGIVILFGLHFLGVFRFAFLYK
 TQSVRLASKSNSMQRFYPFLLGMSFALGWTPCIGPIFTSIVIMSASKDAYGLMLMVVFVM
 GLAIPFLLVALMLERALLFLKSLKKYNRAIEIVSGLVLILMGVLIMTNSLESITNLFQN
 >PROKKA_00717 Dihydroorotase
 MLLKNASFYDDEV LKRADIRLKD SLII EIKENLSPINNEEVIECGNLFVLPSFIDLSVTG
 LEGYENLKQKAFKGGVGLLN VFDCDKSGIKNIMAIKNNQLADIATLKNKGGEILIAPSDA
 FLELISHYAKSYNLP LLVSLENSFEALNSGVLAYELGQNFVENAFENTRLVRFMEVSRAL
 QIPMLLDKVVSIATLKL IAFNNLGARLQAQTPLSHLVLDSEVYEDYEPRFKIAPPLRDK
 EGQNALKEALKNNKIAMLTSLHVFKN SNAELFEESAFGCESIEDAFSVAYTFLVQKKVIS
 LQQLIKVMAINQAKFLKLNAGEVKENQLANLMIVDLNAQTRVNNQNSPFYGLELYGEVQR
 MILKGQTTFIKENACKKS
 >PROKKA_00718 Adenosine deaminase
 MQEIIIGASLVFLCNEKCEVLEDYGVVFDEKIVETGDYKNLTLKYPHLKAQFFENS VLLPA
 FINAHTHFESNNKASF DYGSFSGWLGSVLNNGGAILENCQGAIQNAIAVQLKSGVGSVG
 AISNHLIEVDLLKESPLNAVVFLEFLGSSYSLEKLKAFEAKFKELKDLEDKKLKAALAVH
 APYSVQKDMALSVIQLAKDSQSLLSTHFLESFEELEWVENS KGWFENFYQHFLKESHFKS
 LYKGANDYIDMFKDTHTLFVHNQFASLEALKRIKSQVKNAFLITCPFSNRLLSGKALDLE
 RAKEAGLSVSVATDGLSSNISLSLLDELRAFL LTHNMPLLELAKIALLGATR HGAKALAL
 NNGEIEANKRADLSVFGFNEKFTKEQAILQFLLHAKEVERLFLGGKRV I
 >PROKKA_00719 hypothetical protein
 MKLVLAKNTRKSDAKSVELEDLYRKFS EDKRSIFYFAPTNAHKDMLKAVDFFKEKGHTAY
 LDEV RVSTDEKDFLYELHII
 >PROKKA_00720 (Dimethylallyl)adenosine tRNA methylthiotransferase MiaB
 MKVYIETMGCAMNSRDSEHLLSELSKLDYKETS DPKTADLILINTCSVREKPERKLFSEI
 GQFAKIKKPNAKIGVCGCTASHMGADILKKAPSVSFVLGARNVSKISQVIHREKAVEVAI
 DYDESAYAFEF FEKKAQIRSLN ISIGCDKKCAYCIVPHTRGKEISIPMDLILREAEKLA
 NNGTKELMLLGQNVNNYGARFSSEHAKVDFSDLLDKLSEIQGIERIRFTSPHPLHMNDRF
 LERFAKNPKVCKSIHMPLQSGSSAVLKM MRGYSKEWFLNRVERL KALVPEVGISTDIIV
 GFPNESDKDFEDTMEVLEKVRFD TLYSFIYSPRPFTEAGAWKERVPLEVSSSRLERLQNR
 HKEILEEKAKLEV GKTHVVLVENRREMD DQIVGFEGRSDTGKFIEVACKEKRNPGE LVRV
 EIVSHSKGR LIATT KGN
 >PROKKA_00721 hypothetical protein
 MSLKFFRKNIVLKVVPRLAFGVLWLLHKTCKNRYFLAQNLKEKPFIVSCWHGELGMIGFA
 YLRLQKPSVYVIASQHFDGSIAAGLFESFGFKNIRGSSKKGGVKV LIEGLKRLKEGCDVA
 ITPDGPKGPRHSIADGVIALAQKSGVGISACRVCKNAWRLNTWDQFEIPKPFSEVYYM
 LESV IIPKEWELSKAKEYLKTRMDSVGFEESQ RGLGA
 >PROKKA_00722 hypothetical protein
 MGLKNLKG VVWLKGLKKVLKERFCSQVYISFNVDH NLLSTQVIRIKNDRIKEKFFKTFET
 KVETKNGEVPIQALKIARTYSQKYP TYFSAMSKAKEVLCEKQAFEQIKQENQDYHACEV
 NQKYCVYVESKDFLKD FKRFKIQD VDFLFS PFSLIYDFVRDHLENKPLLYLLERSRFYF
 LIADKKEIFLAKSVFLEE QPEEFIESKEEDFMGMDNEAVDLFLNEIQEDIDSLEE AIGLD
 SSKDNSEKNKDAYSLIEGMTNIPLIANALQEGLRGVYHSRDIDFVEKVVVL DSCQIHQKA
 LMHLQETLMIEVDRLDFSLVERLNILARMENEKHAF
 >PROKKA_00723 hypothetical protein
 MRFSYIEPRAKYLISKLSKIWFYIFLSFV VIGGLVGMHNAIKSTQDNASDLTIQERLY
 RHEISRLQVKTDETLKLIKEAKRLNYNDDIR DVLQGLLNIVPDSITINSIEIDQQSVVV
 SGKTPSKEAFYFLFQNKLNPMFDYSRAEFFPLSDGW FNFVSTNFSNSLLIKNPESIK
 >PROKKA_00724 hypothetical protein
 MKPLHFSHLDREQSGDVGFIIKNLVFLGVFSL LGWLNT EYFLWPSMLELKKILLEENRKK
 SVLEYAQRHFETALTN YRNQKETSESLKIFNDEESRRILEKILKKCFDAYKIKPLLSQN
 PLQKTQFFIMARASELEKTYLFFTLINKYLP SAQSQLPLKISKDDEGLLVQFSV SIDLQ
 >PROKKA_00725 Flagellin N-methylase

MQDFDFSFPKACEGCGAKCCVGESGYIFLNIQEMQQISAFLELEEFSSQKYVKKVGYK
FSLLEKDAKELGLACVFLDLETKKCQIYSVRPKQCQTFPFWESVKTFSKEQKEAFCQSCP
GITQKTKETKVR

>PROKKA_00726 PPR repeat protein

MDIPIKKRFLANLLLFSLFSLKAETLSEDHQILLSSDAFHRGDFAAAQKGYMNLKQTN
KVYAKEAAISAASLGDIKTAMHLAMLYQKITNNRNDVSINKILVDGYVQMGQIDKAIEL
LHKIRKEEKTIATDNVLGTLTYLTQKRLDKAFPLLNKFYNQVHDEDSLEKLITTYFLQNRK
KEGLDLLQSHIDRYGCSEQLCQKALNTFTQFNELDLAKTTFARLYEKNPIVQNAQFYIGV
LILLKEFDKAQKIAELFPFDRLLLDLYTAQKKFAQASKQASLIYQEKKDPKFLGLEAIY
HYESLSAKNKKLTKEEMLPIIQKLEQATKERQQILAKTKDKEDAQDAFFYNFLGYSLIDY
DMDIKRGMDFVRKALALDSGSVLYLDSLAWGYKLGNCLEAKKIFSSIAKESIQAEPELK
EHNKIIQECKK

>PROKKA_00727 hypothetical protein

MQEIGILENLQKSLALKEGMLSYLEMLGKSLSYNPYLPRVIPQTKDCVFVTPDEVLEKLLK
ENTHTDCVIVNFKGLYEIGAPSVFDLEVLGLLRRHASSLIVHQDLFISHYQLLESLSVQGS
DGVILDEELLKEDLKGMVEFSWRGLSVFVETDKPDYTHLKDGLVGLVLENSPHSYNQKK
IVFLD

>PROKKA_00728 Ferredoxin

MSLLVNDECIACDACREECPSEAIIEEGDPIYNIDPDRCTECYGYDDDEPRCVSVCVDAI
LPDPNNAESKEELKYKYESLKEQD

>PROKKA_00729 Guanosine-5'-triphosphate,3'-diphosphate pyrophosphatase

MAKITTVIDIGSNSVRLAVFKKTSQFGFYLLFETKSKVRISEGCYAFNGILQEIPMQRAT
KALSEFKEIALKYKSKKILCVATSAVRDAPNRLEFVARVKKACGLQIKIIDGQKEALYGG
IACANLLHKNSGITIDIGGGSTECALIEKGKIKDLISLDVGTIRIKEMFLDKDLEVKLAK
AFIQKEVSKLPFKHKNAGVGGTIRALSKIIMKRFDYPIDSLHGYEIDVHKNLAFIEKIV
MLKEDQLRLLGVNEERLDSIRSGALILSVVLEHLKTSLSMITSGVGVREGVFLSDLLRHHY
HKFPNINPSLISLKDRLPHEKHSQVKKCECVKLFEALSPLHKIDEKYLFLHLKIAGELA
SMGKILSVYLAHKHSAYFILNALSYGFSHQDRAIICLLAQFSHKKIPKDNAIAHMSAMMP
SLLTLQWLSFILSLAENLCLTDSHHLKYTLEKNKLVIHSNDTLYLAKEMPLPKLVKPIALT
IEFA

>PROKKA_00730 Lipopolysaccharide heptosyltransferase 1

MKIAIVRLSALGDIIVSAVFLAAIKERFTDAQIEWFVDERFSAILEHSPYIDKLHPIALK
SALKTFNPLKIFKLFKSLRAYEYDIIIDMQGLIKSALITQMLKAPKKVGFYASAREGLS
AFFYSQKVSIAAYDEPILKRNFTLLSQALNLPKKEISQSLSSRFKVFSCQDSPKINALNLN
QNKPKILFVLETSKINKTYPTERFKELALALENFQICLLWHADEKKAADVLYHALKHQRDV
LLLPLKLTNEVKALLFKMDLIIGGDTGITHLAWALQKPSITLYGNTPMERFKLESPINVS
LTGNSNANYHKKDFSQNIIEPKKIKECVLNILKEKE

>PROKKA_00731 Lipid A biosynthesis lauroyl acyltransferase

MTYKERLIHEKILNQNDKGLKTELRIISIFIVESLVNIGFILAKMPHSWFLRCIKALAW
LMRTFDRHRYFDAKANLDFVFGDSKSEEEKKRIIKKGYENFAFIILETIRVIFIPKDEYD
ARFTLINEENVWKSINKEGQAITLCMHFGYWEAVGTTLAQYYENYGRGCLGRLTKFAPIN
HMIMSRREAFGVRVFNKVGAMKELIKMYNQGNLVLGILVDQNVVPKDGVVVKFFNKDATH
TTIASILSRRYNIDIQPVFIDFNDDYSHYTATYYPSIRSKITDNAENDILECTQAQASLC
EEVIRNHPESYFWFHRRFKSTHPEIYQR

>PROKKA_00732 Queuine tRNA-ribosyltransferase

MDFQLQATDKHARAGLLNLAHSQVATPVFMPVGTQGCISLDAMDMQEILGAKLILANTY
HMYLRPGEKVVEQLGGLHRFAQFHGSFLTDSGGFQAFSLSDNVKLQEDGIIFKSHIDGSK
HLFTPAKVLDIQYSLNSDIMMVLDDLVLGPAPLKRLEESIKRSKAWANLSLEYHKEKNRP
SNNLFAIIQGGTHLMRSLSVELTHEGFDGYAIGGLAVGESVDEMLETIAHTAPLLPKDK
PRYLMGVGTPENILDAISLGVDMFDCVMPTRNARNATLFTHSGBKISIKNPYKLDTPIE
ENCACYACKRYSKAYLHHLFRAKELTYARLASLHNLHFYLELVKNARNAILEKRFLSFKK
EFLEKYHSRSH

>PROKKA_00733 TrkA-C domain protein

MKKIALILDGIVAKNFLDLVLRHYSNHNFIYIVVKDESIPKNYPSTFAFHCFDATSSFR
LLQVLNDEVSDAFLIIQDFKEQCIHKIIQTHFKRMRVLSVKKDGEKTLENNEENKDEK
LILIDEFEVLANKFISRLPNIPSTPREFGLGKGEIMEIDVPFGSIFAYRHIGSIRQKEYR
IVGLYRNDVLLLSTKSLVIQPRDILLVAGNPEILNVVYHQVKSINVQGFAPAFGKSIYLYI
DMRLQSRKAMMRDVYQALFLHKHLKSYKLYIQVLHPTSPKFYHKFLALETESIEVNFDFY

GKSFIQKLHEDHQKKMGLIVVGRELFLSKKHKALYKTATPVYKTNTSGLFKTSQSVVVL
 NESLDINEDMSSVIFDVSMQMDLGLLLYDFDPNKRYKNEIVNHYENLANTLNRKIEIFQT
 DIRNPIMYLNSLRNPILHFMFPFEECITHTRFWWFLSTKVEKLAFLNDDNPQIFIPVAE
 >PROKKA_00734 3-dehydroquinase synthase
 MQEILIPLKEKSYKVFLGELPKIELKQKALIISSIVAGLHLSYLLKHLKALEVRVCVIE
 SGEKYKNFHSLECI LNNAFEMQLNRHSLMIALGGGVISDMVGFASSIYFRGIDFINIPTT
 LLAQVDASVGKGTGINTPYGNLIGSFHQPRAVYIDLSFLKTLEKREFQAGVAEIIKMAV
 CFDKNLVEILETKDLKDCLEEVVFQSVSIKAQVVVQDEKEQNIRAGLNYGHTFGHAIEKE
 TDYERFLHGEAIAIGMRMANDLALSGLMLTLKEYERIEENLLKKFDLIFHYKIIDIQKFYE
 RLFLDKKSENKTIKFI LPKGIGAFEVASHIPKETILKVLEKWH
 >PROKKA_00735 Mechanosensitive ion channel
 MALRVLLFFCFLFLQAEDKSQELLSIQQMALVDKKLAKDDNVWLKKFENYKIYNQIYTE
 KESVRQELRRLKNKKSDDLKISTLEHTLKALESQQKMFESYGVNPFKDLIERPNIPNIP
 NIANPIAIIIDGISFIKSMHLKHENLKRNTALEEVLRLLDQKHQLLNEWHALDKSAKLS
 EIYQTQAKRLELQGAQNILKTTIGIFQKDSDEAISIVKSQVKNQLFKLIYVFLAALLSVV
 FAWILKIISSKYIENNERVYTVNKAINFVNVSVIILIFLFSYLENVTYLVTVLGFASAGL
 AIAMKDLFMSLLGWFIILIGGSVHVGDVRVIAKGTDIFIGDVLDISMLHITILEDVTFTT
 YTNRRRAGRIIFVPNNYIFTTTFANYSHFGMKTVDGVDGVCITFDSDFKASKIALNIAT
 ELSKEYTDITYKQLNKMRDRYSRLSVKPRCFLMPENNGIKISVWYQTN SYATMSLRSK
 IVAEIVEAFLKEENIHIAYTTSKLLKVDADFLGDGFGNKRQK
 >PROKKA_00736 Threonylcarbamoyladenine tRNA methylthiotransferase MtaB
 MKKVYFKTFGCRTNLFDTQVMGENLKDFSATLEEQEADIIVINSCTVTNGADSAVRSYAK
 KMVRLNKEVLFTGCGVKTQ GKELFEKGFLKGVFHGNKEKINALLQEKRRFFIDNLENK
 HLDTTMVSEFVGKTRAFIKIQEGCDFDCNYCIIPSVRGRARSFEERKILEQVGLLCSKGV
 QEVVLTGTNVGSYGKDRGSNIARLIKLSQIAGLKIRIGSLEPNQINDEFLELLEEDFL
 EKHLHIALQSHDFMLERMNRNRRTKSDRELLEKIASKNFAIGTDFIVGHPGESESVFEK
 AFKNLESLPLTHIHPFIYSKRKDT PSSLMRDSVSLEDSKKRLNAIKDLILHKNKAFRLLO
 LKLNTPKALVEAQKDGEFKALDQFFNPIKIKSDKPLRASFLIKEYEIKERENHAVF
 >PROKKA_00737 ATP-dependent zinc metalloprotease FtsH
 MPFSKFLENLTAPFKRIKNSLVLALGFLILTFCLLLFLVLSDVSR LISGKDFFYVIQSH
 PKQTLIEDENYFYANKGLYKTNKEAFLRAYKIPESMPIEKRENLNKVSKIFLALLFFISS
 MLFGIFWRPLKRLDTKMSLES AHKNELENAFQRYDALGVRFEDIAGVDEVKEELLEVIDY
 LKNPKKYQDLGIFLPGVLLIGPPGVGKTMIAKALASEARVPFFYESGSAFSQIYVGAGA
 KKVRELFMHAKRHAPSII FIDEIDALGKARGGHRSDEREATLNQLLTEM DGFLQNDVVV
 IGATNQMEVMDEALLRSKRFRDRIFISLPDLLERQSILEKLENKRHALNYLKIACICVG
 FSGAMLATLINESALNALKHQRKEIAESDILEVKDKIAYGKKKPQTLDENQKELVALYQS
 AKALSAYWLEIEFDKVPILGEFIAFNENKIHSESEIKNYIKVYLSGTIVLELLYKERYSL
 SKQDLQKAKFLNEFMASELLLASTKESLSVLYEEQLEFLKPQIAACKRLSVLLLEQECLE
 HANFYDLLNG
 >PROKKA_00738 hypothetical protein
 MRFFSGFGFVNESVLFEWLLKGAYDISGFSMGAIKAIEYAYNEILQORRINSLLLFSPC
 MLAHKS LAFKRLQLSSFQKDPQSYMDNFYKEVGLSAQLERFKKEGSLEELEFLLNYKYSD
 SIIRFLLKEGVKIEVFGLKDKITDIQALLEFFMPLVQVWQFKDCNHLQRS
 >PROKKA_00739 hypothetical protein
 MKKFGGLGVYLLLLGILGGSLLIILGAIVAPIVFKASSILPELNLT PFESGKLMAQIFVRFN
 YVLGAIGFVLLLYEII SFIYSKRSLVYLILGVAIGALCLLFVYYTPYILNAQKIGEVAL
 QSAEFARSHAQSEWLFKELFVLVCAFFWRLLGKNAL
 >PROKKA_00740 Vacuolating cytotoxin autotransporter precursor
 MKKFKKKPKSIKRSHQNK TILKRPLWLTPLLIGGFAGGVYADGTDILGLSWGEKSQRVC
 VHPWYALWSCDKWEEKTQ QYTGNQLITKTWAGGNAANYHTQNNQNITANLKNNDNGTYF
 LSGLYNYTGGEYNGGNL NIELGSNATFNLGASSGNSFTSWYPNGHTNVTF SAGTINVNS
 VEVGNRVGSGAGTHTGTATLNLNANKVTINSNISAFKTSQVNI GNANSAITINSVSLNGD
 TCSSLASVGVGANCSSLGPSYSFKGTTNATNTTFSNANGSFTFEENATFSGAKWNGGTYT
 FNKEFNATNNTAFNSGSFNFKGASSFN GDSFSNASYTFENQATFQNSSFNGGTFTFNNQT
 NQSAQHPQIQNSSFSGGDTTLKG FVNFQQA FNNSNHQLTIQNASFNNANFNNTGKITINE
 GASFNNTSFNTSINTNMSVTGGVTL SGKNDLKNGSTLDFGSSQVTLTQGTTFNLTSLS
 EKSVTILNSSGGITYNNLLNHAINGLTNALKTTESPSKPQSFAQGLWDMITYNGVTGQLL
 SENATTPKPADSSPSAPT KDSPQVYQVGYKIGDTIYKLQETFS PNSII IQALES GTYTPP

PVINGSKFDLSASNYINADMPWYDHKYYIPKSNFTESGTYYPVSVQIWGSYTNSEFKQTF
SANGSNLVIGYNATWTGNSVSSSGTVSFGDTSALNGHCGPWPYYQCTGTTNGAYSAYH
VYITANLRSGNRVGTGGAANLVFNGVDSINIANATITQHNAGAYSSSMTFSTQNMDSQN
LNGLNANGTSLVYGTTFTNEAKDGKFIENAGQAVFENTNFNGGSYQFSGDSLNFNSNQF
NSGSFEISAKNASFNANFNNSASFNNSSATTSFVGDFTNASHNLQIAGNAVFGNSTN
SDGSQNTANFNNTGSVNIAGNATFDNVVFNSTNTSVKGVTLNNITLKNLNAPLSFGDG
TIVFSAHSVINIDQITITNGNPITLVSSSKEIEYNNAFSKNLWQLINYQGHGASSEKLVSS
AGNGIYDVVYSFNNQTYNFQEVFSPNSISIRRLGVGMVFDYMDMEKSDHLYYKDVVGFM
YMPNSYNNNLGNNSNTIYYIDKSIDFYASGKTLFTKAEFSQTFGTGQNSTIVFGAKSIWTN
ASDAPQSNITIRFGDNKGAGSNDASGHCWNLQCIGFITGHCWNLQCIGFITGHYEAQKIY
ITGSIESGNRISSGGGASLNFNALQGILLTNATLYNRAAGTQSSSMNFISNGANIQAQNS
YFIDDTAQNKGNPNFSFNLNLDNFSNSSFRGYVGATQSVFKFNAKNAINFTNSTNLSSGL
YQMQAQSVLFDNSNLVSVGTSSIKANAINLSQNASINASNHSTLELQGDNLNNDTSSLN
LNQSTINVSNNATINDYASLIASNGSHLNFNGAANFNNSANITTSLSDDSSIVFKGAVSLGG
QFNLSNNSSLDFOGSSAITSENTAFNFYDNAFSQSPITFHQALDIKAPLSLGGNLLNPSS
SVLNLKNSQLVFGDQGSNLNIANIDLSDLDNDKNRVYNI IQADMNSNWYERISFFGMRIN
DGIYDAKNQTYSTFNPLNALKITESFKDNQLSVTLSQIPGIKNTLYNIGSEVFNYQKVY
NNANGVYSYSDDAEGVFYLTENVKGYNPNQSYQANGSNNTTKNNNLTSEPSVISQTYNA
QGNPISALHIYNKGYNFNNIKALGQMAKLYPEIKKILGNDFSLSSSLDLKGDALNQLTK
LITPSDWKNINELIDNANNSVQNFNNGTLIIGATKIGQTDNTSAVVFGGLGYQKPCDYT
DIVCQKFRGTYLGQLLESSADLGYIDTTFNAKEIYLTGTLGSGNAWGTGGSASVTFNSQ
TSLILNQANIVSSQTDGIFSMLGQEGINKVFNQAGLANILGEVAMQSINKAGGLGNLIVN
TLGSDSVIGGYLTPEQKNQTLSQLLGQNNFDNLMNDSGLNTAIKDLIRQKLGFWTGLVGG
LAGLGGIDLQNPKEKLIGSMSINDLLSKKGLFNQITGFISANDIGQVISVVLQDMVKPSKA
LQSDVAALGKQMIGEFLGQDALNSLESLLQNQQIKSVLDKVLAAKGLGSIYEQGLGDLIP
NLGKKGLFAPYGLSQVWQKGDFFSNAQGNVFNQNSTFSNANGGTLNAGNSLIFAGNNH
ISFTNHAGTLQLLSDQVSNINITTLNASNGLKINAANNVSVSQGNLFVNASCAQQSEPT
AANAANPCTLTQSANASSNNASNNAPIALNNNDESLMVTANDFNFSGNIYANGVVDFS
KIKGSANIKNLYLYNNAQFQANNLTISNQAVLEKNASFMANDLNIQGAFFNNATQKIEVL
QNLTIASNASLSTGIYGLGVGGALNNLGTIHFNLNSQTPVNPLIQAEIINLNTTQTPF
INVNNVMANNTTYTLKSSRYIDYNINPNLSQSYLKLYTLININGNHIEEKNGALTYLGQ
RVLLQDKGLLLSVALPNSNNASQNHILSLSVLHNQIKMSYGNKIMDFTPPTLQDYIVGIQ
GQSALNQIEAVGGNAIKWLSTLMMETKENPLFAPIYLNHSLNEILGVTKDLLNTASLIS
NPNFRDNATNLLELASYTQOTSRLTKLSDFRAREGESDFSERLSELKNKRFSDPNPGEVF
VKYSQPSKHKNLWVQGVGGASFISGGNGTLYGLNVGYDRLVKNVILGGYVAYGYSGFNG
NIMHSLGNNDVGMARAFLEKRNFTLSANETYGGNATSINSSNPLLSVLNQRYNYNTWT
TSVNGNYGYDFMFKQKSVVLKPQVGLSYHFIFGLSGMKGKMNDAAYKQFLMHSNPSNESVL
TLNMGLESRKYFGKNSYFVTARLGRDLLIKSKGGNTVRVFGENTLLYRKGEVFNTFASV
ITGGEMHLWRLMYVNAVGLKMGLOQYQDINITGNVGMVAF

>PROKKA_00741 Diaminopimelate decarboxylase
MFNYEELFQTHKTPFYLYDFDKIKQAFNLNKEAFKGRKSLICYALKANSNLSILSLLAHL
ESGADCVSIGIEYRALKAGIKPYRIVFSGVGKSGFEIEQALKNLILFLNVESFMELTTIE
KIAQSLGIKARISIRINPNIDAKTHPYISTGLKENKFGVGEKEALEMFLWAKKSAFLEPV
SVHFHIGSQLLDLEPIIEASQKVAKIAKSLIALGIDLRFFDVGGGIGVSYENEETIKLYD
YAQGILNSLQGLDLTIICEPGRSIVAESGELITQVLYEKKAQNKRFVVVDAGMNDFLRPS
LYHAKHAIRVITPSKGRKISPCDVVGPVCESSDTFLKDARLPELEPGDKIAIEKVGAYGS
SMASQYNSRPKLLELALEDHKIRVIRKREALEDLWRLEEEGLKGV

>PROKKA_00742 T-protein
MQKNLDSLLENLRAEIDALDNELSDLLDKRLEIALKIALIKQESPIYCPKREREILKRLS
QRDFKHLNEEILTGFYAEVFKISRKFQENALKELKK

>PROKKA_00743 hypothetical protein
MFEKITLAHKDLFSRFLSAQKIVLSVSTNCFWQHARLIQVAVIRDCLVIQTTYENQK
PFYFYPIGKNAFECVKELLKLEKNLRFHSLTLEQKDDLKDNFVGVDFTYNRDRSDYVYS
IEELIALKGKKYHKKKNHLNQFLTNYANFVYEKISSQNKKEVLEASKAWFLESQTDIGL
INENKGIQSVLENYESLDVKGGLIRVNGEIVSFSFGEVLNEESALIHIEKARADIAGAYQ
IINQQLLLNEFSLTYANREEDLGLEGLRRSKMSYNPVFLIDKYEAVAKN

>PROKKA_00744 Aminodeoxychorismate synthase component 1
MIFGDFKYQKSIKKLTATNFNELKNALDFISQNRNGYFVGYYLLYEARLAFLDETQSQ

PFLYFEQFLERKKYPLEPLKEHAFYPKIHSSLDQKTYFKQFKAVKEHLKNGDTYQVNLTMDLFLDTKAKPKRVFKEVIHNQNTPFKALIENEFSGSILSFSPELFFELEFLETAIKIITKPMKGTIARSKNPLIDEKNRFLQNDKRNSENVMIVDLLRNDLSRLALKNSVKVNQLFEIISLPSVYQMISEIEAKLPLKTSLEIFKALFPCGSVTGCPKIKTMQIIEELEKRPRGVYCGAIGMVEEKKALFSVPIRTLEKRAHENFLHLGVGSGVTYQSKASKEYEESFLKSFFVMPKIEFEIVETMKIIKRDQKLEINNKNNAHKERLMHSAQYFNFKYDDNLLDFELEKEGVLRVLLNKKGKLIKEYRTLEPLKSLEIRLSEAPIDKHNDFIYHKTTYAPFYQNARTLIKKGVMFDEIFYNQDLELTEGARSNLILEIHNRLLTYPYFSAGALNGTGVVGLLKKGLVEHAPLKLHDLQRADKIYCINALYGLVEVKIK

>PROKKA_00745 Aliphatic amidase

MRHGDISSSPDVTGVAVVNYKMPRLHTKNEVLENCRNIAKVIGGVKQGLPGLDLIIFPEYSTHGIMYDRQEMFDTAASVPGEETAILEACKKNRVWGVFSLTGEKHEQAKKNPYNTLILVNDKGEIVQKYRKILPWCPIECWYPGDKTYVVDGPKGLKVSIIICDDGNYPEIWRDCAMRGaelivRCQGYMPAKEQQIAIVKAMAWANQCYVAVANATGFDGVYSYFGHSSIIIGFDGHTLGECEGEEENGLQYAQLSVQQIRDARKYDQSQNQLFKLLHRGYSGVFASGDGDKGVAECPFEFYKTWVNDPKKAQENVEKFTRPSVGVAACPVGDLPTK

>PROKKA_00746 Flagellar hook-associated protein 3

MRVTFGSKYNQMNNYQNALQNKINDANTQIASGLKIRYGYQNSDINNQLKFQYEENTLDQGIDVAQNAYTSTLNTDKALQEFSTMEAFKTKLIQSANDVHSETSRAAIANDLERLREHMMNVANTSIGGEFLFGGSKVDRPPIDSEGKYHNGEDLNALISSDNLPYNISGQDLFLGADKDKHKLITTNIKLLNQKLHPDVMdaleHSSLPEEVFIKPSDTLRELIGDNDKDPTNDPKEFFYLQGVRPDGSSFKEKFALSKAYQNKESASKVSDLLDKIAHAYGNTSQNKVVDVSLNNWGQIEIKNLTPGSENLDHFLISSDGDGDFDDLDALRSSGKRVTEYVKSafVTDrsLSQVKAVPNMYPNPKTLEIPSVFVTKDNVLANKNTKLSEIFGDSVETLKINASRLDDESLIKIPNLPVYLDIPILLDVKNSTIKDLKDAIKERFNNEVDVEIETNGRLRIIDNSSKESPISLALSTLDAKGLEVAGIPTNNASEYQKTYFNKEGAKLESNVAQTAQNGAANGSTKLSEAAKGSLENsvfNMKLNDVNGLFLEAQINLDNNGAFLSLPngVKIPLYDPTTADIQASKPNEVTYRQLMDAMsIALNYSNTDPAIYQQISDNPTSkeskerfiELLKQAKDNLSVNLNEEGKVIIQDNMHSNTKMqFmLFdKDANDFSQNALHSDKPSLKLNNANNALIIDKPSVNFfdQLENIITSVRKGIYRPDALGDTYSSDMRNLGIQNGITLIDHLSdHIEKMIakNGahGKAfENIIRNEVLKTQVQSIRGETTGTDMAETYNKFSNLTNNAVlastnKINNLSLTKYL

>PROKKA_00747 50S ribosomal protein L21

MSYAIFKHGGKQYKVVEGDIVLLDKMNKEPKALVELVEVLAVSKEGKLSCGKPFVNGAKIEAEVINEGRGKKVITFKKRRRKDSKTKRGFRDFTRVRITKIVA

>PROKKA_00748 50S ribosomal protein L27

MAHKKGQGSTQNNRDSAGRRLGVKKFGSEFVRAGNIIVRQRGTMHPGNNVGMGKDHTLYALIDGVVKFEHKDRNRKKVSVVSQNFGE

>PROKKA_00749 Heme-binding protein A precursor

MKASESPNATLNPSKENVSVEERKRFGGVLVfARGADGSSMDPALVTDGESYVATGNIYDtlvQfKYGTTEIEPALATsWEISPDGLVYTFHLRKGvYFHQTKYWDKKVEFSakDVLFSFERQMDKAKRYSPGAKSYKYWEGMGMSHIIKSIEALDDYTIKFTLNGPEAPFLANLGMDFLSILSKDYADYLEQNNKKDELAKKPVGTGPFKFFLWNKDEKIIILLKNQDYWGPKAYLDKVvVrtIPNSSTRALALRTGEIMLMTGPNLNEVEQLEKLPNIVVDKSPGLIANWLSLNTQKKYfNNPLVRLAINHAINVDDYIKVIYEGfAQKMVNPFPPTIwGYNyNIkPYEYDLKKAKELLKQAGYPNGFKTTIFTTSTRNPKGAVFIQASLAKIGIDVKIEVYEWGAYLKRTGLGEHEMAfSGWMADIADPDNfLYTLWSKQAASAIPTQNGSfYKSDAFSDLLVKAKRVSDQKEREALYLKAQEIIHKDAPYVPLAYPYSVVPHLSKVKGyKTTGVSvNRRFFKVYLEK

>PROKKA_00750 Dipeptide transport system permease protein DppB

MLSfIIKRILWAIPTLFGVSVIVFMMVHLVPGDPALVILGEKANQAaidALREqFGLNKLPLIEQYFFFfINNVLHGnFGASIMTGEpVMHEfLQRFPATVELALIALfMALVLGISVGVLAaIKRYSVFdySSMTfALAGISMPVfWLGLMLIYIFSVQLGWLpVfGRlSDVYYLDGPTGFYLIDSLIARDYGAfVDTIKHLILPSIVLATVSTAVIARMTRASMAEVSKEDYVRTAKAKGcssFRVIVfVHTLRNALIPVTTIAGLMLAGLLGGSmitETVfSWPGIGKwIVNALNQRDFPIIQSMsLIIAMMYIGANLLVDILYAfIDPRIRLS

>PROKKA_00751 Dipeptide transport system permease protein DppC

MESfREFIQQfKKNAaAVGAWIVLLLVIcAVfAPLLVPHDPYVQNAQDRLLKPIWEHGGNAKYLLGTDDLGRDILSRLIYGARISLTIGIVSMGIaVFFGTILGLIAGYFGGKTDAIIMRIMDIMFALPSILLIVIVAVLGPslTNAMLaIGfVGIPGFARLVRSSVLGEKEKEYVIA

SKINGSSHLRLMCKVIFPNCIIPLIVQTTMGFASTVLEAAALSFLGLGAQPPKPEWGAML
MNSMQYIATAPWMLVFPGMIFLTVMSFNLVGDGIMDALDPKRAS
>PROKKA_00752 Oligopeptide transport ATP-binding protein OppD
MILEVKDLKTYFFTDKGVNKAVDGVSFGLKKSQTLICIVGESGSGKSITSLSILGLIEKPG
QIVGGSIQFLGQDLLQLKEKQMQKEIRGKKIGMIFQEPMTSLNPSYTVGFQINEVLKIH
PNLNKKERLERVVYELERVGIPHAGDKYHEYPFNLSSGGQRQRMAMAMVCEPEILIAD
PTTALDVTIQAIQILELMKELQQKKGTSILFITHDLGVVAQIADEVVVMYKGHVVEQASAK
ELFADPRHPYTKALLSAIPKPGKEYRKKRLETVDENVDYLSFPKELR
>PROKKA_00753 Oligopeptide transport ATP-binding protein OppF
MKLLEIKELKKSAYADRGLFKPKRVIHALNGISFEVEQNEVLSIVGESGCGKSTTAKILA
GIERQDSGAIYFNGKRHLHFSKQDWFDRKKVQMIFQDPYSSLNPRWKVGEIIAEPLLLN
SHFSKKEIKTKVLEIMQKVGLKLEWIDRYPHQFSGGQRQRIGIARALILHPSVICDEPV
SALDVSIQAQVLNLLDLQKEMGLTYIFISHDLGVVEHISDKIIVMNQGGIVETGDVDSV
ISAPKHPYTQKLLNAVPHLEKSMQRFK
>PROKKA_00754 GTPase Obg
MFVDSVEITIASGKGGPGMVSFRRKFFVIKGGPDGGDGGDGGDVYFEVDNNTDTLASFRG
TKHHKAKNGAPGGTRNCTGKKGEDKIIVPPGTQVFVGDKLWLDLVEPKRVLALKGGKG
GLGNAHFKSATKQQPTYAQKGLEGEKCVRLLEKLIADIGLVGFPNAGKSTLISTISNAK
PKIANYEFTTLVPNLGVSVDEKSEFLMADIPGIIEGASQGKGLGISFLKHIERTKVLA
FLDASRLDLGMKEQYQRLRLELEKFSALANKPFGVLLNKCDVVENIDEMAKDFCAFLNL
GAQKLNADFLEPHLGLFHPQLISDFENNPNEKSVLFVLPPLSAVSALNTHALKFVLEALP
>PROKKA_00755 Alginate lyase
MKRFVLFLSFMFVCVCVQAHIEQDYFFRDFKSKDLPKKLHLDKKLSQTIQPCTQLNASKH
YTSTGVREPDVCTKSFKKSALMSYDLALGYWVSKNKQYGLKAIEILNAWAKELQSVDTYQ
SEDNINFYMPYMNMAWYFVKAFPSPEYEDFIKMRQYSQSALNTNHGAWGILFDVSSAL
ALDDHTLLQNSANRWQEWVFKAIDEDGVIASAITRSDTSDYHGGPTKGIKGIAYTNFALL
AITISGELLFENGYDLWGSGAGKRLSVAYNKVATWILNPETFPYFQPNLIGVHNNAYFII
LAKHYSSPSADELLKQGDLDHEDGFRLKLRSL
>PROKKA_00756 YceI-like domain protein
MKKMVLVSVLLAGFLQAVNLDLSSAKLTWTAFTKAKTPVNGSFESITYKLGKSQDSLKT
LLEGASASMSDLKVNLDGDDIKNKNVKEAFFALFKNTNIKVTFRNVIEGDHAGSLTAYVRM
NEKLKVKPMQYTTAEDKLKLVKGVLDLLNFGGLKNELASLAKRCESFHEGLTWSQVEIQFES
MVKG
>PROKKA_00757 Glutamate-1-semialdehyde 2,1-aminomutase
MELLHSINDFNEAKQVIAGGVNSPVRAFKSVKGTPTFFILKKGAYLYDVDNNHYIDFVQS
WGPLIFGHADEEIEENIINTLKKGTSFGAPTELETTAKEIISCYEGLDKVRLVNSGTEA
TMSAIRLARAFSQKDDLIKFEFGCYHGHSDSLLVKAGSGCATFGSPSSSLGVPNDFSQHTLV
ARYNDLNSTEECFKKGDVGCVIIPIAGNMGLVPAQKEFLLGLKALCEKYQAVLILDEV
SGFRASLSGSQEFYGVVDPDLVTFGKVIAGLPLACFGGRAEIMDLLSPIGGVYQAGTSLG
NPLAVCAGLSALYKIKRDKTLYTRLNALAIRLTQGLKKSQNYNIALETLMGSMFGFFF
NENAVCDFDDALKSDTEMAFKFHQKMLFKGVYLACSSFETGFICEPMTEEMIDLAVAKAD
ESFDEIIGV
>PROKKA_00758 Putative F0F1-ATPase subunit (ATPase_gen1)
MKKPKYYKFIEGANYLSLGLSMVVAIIMGVAIGYGLKKLTHISWLFWLGVWGLASFLN
VYKAYKNMQKDYEELAKDPKYTQNKTK
>PROKKA_00759 hypothetical protein
MCQIQCLLILFFINIVSAIIIVYFFQAFQGVNLFEGGFLGFFIVALSSYYGVKKRLDLRKQ
NGEKEEKQKFQKQFALGLEMSFNVWRLGGYGVLLGILGALLFLHLFNLGSLFLIGVSVSSLS
SALLRFLNNNGKF
>PROKKA_00760 2-oxoglutarate amidase
MKTKNPAKRILKTAVIQMSKPYALNENLQALNLAKEAHNKGANLIVLPELFDSGYCVN
DKDADFGIDFKAIEHGELKSETLRALSNAFASNKVHLVACSIEKTDKKLYDSAYIIPPKG
GISGKHKRIYLWDEKSRFRGKKYEVFTLDFGDFSAKVGLQICYEIGFGVGANLLALQG
AEILIYPSAFGKARAYNWDLLSKARALENGCFVCACNHSGEETNAKLKQTLFAGDSRII
APNGKIIAQATKLNVEVIAEMDLNEAALQRQKIPYLQDFDTKLTKKGFGLT
>PROKKA_00761 Peptidoglycan deacetylase
MAKEILVAYGVDIDAVAGWLGSYGGEDSPDDISRGLFAGEVGIPRLKLKFKYHLPATWF
APGHSIETFPEQMKMIVGAGHEVGAHGYSHENPIAMSAKQEEDVLLKSVELIKDLTGKAP

TGYVAPWWEFSNITNELLLKHGFKYDHSMLMHNDFTPYYVRVGDSWSKIDYSLEAKDWMKP
 LIRGVETDLVEIPANWYLDLPPMMFIKKS PNSSFGFVSPHDIGQM WIDQFDWVYREMDYA
 VFSMTIHPDVSARPQVLLMHEKIIIEHINKHEGVRWVTFNEIADDFLKRNP RKK
 >PROKKA_00762 hypothetical protein
 MCVLCGELISSFHWTDGTYGSDSCEIDENLKEPNALISVNENARERKRARLKRVRLLNQI
 LAFYGLKINDWQGAKFVLCDDKKGSVIVNDLGLDWDKAQNLAKKEMDALDSHLLAFLNQNTSATH
 >PROKKA_00763 putative GTP-binding protein YjiA
 MPKIPITLITGFLGSGKTSFLSEYLNQIDHQGVALIINEIGQAALDQRILSVQYCGEKML
 YLNAGCVCCNKRLDLVESLKATLNNEYWRGEILKR IIIETTGLANPAPILWTLSDVFLG
 AHFEIQSVVACVDALNAKTHLTNNEAKEQIVFADSVLLTKTDLQND SRALIKLKERIQAL
 NPSAEIFDKKNIDYESFFSRKNRVRNFMPRMPKDSHSQGFETLSIRFEGAIEWSAFGIWL
 SLLLHQYGTQILRIKGIIDIGSDLLVSINGVMHVIYPPKHILKDQNGSNLVFIMRHLECE
 KILNSLKGFKDFLGIKFETP
 >PROKKA_00764 Major Facilitator Superfamily protein
 MCVFVCFLGVFVSNGLARFGYVVLIPLLILSGSLTPHQSFQLGIAVLMGYVFGSFLIQFL
 SPLMSLENIAKISFGLIALSFLICYFDSIPFFWLWIWRFIAGVASSALMILVAPLSLPYV
 KEHKKALVGGLIFSAVGIGSVFSGFVLPWISSYNIKWAWIFLGGSCLI AFILSLVGLKTR
 SLRKKSVKKEESA FKIPFHLWLLLVSALNAIGFLPHTLFWVDYLIRHLNISPTIAGTSW
 AFFGFGATLGSLISGPMAQKLGAKNANIFILILKSIACFLPIFFHQISLLNLSIFIMGAA
 TTANINLFSMMALKIAGAKHFAQASSWVVF SFGIFQALFSYLFTIFLGLGYVLIFVICG
 ACLVLSFIVLFP IKMQTATNK
 >PROKKA_00765 Pyridoxamine 5'-phosphate oxidase
 MLNRIIEHMNAHHVEDMKGLLKKFGQVHHAENVAFKSVDSQGIVIGYNNNQTLRIEFNHE
 VKDPKDYKNAIIELCQSVEKTHDLKGVEEEVKAFKESFDSVCLATLHPNGHVVC SYAPLM
 SDGKQYYIYVSEVAEHFAGLKNNPHNVEVMFLEDESKAKSAILRKRLRYKTNARFIERGA
 EFDKAFDSFIEKTGGAGGIKTIRTMQDFHLIALDFKEGRFVKGFQAYDILGDKIAYVGD
 KGNPHNFPHKK
 >PROKKA_00766 Arginine--tRNA ligase
 MHTLIKGVL EEIIEEEV IIEYPKDREHGHYATPIAFNLAKVFKKSP LVIAEELALKISTH
 EKTQGGFDSV VACKGYINFTLSLDFLERFTQKALELKERFGSQPKSEHSQKIFLEFVSAN
 PTGPLHIGHARGAVFGDSLAKIARFLGHEVLCEYYVNDMGSQIRLLGLSVWLAYREHVLK
 ESVTYPEVFYKGEY IIEIAKKANN DLEPSLFKENEET IIEVLSGYAKDLM LLEIKDNLDA
 LGIHFD SYASEKEIFKHKYAVFERLEKANALYEKDSKIWLRS SLYQDESDRVLIKEDKSC
 TYLAGDIVYHDEKFKQNYTKYINIWGADHHGYIARVKASLEFLGHDSNKLEVLLAQMVRL
 LKDNEPYKMSKRAGNFILIKDVIDDVGRDALRFIFLSKRLDTHLEFDVNTLKKQDSSNPI
 YYIHYANSRIHTMLEKSPFSKEEVLQTPLTNLNAEEKYLLFSALS LPKAIESSFEEYGLQ
 KMCEYAKTLASEFHRFY NAGKILDTPKAKELLKICLMVSLSLTNAFKLLGIEIKTKISAK
 D
 >PROKKA_00767 Sec-independent protein translocase protein Tata
 MGGFTSIWHWVIVLLVIVLLFGAKKIP ELAKGLGSGIKNFKKAVKDDEEEAKNEPKTLDA
 QATQTKVHESSEIKSKQES
 >PROKKA_00768 Guanylate kinase
 MHNDFNLLILSGPSGAGKSTLT KYLQEKIPKTHFSLSTTRKPREGEVDGLHYNFVSEEE
 FKQGIEKGQFLEWAI VHNHYGTSKIPVEKALKEGKIVIFDIDVQGHEILKKHYPNACSV
 FISTKNQEILKERLLL RGTDSKETIEKRLINAYKEMQCLESFDYLI INEDLEKSKEIILS
 IAKTLVHRLKAFNF EKICKAWKNESL
 >PROKKA_00769 hypothetical protein
 MKMILFNQNP MIAKLLESVSKKLELPMENFNHYQELSAHLKKDPEWLLIADDECLEKLDQ
 VDWLELKEI ISQNKNSVC MYKKGNEAQPFLEGFEVKIKKPFLPTEMLKVLQKKLGSNISE
 LEPSQNLDP TQEVLETNWDELENLG DLEPLAKEESNNEEQLLPTLDVQEKEEVKETPQEE
 KPKDDETQES ETPKDEEVSKELETQEEVKEETQEEQVKEQEPIKEETQEIKEEKQEETQD
 SPSAQELEAMQELVKEIQENSNGQENKEETQEIVETPQDVEVPQSQEKETPQEETQEVQE
 TQEKETPQTQDEK PQEDHYESIEDIPEPVMAMAGEELPFLNEAVAKTPNSENDAETPKE
 SVIKTPQEKEESDKTSSPLELRLNLQDLLKSLNQESLKS LLENKTL SIKITLEDKKNPA
 >PROKKA_00770 putative cardiolipin synthase YwiE
 MFNKFKKIVGVGLVGLGVLQAKNSL FVLPHYEQRDALNSLVSGISNARES VKIAIYSFT
 HRDIARAISKVASRGIKVQIIYDYESNHNKQSTIGYLDKYPNTKVCLLKGLKAKNGNYY

GIMHQKVAIIDDKIVFLGSANWSKNAFENNYEVLLKTDDTETILKAksyyQKMLGSCVGF
>PROKKA_00771 Helicobacter outer membrane protein
MRTLLKMLVGASLLTHALVAKEESAAPSWTKNLYMGVNYQTGSINLMTNIHEVREVTNYQ
TGYTNIITSVNSVKKLTNMGSGNIGLVMGYNHFFHPDKILGLRYFAFLDWQGYGMRYPKG
YYGGNNMITYGVGVDAMWNFFQGSFYQDDISVDIGVFGGIAIAGNSWYIGSKGQELLGIT
NSSAVDNTSFQFLFNFLGLKALFVDEHEFEIGFKFPTINNKYTTDALKVQMRRVFAFYVG
YNYHF
>PROKKA_00772 Flagellar L-ring protein precursor
MKKALCLGAVAFSVAFNMASANEPKIDFNPPNYVEETPSKEFIPELNKLSLFGQGERPL
FADRRAMKPNDLITIIIVSEKASANYSSSKDYKSASGGNSTPPRLTYNGLDERKKQEAeyL
DDKNYNFTKSSNNTNFKGGGSQKKSEDLIVLSARIIKVLENGNYFIYGNKEVLVDGEK
QILKVSgvirpydiERNNTIQSKFLADAKIEYTNLGHLSDSNKKKFAADAMETQMPY
>PROKKA_00773 CMP-N,N'-diacetyllegionaminic acid synthase
MRAIAVVLARSSSKRIKNKNIIDFFNKPMLAYPIETALNSKLFEKVFISSDSMEYVNLAK
NYGASFLNLRPKNLADDRATTLEVMAyHMKELKDEDIACCLYGISVFLQEKHLKNAFE
TLKENTDYVFTCSPFSASPYRSFSLENGvQMAFKEHFNTRTQDLKTLyHDAGLLYMGKAQ
AFKEMRPiFSQNSIALELSPLEVQDIDTLEDLAKLKYSRLKNACQ
>PROKKA_00774 hypothetical protein
MPVKILCDSFLTSGLGHVRRCEKILSFIEKLGIKADLYLYKHNDIGAFLEGVGNDDLLIT
DSYRLNSKDFYLLKEKAKSLMVIEDDEEHAKGFYPKNTKIMNFTLNALKHYNHASKDHYLG
VGfYPVDTRFIYDRPINTENKEVLITLGGSEQEMLKEIVKILENKNVNLHIISPYIPKNP
PKNARYYSPLSPLEFSSLMKFCACAISAAGQTLyELALSQTPSLILPIASNQILQSQEF
NSGIFKQTSKTLARDFEKLQIQKNQAWAKNLAFGSELEGALREFLEI
>PROKKA_00775 UDP-4-amino-4,6-dideoxy-N-acetyl-beta-L-altrosamine N-
acetyltransferase
MKKNYSYGNiQAIDFTNLNDGEKLLVLEFRNHPNTALWMySAFISLkTHLQFIEDLKNSP
NHcyFLfKEEGVYLGVGSITKINFFHRHGYLGIYKNPFLKNKGAILKALERIAFEFQ
NSLHLEVmenNFKAIAFYEKNHyeLEGRlKGfIFKEKEFIDVLLYYKDKKGYNDQFSLKF
>PROKKA_00776 Tetraacyldisaccharide 4'-kinase
MKSDKPFLEryfyDPTLLQKGLIFALYPFSliYQGIATLKRKTAKKHDFKIPLISIGNLI
AGGSGKTPFILEIAPRYQeVAIVSRGYQRDSKGLVVSVKGNILVSQKTAGDEAYLLALN
LKQASVIVSEKRELGVlKALELGAKIVFLDDGFRFNFNQFNLLLKPVPpyPFCPLPSGL
YRESIKSYKEAHLIITEDKDYQRITSISHHTKRMILVTAIANPSRLDAFLPKEVVKKLYF
RDHALFDLELLEKEFEHNNNATSLLVTSKDLVKLQDCNLPLSVLDLkLEIDPKILERIDRY
ILSYPCNTKERL
>PROKKA_00777 NH(3)-dependent NAD(+) synthetase
MQKDYQKLIAyLCDFLEKEVQKKGFKKVVYGLSGGLDSAVVGVLcQKVfKENAHALLMPS
SVSMPEskTDALNLCETFSiPYTEYSIAPYDKIFGSHFKDASLTRKGNFCARLRMAFLYD
YSLKSDSLVIGTSNKSERMLGYGTlFGDLACAINPIGELFKTEVYELARRLNIPKKILNK
PPSADLFVGQSDEQDLGYPYSMIDPLLKDIEALFKNKPIDAETLTQLGHDEILVKNIISR
IQKNAFKLELPTIAQKFNPk
>PROKKA_00779 Ketol-acid reductoisomerase
MALPVYYDKDIDLGVlQSLQVGIIgYGAQGEAQALNLRDSKVVRIGLYQGSLSVSKAKA
EGFEVLEVkelVQNSDVIMVLLPDELHKEVLEKEVIPFLKEGQIVGFAHGFSVHFNQVVL
PKGVGAILVAPKGPgSALREEYLKNKGLYHLIAIEQESSKNNAKAVALSYAKAMGGGRMG
VLETsfKEECESDLfGEQAVLCGGLEAIVRMGFETLIKAGYPEELAYFECVHEVKLVADL
LHYKGVEGLRKHisNTAEFGAIKHREPMGKlLEKRMQKILKKIQNGSFAKDFLLEKSLNY
PRLNTERKALKETKIEQIGeILRAPFNHKK
>PROKKA_00780 Septum site-determining protein MinD
MAIVVTITSGKGGVGKSTTTANLAIGLAESGKKVVAVDfDIGLRNLDmILGLENRIVYDV
VDVMEKNCNLSQALITDKKTKNLSFLAASQSKDNILDKEKVAILINALRADFDYILIDS
PAGIESGFEHAILHADMALVVVTPEVSSLRDSDRVIGIIDAkSNRAKRGEeVHKHLIINR
LKPELVANGEMiPIEEVLKILCLPLIGIIPEDHHVISATNKGEpVIRTDCEsAKAYQRIT
RRILGEEVEYVEfKAKRGFFSALKGIFS
>PROKKA_00781 Cell division topological specificity factor
MSLFDFfKNGSAAATATDRlKLILAKERTLNLPYMEEMRKEIIAVIQKYTKSSDIHFkTL
DSNQSVETIEVEIILPK
>PROKKA_00782 hypothetical protein

MKSHFQYSVLENIPKAFDILKDPPKQLYCVGDIKLEAPLKVAIIIGTRRPTPYSKQHTIT
LARELAKNGAVIVSGGALGVDIIAQENALPTTIMLSPCSLDFIYPTNNHKVIEIAQNGL
ILSEYEKDFMPIKGSFLARNRLVIALSDAVIIPQADLKSGSMSSARLAQKYQKPLFVLPQ
RLNESDGTNELLEKGOAQGIFNIQNFINTELLKDYHLKEMPEMEDEFLEYCAKNPSYEEAY
LKFGDKLLEYELLGKIKRINHIVLA
>PROKKA_00783 Putative Holliday junction resolvase
MILACDVGLKRIGIAALLNGVILPLEAILRHNRNQASRDLSDLLREKNIQVLVVGKPNE
YADTNARIEHFIKLLDFKGEIVFINEDRSSIEAYENLEHLGKKNKRLAIKDGRDLSLAC
RILERYCQQVLKCC
>PROKKA_00784 Putative beta-lactamase HcpC precursor
MLENFKKAIFRALCLCCVLMGGLMAEPNPEELFDLGMLSYNKQDFSKARKYFEKACNLKE
GGGCGALGDLYENGEGVEKNSKAAQFYKACELKDGRGCGALGRLYYQNGQGVEKNLTK
AAHFCSKACDLNDSEGCKNLGSLYNGEGVKQDSKAAALFEKACKLGDQEACEMLKELK
>PROKKA_00785 hypothetical protein
MKIMHQVMDFLDSCQTIQKNSNEALQKTKRALFKEKSQKIREQAVKIVEKRLKKEGMKL
SDFNEEELKIMFEAEKRLLEQIHAKSEENQEKSARKHFKEVWERGDNEQER
>PROKKA_00786 hypothetical protein
MSKKDSFISSVMGFFSEKNERWLLAHRHTRGFVIVAWLFRFKSIAFSILITLLVILVDIW
VYSDVRQFLDLDASSPILLIAALLIKWGVIVISVRQCYKFSKKMFALIQKKQIRESLLK
RSHQQEGKTPPKERLSSVTEEIIIAKKQEETQRQESLEDTHERLSSTTEEIIISKLEELKA
RNDKGN
>PROKKA_00787 tRNA pseudouridine synthase C
MPFVEEEFEILKPTKALFFVRDVLKCSLKEAQRHLQKQAVRKSQMIQGVVSLI
HFKPNEKTQALVFETKDFGVFDKPPQVYTHPKGYFYHESLLDCIQSHFGKNAHPAHRDLY
ETSGVLVLAGKTLQSAKDLKALFMQKVKKTYLALAHGLVDKNIVIDKPILTPQNIQKDLR
IRSKISPLGKPSTTLVEPLSYNPFLLDISLLKITPITGRTHQIRLHLSSVGHRIVGEGLYG
VIDENAREYLQLKRENNAPLLMLHASSLEFEFFKAFYEITSPMPKRFMPFLKDLALFY
>PROKKA_00788 Single-stranded-DNA-specific exonuclease RecJ
MKQKLKAQIKERVASIAAYNEKGFPSPFLFKDLKKAALKIIEAMRANTEILVVGDDYDADGV
ISSTIMAKFFESLNKHYVRIAPNRFIDGYGISKKFLEKHHAPLIITVDNGISAFEVARF
CKEKNYTLIITDHHCLHHDEVPDAYAVINPKQPDCEFIQKEVCGALVAFYLCYGIHQLLG
KEKSHSSELLCLAGVATIADMMPLTFFNRFVSKALYFLQKESLGAMGFLRQREVFRKRS
LKASDISFNIAPLINSAGRMQDAKMALDFLSANNFQDGYFLYERLKACNMKRKMIQQQVF
EEAFKHAMVGEKIIIVAFKDDWHEGVLGIVASKLVEATQKPSLVFTFKEGVYKSGSRSSSN
IDLIDALNGVSSLLGYGGHRQACGLSVGKNNIISLFETLENFDFKILPFCEKEPPLTLA
LKDIDRELLEIIIEAGEPYGQENPEPLFQAQNLEVIEEKIIESHQVLRFKDKECVKEAII
FNTERFLKAGEKVSIFSVLEDEYSNEPKMFVKSLL
>PROKKA_00789 CTP synthase
MDRAKFIFVTGGVLSSLGKGISSSIATLLQHCNYQVSILKIDPYINIDPGTMSPLEHGE
VFVTSDDAETDLIDIGHYERFLNRNLTRLNNFTTGQIFSSVIENERKGEYLGKTIQIVPHV
TDEIKRKISAAKGLDFLIVEVGGTVGDMEGMFYLEAIRQLKLELGNEKVINHVHTLIPY
IQTTELKTKPTQHSVQELRRLGVTPQIILARSPKPLDKELKNKIALSCDVEQDSVIVAT
DTKSIYACPIFLQEGILTPIARRFNLNKLHPKMAAWNTLVEKIIAPKHVKVIGFVGKYL
SLKESYKSLIEALIHAGAHLDQVNIWLDSENFNEKTDLEGVDAILVPGGFGERGIEGK
ICAIQARARLEKLPFLGICLGMQLAIVEFCRNVGLKGANSTEFNQRCEHPVVYLIEDFID
QNHQKQVRYTNSPLGGTMRLEGEYECIEVPNSLLEKAYKKPSIKERHRHRYEINPKYRQEW
ENKGLKVVGFGANHLIEAIELEDHPPFFVGVQFHPEFTSRLQSPNPILDFIKSALHKS
>PROKKA_00790 hypothetical protein
MDNAIFNFARSTPFNSSPILTLILQNIANLGSSQFVLPLSLLVGVFLSLYRRNLALGVWF
VLSVVFIEALLESKLHLLAHSIQWLSHSANFPSAIALSLTFYGLLVLLLPHLITHQIFQ
NILSCSLLGLIFLIGLALIVLGVSFSSVLGGFCLGALGACFSIGIYLSVFQKI
>PROKKA_00791 Flagellar M-ring protein
MDLKVLLQRIVDFFIKLNKKQKIALIAAGVLITALLVFLLLYPFKEKDYTQGGYGVLFE
LDSSDNALILQHLQQNQIPYKVSDDTILIPKDKVYEERITLASQGIPTSKVGFEIFDT
KDFGATDFDQNIKLIRAIIEGELSRTIESLNPILKANVHIAIPKDSVFAKEVPPSASVML
KLKPDMLKSPTQILGIKNLIAAAMPKLTIDNVKIVNENGESIGEGDIPENSKELALEQLH
YKQNFENILENKIVNILAPIVGGKKNVVARVNAEFDFSQKKSTKETFDPNNVVRSEQNLE
XKKEGTPKKQVGGVPGVVSNIQPVQGLKDNKEPEKYEKSQNTTNYEVGKTISEIKGEFGT

LVRLNAAVVVDGKYKIALKDGTNTLEYEPLSDESLOKINALVKQAIGYNQNRGDDVAVSN
FEFNPIMPMLDNATLSEKIMHKTQKILGSFTPLIKYVLVFIVLFIFYKKVIVPFSERMLE
VVPDEDKEVKSMFEEMDEEDELNKLGLRKKVEDQLGLNATFSEEEVRYEIILEKIRGT
LKERPDEIAMLFKLLIKDEISSDSAAG

>PROKKA_00792 Flagellar motor switch protein FliG

MATKLTTPKQKAQLDELSMSEKIAILLIQVGEDTTGEILRHLDIDSITEISKQIVQLNGTD
KQIGAAVLEEFFAIFQSNQYINTGGLEYARELLTRTLGSEEARKVMDKLTQSLQTKNFA
YLGKIKPQQADFIINEHPQTIALILAHMEAPNAAETLSYFPDEMKAIEISIRMANLGEIS
PQVVKRVSTVLENKLESLSYKIEVGGLRAVAEIFNRLGQKSAKTTLARIESVDNKLGA
IKEMMFTFEDIVKLDNFAIREILKVADKKDLSLAKTSTKDLTDKFLNNMSSRAAEQFVE
EMQYLGAVKIKDVDAQRKIIIEIVQSLQEKGVITGEEEDVIE

>PROKKA_00793 flagellar assembly protein H

MNSRKNLIQKDHLNKHDIQKYEFKSMANLPPKTNPNGASLETPNPQEPLEKKAIENDLID
CLLKKTDELSSHLVKLQMQFEKAQEESKALIENAKNDGYKIGFKEGEEKMRNELTHSVNE
EKNQLLHAITALDEKMKSENHLMALKEKLSAIAIDIAKEVILKEVEDNSQKVALALAE
LLKNILDATDIHLKVNPDPYPLNERLQNASIKLESNEAISKGGVMITSSNGSLDGNLM
ERFKTLKESVLENFKV

>PROKKA_00794 1-deoxy-D-xylulose-5-phosphate synthase

MQNKTFDLNPNDIAGLELVCQTLNRNILEVVSANGGHLSSSLGAVELIVGMHALFDCQKN
PFIFDTSHQAYAHKLLTGRFESFSTLRQFKGLSGFTKPSSEAYDYFIAHSSSTSISIGVG
VAKAFCLKRALGMPIALLDGSGISAGIFYEALNELGDRKYPMIMILNDNEMSISTPIGAL
SKALSQLMKGPFFYQSFRRSKVKILSTLPESVNYLASRFEESFKLITPGVFFEEELGINYIG
PINGHDLNAIIETLKLAKELKEPVLIHAQTLKGKGYKIAEGRYEKWHGVGPFDLDTGLSK
KSKSAILSPTAYSNTLLELAKKDEKIVGVTAAMPSTGLDKLIDAYPLRFFDVAIAEQH
ALTSSSAMAKEGFKPFVSIYSTFLQRAYDSIVHDACISSLPKLAIDRAGIVGEDGETHQ
GLLDVSYLRSIPNMVIFAPRDNETLKNVHFANEHDSSPCAFRYPRGSFALKEGVFEPSG
FVLGQSELLKKEGEILLIGYNGVGRAHLVQLALKEKNIECALLDLRFLKPLDPNLSAIV
APYQKLYVFSNDYKLGGVASAIIEFLSEQNILKPVKSFEIMDEFIMHGNTALVEKSLGLD
TESLTDAILKDLGQER

>PROKKA_00795 Elongation factor 4

MKTKAPMKNIRNFSIIAHIDHGKSTLADCLIFECNAISNREMKSQVMDTMDIEKERGITI
KAQSVRLNYTFKGEDYVLNLIDTPGHVDFSIEVSRSLCSCEGALLVV DATQGV EAQTIAN
TYIALDNNLEILPVINKIDLNPANVLEV KQDIEDTIGIDCSNANEVSAKARLG IKDLLEK
IIT TIPAPSGDFNAPLKALIYDSWFDNYLGALALVRIMDGSINTEQEILVMGTGKKHGV
GLYYPNPLKKIPTKSLECEIGIVSLGLKSVTDIAVGDTLTD AKNPTFKPIEGFMPAKPF
VFAGLYPIETDRFEDLREALLKLQ LNDCALNFEPESVALGFGFRVGF LGLLHMEVIKER
LEREFGNLNIATAPT VVEVHLTDNSIKYVQNPSELPPENHIACIKEPFVRATIITPSEF
LGNLMQLLNNKRG IQEKM EYLNQSRVMLTYSLSNEIVMDFYDKLK SCTKG YASF DYEP
ENREAH LVKLDVRVAGDVVDALSIIIDKNKAYEKG RALVETMKELIPRQLFEVAIQASVG
NKIIARETIKSVGKNVTAKCYGGDITRKRK LLEKQKEGKKRMKAIGKV ELPQEAFLAILK
ID

>PROKKA_00796 hypothetical protein

MAHKHNYEILNLIGYALAKFDNDFIKEFGFSTKNAFFEYCVQIGLADTTGVIKNRMDLFD
YFFPNKRKGWWQKGDAYIHRKLWIDSLFGNESVKGF SHIMKWFLQE QYGIKDLGITPNAY
LKTRYKSMQESGLEAE LYFLNHYKNIKIFSCGHLKDMRLFGDGYDFYIQTNKQ AFLVEVK
GIREKQGT LRLTQKEYEQAQTYSHDYVLVVVLNLSEKPHLLSIANPLKHLEFKACERKQK
SILEYHLIGQIK

>PROKKA_00797 hypothetical protein

MREEAKKHFRIPLN RANKITL NFTGGYRSGVQIDRNAPKRTYKYTKKDCDLILGIDTRTS
ECYIIPIEDTQEWGNTKSLSQLQHYKENWQILIDLAL E

>PROKKA_00798 Flagellar basal-body rod protein FlgG

MQNGYYATTGAMATQFNRLDLTSNNLANLNTNGFKRDDAITGDFLRLYQQYREQLPLEDQ
TKASAKYLNRLNLRVPVLSEIYTD RSLGAFEGTHNPLDFALTSPNLYFAIQTN EGVAYTR
DGHFSVDKDGFLVTLNGFRVLRSRGLNEKGGIMLPNAEIEVNQNGGITFRDNEAQIQAG
ALALVSFSEPKNLKKIGQNLTYTQGEVGHQVSDSGALRQYMLEKSNVNAVREMSALIEIN
RFLDMYSKVLKTHQDDMNAEAINKLATKA

>PROKKA_00799 Alpha-ketoglutarate permease

MSSQTATKKPLKSLLAASSGNLVEWYDFYAYAF LOPYFAKEFTHTNDPTLALISAF LVFM

LGFFMRPLGSLFFGKLGDKKGRKTSMVYSIILMALGSFLLALLPTKEIVGEWAFLLLLA
RLLQGFSVGGEYGVVATYLSELGKNGKKGFYGSFQYVTLVGGQLLAIFSLFIVENIYTHE
QISAFAWRYLFALGGILALLSLFLRNIMEETMDSKTTSKTTIKEETQRGSLKELLNHKKA
LMIVFGLTMGGSLCFYFTFVYLKIFLTNSSSFSPKESSFIMLLALSIFYFLQPLCGMLAD
KIKRTQMLMIFAITGLIVTPVVFYGIKHATSVHEALFYEILALSSMSFYTCIAGVIKAE
FPEHVRALGVGLAYAIANALFGGSASYVALEFKQHGFAGFVGYVMLSIIIFMVMVIFP
KKTYLE

>PROKKA_00800 DNA translocase SpoIIIE

MKSKKLYLALIIGVLLAFLTLSSWLGNSGLVGRFGVWFAALNKKYFGYLSFINLPYLAWV
LFLLYKTKNPFTEIVLEKTLGHLLGILSLLFLQSSLLDQGEIGNSARLFLRPFIGDFGLY
ALITLMVVISYLILFKLPPKSVFYPIYINKTQSLLEKIYKQCLQAFSPNFSPKKEGFENTP
PDIQKKETNNDEKENRKENPIDENHKTPNEESFLAIPTPYNTTLNDLKPQEGLVQISSH
PPTHYTIYPKKNRFDDL SNPTNPTLKEIKQETKEREPTPKQETLTPTTPKMPASAPNTE
NDNNTQNHKAPNHPTKEENAQENTQKEMIEERIKENLKEEEKEAQNAPGFS PETPTS
AKK PVMVKELSENKEILDGLDYGEVQKPKDYELPTTQLLNAVCLKDTSLDENEIDQKIQD
LLS KLRTFKIDGDIIRTYSGPIVTTFEFRPAPNVKVSRIILGLSDDLAMTLCAESIRIQA
PIKG KDVGIEIPNSQSQIIYLREILESELFQKSSSPLTLALGKDIVGNPFITDLKKLP
HLLIA GTTGS GSKSVGNAMILSLLYKNPPDQLKLV MIDPKMVEFSIYADIPHLLTPIIT
DPKKA I GALQSVAKEMERRYSLMSEYKVK TIDSYNEQALNNGVEAFPYLIVVIDELAD
LMMTGKKE AEFPIARIAQMGRASGLHLIVATQRPSVDVVTGLIKTNLPSRVSFRV
GTKIDSKVILDTD GAQSLLGRGDMLFTPPGANGVLRLHAPFATEDEIKKIVDFIKAQ
KEVQYDKDFLLEESRM PLDTPNYQGDDILERA KAVILEKKITSTSF LQRQLKIGYNQA
ATITDELEAQGF LSPRNA KGNREILQNF

>PROKKA_00801 PD-(D/E)XK nuclease superfamily protein

MNLEKLFLEKTPLFVFSSTRRLKHFYLEQGEGLPNAMSMGSFFEQA FYIPNKKKIPNST
RLILMIDTIKAIKEKKSILEGLLLFENSFLGYLESTSF LFDLDELSSACIKLNELSSK
DIYLDYEKHLEVLEMIYDRIYK KLEELGFYDKIMQKKPTILKEFFEHEFSSIEWHLDG
FMS VFERQCLLEVAELVPITLHLSCDKYNQKFLEFLNLKLETDCDYSIDFKTQKILSQT
PKRQ KIEPKLYANSSYLKQSALVLQTIEEYLQKDNDPNKMAIITPNADFLPFLKLLD
KNNNLNF AMGLGAKNSPYYTELVKILEDLETSGFDWSASPLLDLENLTLALLEQQSS
KEKAPLKEVH SQIMHQYHLLKDTLQNYNLKDLLHLYLQEF EANFRLLDSSGGKIRVMD
TLETRGMQFDKI VIVDFNETCVPSLKDCDLFLNSALRQSLNLP TLLDKNLQKHYYYQL
FKNSKEIVLSYIE SETSKASNMLELNLYIEPIKDAYTLFETSPLKDYQEEIEKATIPK
DFSFSASSLNAFLT CKRRFYHYMKRFKESPKDENNSAVGSL LHELLKETYEKDKNPYA
LEERLIQ LLETRENI TPKERLDTLIALKKIQAFYIKEKERFKAKIKILDLEKS FETTIQ
GVAFKGRIDRIDKTAD NEVILLDYKFKSDLKLDNMSKTQREGLSPIEIAQISTDYQMTIY
VFALKNLGYKEPIKAF FYDLRKGEELLEDELTLQAKMDHLEFSLIPK LKQEIDFEKT
LEAKDCEYCSFKDMCNR

>PROKKA_00802 Transketolase

MRLSNADLERLKS MANTLRFLCADMIDKANS GHPGVCLGLADVMV VLSLHLNLNPINPKW
LNRDRLVFSGGHASALAYSL LHLWGFDLSLEDLKHFRQLH SKTPGHPELHHTEGIEITTG
PLGQGFANAVGFSMASQYAQNLLDKEVISHKVYCLCGDGLQEGISYESASLAGHLRLDN
LIVIYDSNQISIEGAINISFSEQVKTRFLAQNWEVLECDGHDYQAIHNALEEAKKSHKPT
LLIAHTIIGKGA VGLEGSEKTHGSPLNKEVLKQSKENAQINPDESFIISPKNKMHFEV
KVRGASLEALWEKSLSPKTKEKIHAKDFDFSTIHYPTFKKGESLATRVSNGMILNAIAKE
CEGFLGGSADLAPSNN TQLKHSGDFPLGQNLHFGIREHAMGAITNALAAYGLFLPFCATF
FVFS DYLMPSMRLSALMKL KALFIFTHDSIGVGEDGATHQPIEQLSHLRALPNFYAFRPS
DAFENKACMQVALSLNAPSGLILSRQNL PVLDEVSKERV LKGAYVKYDAKDPIITLVASG
SEISLALES AKMLERENIPTQVVSAPCFD LLLIEQDES YLKELFKGKVLVIEASRAIEWYR
FADKIIIGMDSFGSSAKGDKLFEKFGFSVENITAQAKRLLNA

>PROKKA_00803 Riboflavin biosynthesis protein RibF

MLNFLSISSEHEVKS LAIGKFDGLHLGHQALFKELKDPKALLIEKKHYTKGYLTPLKYR
AKLVGMPLFFVYLEEISQLNALEFLELLKKKFPHLERLVVGYDFRFGHGRQNDALFLKER
FEKTIIVPEVKIKEISVH SKTIKLALSHGDL SLANKLLGRPYEVC GKVISDQGLGHKELV
PTLNIKTKDFILPSFGVYASLVKIKDP IYQKSVSFIGNRLSTDRNFAIECHVLDTIIEKP
PKEIALRWVQKIRD NMRFSS LKELKNQIQQDILRAKEILR

>PROKKA_00804 16S/23S rRNA (cytidine-2'-O)-methyltransferase TlyA

MRLDYALFSQHVLVNSREKAKALVLKNQVLVNKMVVS KPSFMVKEDDKIELIATNL FVSRA
GEKLGAFLETHFVDFKGVVL DVGASKGGSQVALLKGAKKVL CVDVGKMQ LDES LKNDQ

RIECYEECDIRGFKTPETIDLVLCDSFISLYCILEAILPLSDEFLALFKPQFEVGRAAK
RNKKGVVVDKEAILNALENFKNHLKTKDFQILKIQESLVKGKNGNVEFFIHFKA
>PROKKA_00805 hypothetical protein
MKNLSTLLVFLFFCLGCVSNFNETYTLDLVLEKKIQASRKEITQDNVPIITAIATHLN
DVDSGTYDDHEYFLVEIFTQNNWDIDGYISYELFGTKPTGSEPLWMREITKDEFDGILE
TTNRWSRAFLAFDKLDYLAVQEAKLELDAYSISLGKIVFNFAFQVPLPQF
>PROKKA_00806 Aspartate carbamoyltransferase
MPKKCRHLLQTSDSLDEIKLLLNKASVYANDFNAVSLKETKMYNKIIVALFFENSTRT
VSSFEIASLRGAKIVKLNMQTSSTSKGETLTDTFKNIHAMQPDAIITRHAFSSAPFKLA
EFSQCPLINAGSGVSAHPTQALLDLLTLRYHFGSLENLKGKKIAFIGDVKNSRVANSNIK
LLQRLGLEIMLCAPSSMLPSVSLKTTTHNIEEAIADFADILMSLRTQTERHNAPIFASLKDY
GNAYCITQORLKAHAKNKEVIIHHPGPVHRDIDIESAVLEDKRSKVLEQVKNGVAMRMAV
LEFLLLD
>PROKKA_00807 Putative outer membrane protein
MKNSAPLKNQVFRGLYVLSLSASLQAFDYKIEVLAESFSKVGFNKKKIDIARGIYPTETF
VTAVGQGNIIYADFLSKDLKDQGHVLEGGVGGTLGGIAYDSTKFNQGGSVIYNYIGYWDGY
LGGKRALLDGTSLHECALGSDGKVIDSIACGNARANKIRRNLYMNAFLEYRYKDIFLAK
GGRYQSNAPYMSGYTQGFESISAKVKDKNEGSHKLWWFSSWGRAFAYGEWIYDFYSPRTVV
KNGRTLNYGIHLVNYTYERKGVSVSPFFQFSPGTYSPGVVVGYSNPNFNGVGRSETK
AYILLPVHAPLRDRTYRYAIKAGTAGQSLLIRQRFDYNEFNFGGAFYKVKNNANAYIGTT
GNPLGIDFWTNSVYDIGQALSHVVTADAVSGWVFGGGVHKKWLWGTLWRWTSGLTANEAS
AAVNVGYKISKSLTASVKLEYFGVMTHAGFTVGSYRPTPGSKALYSRSHLMTTLSAKF
>PROKKA_00808 Lipid A export ATP-binding/permease protein MsbA
MKLFFKRYSKYLKEHYKSFVVVLFSSSLVVALSTAWGTYLVKPTLDEIFINKDTHMLKILP
FLVILAYLGKSGMYLGTFTNFIFGLDIVKKIRNTMLESLLKMEMDFFNRTKKGELIARI
TNDIGLIRASLSNYLSESLREGLTIVGLVGVIYQSPKLALVGLVIMPLAAIPISKIIRK
VKKLAKSHQESNAKITARLSEVFNNVEAIKISNGEKLHRAFKENEAFKIGIKNIABA
EISSPLMEFLGSIAIALVIYLGNGEVIRGYISVGAFSSFITALFMLYTPIKRLTRIVSNF
QEALVASDRIHEILEREPAIVDGELTLNNAIHTIEFKKVLAYTLDNQEHYVLSDISLKF
QQNEIIALKGESGSGKSSLVNLILRLYEPSKGEIFINDQKIESITQKSLREKISVVTQRV
FIFNGSVAENVAYGLEIDEVKIKECLKKAQALDFVEKMPHGIESVLDEFGANLSSGGQRQR
IAIARALYKDAQVLIFDEATSALDNNTESVKQSILELKQNRLLIILISHNPSTLKLATKH
VKLEHGRLIEC
>PROKKA_00809 hypothetical protein
MLRILSVGVVFLVLLGCQFFNKTTLHLKYKDYPKNSALKTAFTLTPPKIFFNARFVPPFYQ
KEFKKALAQQIAYFLKDKSAFTFNVSGNVFFSFEENPKDLKAIKERLKKTIENADPKSV
MRFLNLQASLILECVPQTTCFPTDILLIPTAFSVPVYANRLGDNPSLFSQEDKTYHNALI
KALNKAYYSLMEGLEKRLNAIENAEWL
>PROKKA_00810 Restriction endonuclease
MKKIAFFIFVILFSVGIYLIWHVLLQKALELKLVTSSANDLLLKLLSILGVFSMLVLFQGV
ISSYKKRQLKRVLQKIDAMNGFEFEEYSKIFFTSKGFVSITQKSGDYGADLIEKDGVK
WAVQAKRYSHKVSPKAIQEVVSSKAYYACEKACVITNSYFTQAAQKLAQANEVLLIDRDE
WVRFLNEKR
>PROKKA_00811 ATP/GTP phosphatase
MIQSVRIKNFKNFKNTKIDGFTKLNIITGQNNAGKSNLLEALYYLVGKSMHPCTNVLEIY
DNIRKEPLTSESXSLMFYGLDTEKKIQITTTLENNQTLDLQIKFIANEKQKVIESQIIPT
VEHAQEF SRLNFTLKKNNREIYNDHLNITKIPNFPPIPNQSGYNRQFKNFEP SQLQELLP
FESTAIITPSDVARKRDIMPNNAIHIMDPNIIQAMRKILDENQLEKELNERLNQFDKSIQ
AIRFSANNQLKLKVGIREKIPLSMFGDGLIKYLYIVSTFIANNAKTIYIDEVENGLHFS
RMGLLLKNIIDFINNNKEGNLQVFMTHNQEFIEILDQVIREKDFAHQTKLFCLEQYNGS
IVAEPYYGENLSLYFKNHANLFGGKERFKENNYE
>PROKKA_00812 hypothetical protein
MSKKTLIYTEGKSDKNFLSWCLNVWKNEDHDFDRDHFIIYVEGKNKLSDEFCKRIENIL
NNKDQAYKQVCIIFDADIKEENQESDAGFDNKLKHIRKEFKGIDFPKEQIFLFPNNQD
DGDLETLLLEIARHDEFKLCFERYLECIKNKEHYKPIKDIRKNMLYAYLEALGLENLTKT
NIDVFDSKGKIKEKHQEEYEKLKEVIDFNSNSLIPLKNFLGQFAENNQKTNPKIF
>PROKKA_00813 High-affinity nickel-transport protein NixA
MKLWFPYFLAIVFLHVLGLALLFMANNASFYAAASMAFMLGAKHAFDADHIACIDNTIRK

LTQQGKNAYGVGFYFSMGHSSVVILMTIISAFIAWAKEHTPILEEIGGVVGTLSGLFL
 LIIGLLNAIILIDLLKIFKKSHSNESLSQQQNEEIERLLTSRGLLNRFKPLFNFVSKSW
 HIYPVGFLFGLGFDTASEIALLALSSSAIKVSVVGMLSLPILFAAGMSLFDTLDGAFMLK
 AYDWAFKTPLRKIYYNISITALSVFIALFIGLIELFQVVSEKLHLKFENRLLNTLQSLEF
 TDLGYLLVGLFVIAFLGSFFLWKIKFSKLES
 >PROKKA_00814 hypothetical protein
 MDILKTLQKHLGGVETSDFKTNAIEKSQQIAKFSRDMKNINESVGALQVLQIACKKLFNK
 SMGLEDKDALQASIIKQELREIVKNCQFLASPLFDTQLNITINDEVFSMIVANPLDLLEN
 AGEFQAYLEEKLNIEIKELLSYLSESLSNPKAFMPSFSNQSLKDLLSDNLRA
 >PROKKA_00815 hypothetical protein
 MKKIWLLVWGLYSLVFLHAETIEKAPTNVEDRDKAPHLALLLAGIQGDEPGGFNATNLFL
 MHYSVLKGLVEVVVPLNKPSMLRNHRGLYGDMMNRKFAALDKNDPEYPTIQEIKSLIAKPS
 VDAVLHLHDGGGYRPIYIDATLNPKRWGNCFIIDQDEVKGAKFPNLLAFANNTIESINA
 HLLHPIEKYHLKNTRTAQGDTEMQKALTIFYAINQKKSAFANEASKELPLASRVFYHLQAI
 EGLLNQLNIPFKHDFELNPSSVHALINDKNLWAKISSLPKMPLFNLRPKLNHFPLPNNTK
 IPQIPIESNAYIVGLVKNKQEVFLKYGNKLMTRLSPFYIEFDHSLSEEVKMQIDNKDQMVK
 IGSVVEVKESFYIHAMDNIRANVIGFSVSNESKPNEMGYTIRFKDFRKRFSLDKQERIYR
 IEFYKNNAFSGMILVKFV
 >PROKKA_00816 hypothetical protein
 MPDIRNLYLSPDLTERLSKISNRLYVMHSSCSFSTREYRVLEIMHEIIQKGM DGLTQRL
 DNDLKDDCSRILVPIECIQGLCLVGNMKPDEL CMLFVKICELICSQCRSNLSRSSLSFIG
 GVISGLNKQIRLYQKWK
 >PROKKA_00817 COP-associated protein
 MKVTFQVPSITCNHCVDKIEKFVGEIEGVSFIDASVEKKS VVVEFDAPATQDLIKEALLD
 AGQEVV
 >PROKKA_00818 Copper-exporting P-type ATPase A
 MKESFYIEGMTCTACSSGIERSLGRKSFVKKIEVSLLNKSANIEFDENQTNLDEIFKLIE
 KLGYSPPKTLAKEKKGFFSPNVKLALAVVFTLFVVYLSMGAMLSPSLLPKSLLTIDNHSN
 FLNACLQLIGTLIVMHWGRDFYIQGFKALWHRQP NMSSLIALGTSAALISSLWQLYFVYT
 NQWSYGHYYFESVCVILFMFVMVGKRIENVSKDKALDAMQALMKNAPKTALKMQNNQQIEV
 LVDSIVVG DILKVLPGSAIAVDGEIIIEGEGELDESMLS GEALPVYKKVGDKVFSGTFNSH
 TSFLMKATQNNKNSTLSQIIEMIHN AQSSKAEISRLADKVSSVFVPSVIAIAVLAFV VWL
 IIA PKPDFWNNFGIALEVFVSVLVISCPALGLATPMSILVANQKASSLGLFFKDAKSLE
 KARLVNTIVFDKTGTLTNGKP VVKSVYSNIELLELLSLASSIEKSSEHVI AKGIVEYAKE
 HDAPLKEMSEVKVKTGFGISAKVDYQGAKEIIKVGNSEFFNPINTLEIKENGILVFVGRM
 INEKEDEL LGAFVLEDLPKKG VKEHVAQIKNLGINTFLLSGDNRENVQKCALELGIDGYI
 SNAKPQDKLNKIKELKEKGQIVMMVGDGLNDAPSLAMSDVAVVMAKGS DSVSQAADIVSF
 NNDIKSVYSAIKLSQATIKNIKENLFWAFCYNSVFIPLACGVLYKANIMLS PAIAGLAMS
 LSSVSVVLNSQRLRNFKIKDH
 >PROKKA_00819 phosphatidylglycerophosphate synthetase
 MPINPLYLFPNLFTASSIFLGMSIF YASSYQFVMACWL VVASLILDGLDGRVARLTNTT
 SKFGIEFD SLADVIAFGVAPSLITYFYVGYNFG RIGMAVSALFVIFGAIRLARFNISTNT
 SDPYSFIGIPIPA AAVLVVLCVLLDNKYHFLEGNTEKLFLGFI VLLGVL MVSNIRYPNFK
 KVKNLKL FILVLIFLSLVFVCPLEALSVFMGLYLIYGIIRWLF LMVKIIFNKNKST
 >PROKKA_00820 ATP-dependent zinc metalloprotease FtsH
 MKPTNEPKKPFQSP IILAVLGGILLIFFLRSFNSDGSFSDNFLASSTKNVSYHEIKQLI
 SNNEVENVSIGQTLIKASHKEGNRVIYIAKRVPDLTLVPLLDEKKINYSGFSESNFFTD
 MLGWLMPILVILGLWMFMANRMQKNMGGGIFGMGSAKKLINAEKPNVRFNDMAGNEEAKE
 EVVEIVDFLKYPERYANLGAKIPKGVL LVGPPGTGKTLLAKAVAGEAHVPFFSMGGSSFI
 EMFVGLGASVRDLFETAKKQAPSII FIDEIDAIGKSRAAGGVVSGNDEREQTLNQLLAE
 MDGFGSENAPVIVLAATNRPEILD PALMRPGRFDRQVLVDKPDFNGRVEILKVHIKGVKL
 ANDVNLQEVAKL TAGLAGADLANI INEAALLAGRNNQKEVRQ QHLKEAVERGIAGLEKKS
 RRISPKEKKIVAYHESGHAVISEMTKGSARVNKVSII PRGMAALGYTLNTPEENKYL MQK
 HELIAEIDVLLGGRAAEDVFLEEISTGASNDLERATDIIKGMVSY YGMSSVSGLMVLEKQ
 RNAFLGGGYGSSREFSEKTA EEMDLFIKNLLEERYEHVKQTLSDYREAIEIMVKELFDKE
 VITGERVREI ISEYETANNLESRLIPLEEQAS
 >PROKKA_00821 Ribosomal protein L11 methyltransferase
 MYEFFFFIFPKERELFESFLD TTHLAL EESSLEDLKA FDDKETIEFISQSSWHYFATHD

PLKKDLKEHLKEKPPYLKNFVILRSEENLNNSLILALEAFCLNLKQNLQNEFDFFYL SRN
LASKDWLEAYKQAILPVQCAKFYIHPSWHQKPSHVVTNDCIMIDPALAFGSGHHESTSMC
LELLSNLDLKRKNALDVGCGSGILSIALKKQGV SALVACD TDSLAVEETLKNFSLNQIPL
LAQDKVIYGSTQKIEGRFDVIVANLVADVIKSLYSEFVRLCNHTLILSGILETHLNSVLQ
IYYNGFEILEQRQRNEWVALKLLKKQPIN
>PROKKA_00822 hypothetical protein
MKLLVDDSSMTMRRIKNTLSRLGYEDVLEAEHGVEAEKLDANADTKVLITDWNMPERN
GLDLVKKVRSDSRFKEIPIIMITTEGGKAEVITALKAGVNNYIVKPFPTQVLKEKLEVVL
GTND
>PROKKA_00823 Helicobacter outer membrane protein
MYLRLILILGISFLNINLAENLSYMSSSYQVGTVMRPLNTNKLQGASILQGYEVNPKN
DWAYSRYFFIDYGNVLFNNDSTLQANMFTYGVGGDFMVAYAKNPINRWAFFFGQLAAN
TWILNNKVKDLVVNTWDSLKDFNFHNTYFRAIGKFGVQFRTIVLYHKVDVEIGMKIFLTP
ERRSLFERSFLFFISHSWHF
>PROKKA_00825 hypothetical protein
MLILGHPLIPSARFVFIKNTDAIHSSANNDIVCFEAHPKNLELAKYCCENG VHF SVIFLS
HKIETDAFFLFNAFKPLYCIFKDIKQAILAQQHATNYLLDSKILFSMDLNDTDSWEICAK
NQIDGVISKDSL LK
>PROKKA_00826 hypothetical protein
MLRILIPLLIIVVWLWRLFLRQKPLKDNHSYTQQTPELEDHMIVCSKCQTYVSSKDAIY
SGAVAYCSETCLKDKR
>PROKKA_00827 Ribosomal RNA small subunit methyltransferase G
MNPLLQDYARILLEWNQTHNLSGAKNLSELEPQITDALKPLEFVKDFKSCLDIGSGAGLP
AIPALAEKPKETKFILLEPRIKRAAFLNYLKS VLP LKNIEI IKRLEDYQNLQVDLITSR
AVASSSFLIEKSQCFLKDKGYFLFYKGEQLKDEIACEDTECFVHQKRVYFYKSKESLC
>PROKKA_00828 S-adenosylmethionine:tRNA ribosyltransferase-isomerase
MKEFDLESYDYDLPEKELIANYPILPKEKAKLLVYERRSQKITHTTFEHVL DFFPKNALVV
LNDTKVIKARLFGSKHAF LPSKTTEVFFHRRFFKNNTALTQIKGKIKVGDKIFFDANYYAE
VLELLHNGQRLIAFYDNQTPLNQENILKLLQYGHMPLPPYIKRADES LDAHEYQSVFAK
HIGAVAAPTASLHFSQHCKLEKLLKDFKHAF LTLHV GAGTFLGVETKDIREHQIHTEVLHI
PKKSQEILRESQEILCIGTTALRSVEYFKRLKNPNQESFECDFLHLANPILHVNYLLTN
FHLPKSSLLMLVSAMIGLDKTEIYKIAIEKKYRFYSYGDGMLIL
>PROKKA_00829 Sec-independent protein translocase protein TatC
MFEDLKPHLQELRKRLMVS VGTILVAFLGCFHFWK SIFEFVKNSYKGTLIQLSPIEGVMV
AVKISFSAIVISMPIIFWQLWLFIA PGLYKNEKKVILPFVFFGSGMFLIGA AF SYVVVF
PFIIEYLATFGSDVFAANISASSYVSFFTRLILGFGVAFELPVLAYFLAKVGLITDASLK
AYFKYAI VVIFIVAAIITPPDVVSQIFMALPLVGLYGLSILIAKMVNPAPKDNEEDSENN
AKEHAKSES
>PROKKA_00830 Sec-independent protein translocase protein TatB
MGFFEILVVLIVAIIFLGPEKFPQAVVDVVKFFRAVKKTLNDAKDTLDKEINIEEIKKET
LEYQKLFENKVESLKGVKIEELED AKITAENEIKSIQDLMQDYQRSLETNTIPNHLNEEV
SNEEALNKEASSDESPKEVQLATDNNTKEHDKEKEHV
>PROKKA_00831 Holliday junction ATP-dependent DNA helicase RuvB
MKERIVNLETLD FEISQEVSLRPSLWEDFIGQEIKSNLQISICAAKKRQESLDHMLFFG
PPGLGKTSISHIIAKEMETNIKITAAPMIEKSGDLAAILTNLQAKDILFIDEIHR LSPA I
EEVLYPAMEDFRLDIIIGSGPAAQTIKIDLPPFTLIGATTRAGMLS NPLRDRFGMSFRMQ
FYSPSELALIIKKAAAKLNQDIKEESTDEIAKRSRGTPRIALRLLKRV RDFALVKNSSLM
DLNITLHALNELGVNELGFDEADLAYLSLLANAQ GKPVGLNTIAASMREDEGTIEDVIEP
FLLANGYLERTAKGRIATPKTHALLKIPTLKSQTLF
>PROKKA_00832 3-methyl-2-oxobutanoate hydroxymethyltransferase
MSMQTAPIKKITLNLHQA KKNQEKIIAITAYDALFAQIFDPLVDVILVGD SLNMSFFNQ N
DTLSASVKMMLYHTKAVCAGAKTPFIITDMPFGSYKDEKTALKNAIRVYKETQANA IKLE
GGKEKAKLVKTLTDEGVI VVGHIGLMPQFVHLDGGYKIKGKNEEQQKKLLEDALSLEEAG
AGLLVLEGITTPIAQ TITQKIKIPTIGIGSGKDCDGQILVWS DMLGFFDSFKPKFVREYL
KGKELVQNAIKQYADDVKGNFPNELESYH
>PROKKA_00833 hypothetical protein
MGLKKRAILWSLMGFYTGLSALDYDTLDPKYKYIKYKAYGDKEVEELIRDLKRANAKS
GLILGINTGFFYNHEIMVRTNSSSITGNILNYLFAYGLRFGYQTFRPSFFARLVKPNIIIG

RRIYIQYYGGAPKKAGFGSVGFQSVMLNGDFLLDIPLPFVGKYLYMGGYMGLGLGVVAHG
 VNYTAEWGMSFNAGLALTVLEKNRIEFEFKILNNFPFLQSNSSKETWWGAIASIGYQYVF
 >PROKKA_00834 Helicobacter outer membrane protein
 MCSKKIRNLILCFGFILSLHAEENTAQENTTEENTPKDAPILLEEKRAQTLEFEENKESK
 KKIDEKSLLEEIHKKKRQLYMLKGELHEKNESILFQMAKNKSGFFIGVILGDIGVNAHP
 NARSYESFEPSSSIQASPLLYGLRSGYQKYFANGISALRFYGEYLGGAMKGFKSDSLASY
 QTASLNIDLLMDKPIDKEKRFALGIFGGVGVWNGMYQNLKEVKGYSQPNAFGLVLNLGV
 SMTLNLKHRFELALKMPPLKETSQTFLYYFKSTNIYYISYNYLL
 >PROKKA_00835 Helicobacter outer membrane protein
 MKLKYWLVLAFIIGLQATDYDNLEENQQQLDEKINNLRQLTEKGVSPKEMDKDKFEEE
 YLERTYPKISSKKRKLLKSFSIADDSKGVFLGGGYAYGELNLSYQGEMNDKYGANAPSA
 FKNNININAPVSMISVKFGYQKYFVPYFGTRFYGDLLLGGGALKENALKQSVGSFFYVLG
 AMNTDLLFDMPLDFKTKKHFLGVYAGFGIGLMLYQDKLNQNGRDLVVGGYSSPNFLWKSL
 IEVDYTFNVGVSLTLYRKHRLEIGTKLPISYLRMGVEEGAIYQNKENDERLLISANNQFK
 RSSFLLVNYAFIF
 >PROKKA_00836 hypothetical protein
 MGLGFKILALVVLILGGYLIFNAFITKPRALSFSLSKSEENALDDNNEVLFWDLKKPIKVK
 IVAPKGIKRYDLKVTTKDNLILYEKENLVLDKPKSLEVPLIKPEIMGLEDKCLLYEIQAS
 DWSYANFFNGNKASFQEVICIDTIKPSITILSRSPSIAYGGSAAVVVFEALDKNLSQAFVH
 VKKKDKFAFRLLEFKQRNVFIALVPWSYKNKDFKAFIIAKDKAYNSNTIPLLLKRKTRHV
 REKDIDLSALKDKIAQQEKFNHTEQTLLEFRFSNARLKDLEKIQKIALDQGDFYKDFSHP
 QALKPLNGPFKMVSNFLENRRFLKDNQVLFKFLHLGVLDLIPSKDLSLAFDPSVKRVFKGE
 LDFYGNLSLINCYGLGLCVFLAHLKDDKSVGSSGLKLGSLHLGMLLQGVFVRPNEWLNEQ
 WIKTNIITPIEQAKRLLMKG
 >PROKKA_00837 putative septum site-determining protein MinC
 MLKTNQKNVHAFFEIEKQEPEAVIGFLEKNHALLQYFLIIFKYDIEPEVKAILHKKHQLLFL
 ETNRTLNGRHIKTMSLKDETDHSPNNPKTEPKTTIYERHIRSGEEIYSANHLIFLGNIH
 NGAKIISEGCVSVYGVCEGAIVCFGERLILKEVKSAQIVFQNKILSLKEVERLLVNKNIK
 IITKNDDILDIKEVL
 >PROKKA_00838 UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase
 MKQTTINHSELVIGLGHKGVPVKLVLEPLGENQGIVFYRSDLGVKLPLKPENIVDTKMA
 TVLGKDNARISTIEHLLSAIHAYGIDNLKISVNNEEIPIMDGSALTYCMLLDEAGIKELD
 APKKVMEIKQAVEVREGDKFVKIEPDSQLSLNFTIDFNHPVIAKQAHHFIFSKAAYKEQV
 AKARTFGFLQEVNLYRSIGLAKGGSLLNCIVLDENSILNKEGLRCEKEFVCHKILDAMGD
 LMVLGMPVMGKYTSFSGSHKLSMLVKAILADTKNYEVLATDPAKEFALQKAFA
 >PROKKA_00839 hypothetical protein
 MELDLVLISLGEGLVLLGVYQNNFLCASYTSKSKTSEALVEVFSQLFKDFKNPNLPAIKGV
 YYAKGPGSFTSLKLTHVFLHTLALIHDFELYSTTGDFDNDNTPILAYANKYFVSKERESL
 NDFKDLKIAPKDFMLPPFLEKDKFTQLNTPFYILPPI
 >PROKKA_00840 Homoserine kinase
 MVVSVPATSANLGPFGDCLGLSLNLRNRFFIEPSNIHAVKLVGEGERIPKFLTNNIFTKV
 FYEILKKKHGDDGSFKFLLHNKVPITRGMGSSSAMIVGAVASAFALGFPFDRENILNTAL
 IYENHPDNITPAVFGGYNAAFVEKKKVISLTKIPSLKAVMVI PNRAISTKQSRHLLPK
 RYSVQESVFNLSHASLMTMAIVQGWDLRLCCSKDRMHQYKRMQTPVLFQKLALENN
 ALMSTLSGSGSSFFNMCEEDAPKLKQVLSKKFPKFRVAVLDFDNDGVVLEKD
 >PROKKA_00841 hypothetical protein
 MRKIEIKIRMCVACRMRPQKDLLRLKSFENQIVEFDGKGRSFYVCGNCLKNGEKKLLKA
 VSKIKNAPKDTKNIINWIKERSIA
 >PROKKA_00842 Translation initiation factor IF-2
 MSDMVDLKKFVTELGKTQKELKNVIEQAKDIGLELKTNSKMTPEQAGKLYKYIVDGIKEQ
 IQANQPTKNPEQDNKDDLNTAATPKPLVKKASKTPKKEETKAQPKPKKTKEKKKEAPVPI
 IKKKGIEIVNTFENQTPPIENTPKAVSHSQIEKAKQKLQEIQKSREALNKLTQSNTNTTN
 NANSASNVSNAKKEISEVKKQEQEIKRHENIKRRTGFRVIKRNDENENETENSVTESKPP
 TQSAAAFEDIKKEWQEKDKQETKKTKKPSKPKATPTAKNNKSHKIDFSDVRDFKDNDIY
 DDETDEILLFDLHEQDNLNKEEEEKEARQNINDRVRVQRKNPWNNEGGIKRQSKKKRAFR
 NDNSQKVIQSAIAIPEEVRVYEFAQKANLNLADVIKTLFNLGLMVTKNDFLDKDSIEILA
 EEFHLEISVQNTLEEFEEVLEGVKKERPPVVTIMGHVDHGKTSLLDKIRDKRVAHTEA
 GGITQHI GAYMVEKNNKWVSFIDTPGHEAFSQMRNRGAQVTDIAVIVIAADDGVKQQTIE

ALEHAKVANVPVIFAMNMKMDKPNVNPDKLKAECAELGYNPVDWGGEHEFIPVSAKTGDGI
 DNLLLETILIQADIMELKAIEEGSARAVVLEGSVEKGRGAVATVIVQSGTLSVGDSFFAET
 AFGKVRTMTDDQGKSIQSLKPSMVALITGLSEVPPAGSVLIGVENDSIARLQAQKRATYL
 RQKALSSTKVSFDELSEMVANKELKNIPVVIKADTQGSLEAIKNSLLELNNEEVAIQVI
 HSGVGGITENDLSLVSSSEHAVILGFNIRPTGNVKNKAKEYNVSIKTYTVIYALIEEMRS
 LLLGLMSPIIEEEHTGQAEVRETFNIPKVGTIAGCVSDGVIARGIKARLIRDGVVIHTG
 EILSLKRFRKDDVKEVSKGYECGIMLENYNEIKVGDVFETYKEIHKKRTL
 >PROKKA_00843 Ribosome-binding factor A
 MNAHKERLESNLLELLQEALASLNDGELNSLSVTKVECSKGKHHAYVFVLSSDHKILSKL
 KKAEGLRQFVLQASGWFKCPKLSFVLDNSLEKQLRLDAIFNEITKGKDND
 >PROKKA_00844 Ribosome maturation factor RimP
 MTKRIEEKIEGVIESLGYLLYDVSLVKENEQHVLRVSLKNPNGAVSLDICQQVSEVISPL
 LDVCDFIQDAYTLEVSSMGLERTLTKTPKHFKLSLGEKVEVKLINKESFQAVLKDANDLSA
 DFELEDHAIKNVEYKDLKKVKTLFEW
 >PROKKA_00845 Acetyl-coenzyme A synthetase
 MQLDEDLFAKKIFNPNRAFAKQARIKNMCEYKDLVHEANEDYEHFWGELAKQKLTWFKP
 FDKVLNSDNAPFFFKWFENGKINVSYNCIDRHLKDKKNKVAIIFEGEMGDYNVITYRKLHS
 EVNKTANLLKNEFNVKKGDRVIIYMPMIVESVYMLACARIGAIHSIVFAGFSPEALDRD
 INDAQAKLVITADGTFRKGKPYMLKPALDKALENNACPSVEKALIVIRNAKEIDYVRGRD
 FVYNEMVNYQSDKCEPEMMDSEDPLFLLYTSGSTGKPKGVQHSSAGYLLWAQMTMEWVFD
 IRDNDNFWCTADIGWITGHTYVVYGPLACGATTLILEGTMSYPDYGRWWRMIEEYRVDF
 YTSPTAIRMLHAKGENEPLKYNLESCLKVLGTVGEPINPTAWKWFYEKIGNSKCSIVDTWW
 QTETGGHIIISPLPGATPIRASCTTLPLPGIHAEVLNEDGKTKPKGEQGFLCITKPWPSMV
 RNIWGDEKRYIDSYFSQIKLNGEYVYLSGDGAIVDENGYITIIIGRTDDIVNVSGHRIGTA
 EVESAISKHEMVVECAVVGIPDAIKGEGLFAFVVLCDGAKCNLGESLELLKEMNHILSVE
 IGKIAKLDNMVYPGLPKTRSGKIMRRLKLSIAKKEPITQDLSTLEDVNVVKEIMSIAQM
 EE
 >PROKKA_00846 putative metallophosphoesterase
 MLISIAFLVLCLLNYSSFRMLKSFLTLLKISQHAYLWFFVLLSAGEVAFACRNTMPSH
 LFLVTSACSFVSFIVFIFSLSFYGFYSYIEKIDFLHSRRKSLKNFLKLGFYLALLGYFWR
 GFYEGLARPKIKETPIYLDKLDKELKIIILLTDMHVGSLQKDFVDYIVEEVNQKEVDMVL
 IGGDLVDESIEKVKSFLLPLNNLKSTHGTFFYVPGNHEYYHGIEPILSFLDTLNLTLILGNE
 CVHLGGINLCGVYDYFARKRQNFAPDIDKALKKRDSSKPTILLAHQPKQIRSLKEGHSVD
 LVLSGHTHAGQIFPFSLLVKLAQTYLHGLYKHSPTTQIYVSSGAGYWGVPFLRFLAPSEIA
 YLRLLPKNQA
 >PROKKA_00847 Transcriptional regulatory protein CusR
 MRVLLIEKNSVLGGEIEKGLNIKGFMADVTESELEDGEYLMDIRNYDLVMVSDKNALSFSVS
 RIKEKHSSIVVLVSSDNPTSEEEVHAFEQGADDYIAKPYRSIKALVARIEARLRFWGSNV
 IEIGDLTISPDEEKIYYKGREVEVKGKPFVLTHLARHRDQIVSKEQLLDAIWEPEMVT
 PNVIEWAINQIRQKMDKPLGISTVETVRRRGYRFCYPKPACEE
 >PROKKA_00848 3',5'-cyclic-nucleotide phosphodiesterase
 MQVYHLSHIDLDDGYACQLVSKQFFKNIQCYNANYGREVSARIYEILNIAAQSESEFLIL
 ISDLNLLNNEAEYLQDKIQEHLRQNKNIQIQLLDHHISGKEVAESFHWYFLDTNRCATKI
 VYEFLKKHYAILEPKNTTWLEPLVEMVNSVDIWDTQGYGFELGKVCMRMINQSSELNRFM
 FDDENRDYKLKLLLEVKNYLFLENAPVAYDNDLFKLKKIALGGDPDAETMDNISSNAQTH
 LLSLKKHDCSVYYQDKKGFLSYSMGGISVLANLFLTQNPDFDFYIDVNSKGNVSLRANGN
 CDVCELSQMCFNCGGHRNASGGKIDGFRESFNRYRDIKEQIEEIFNNA
 >PROKKA_00849 Flagellar biosynthesis protein FlhA
 MANERSKLAFKKTFPVFKRFLQSKDLALVVVFVIAILAIIVPLPPFVLDLFLTISIALSV
 LIILIGLYIDKPTDFSAFPTLLLIVTLYRLALNVATTRMILTQGYKGPSAVSDIITAFGE
 FSVSGNYVIGAIIFSILVLVNLVVTNGSTRVTEVRARFALDAMPKGQMAIDADLNSGLI
 DDKEAKKRRTALSQEADFYGAMDGASKFVKGDIAISIIITLINIIGGFLVGVFQRDMSLS
 FSASTFTILTIGDGLVGQIPALIIATATGIVATRTTQNEEEDFASKLITQLTNKSCTLVI
 VGAILLLFATIPGLPTFSLAFVGTFLFLFIAWLISREGKDGLLTKLENYLSQKFGLDLSEK
 PHSSKIKPHTPTTRAKTQEELKREEEQAIDEVLKIEFLELALGYQLISLADMKGQGDLL
 RIRGIRKKIASDYGFLMPQIRIRDNLQLPPTHYEIKLKGIVIGEGMVMPDKFLAMNTGFV
 NKEIEGIPTKEPAFGMDALWIETKNKEEAIQGYTIIDPSTVIATHTSELVKKYAEDFIT
 KDEVKSLLERLAKDYPTIVEESKKIPTGAIRSVLQALLHEKIPKDMLTILETITDIAPL

VQNDVNILTEQVRARLSRVITNAFKSEDGRLKFLTFSTDSEQFLNKLRENGTSKSLLLN
VGELOKLIEMVSEEAMKVLQKGIAPVILIVEPNLRKALSNOQMEQARIDVVVLSHAELDPN
SNFEALGTIHINF
>PROKKA_00850 30S ribosomal protein S15
MALNLEKKQEIIKAFATKQNDTGSCVQVALLNERIKLLTEHLKANPKDHSSRLGLLKL
AQRRNLLKYIKRTDHARYVVLIEKLGIKDR
>PROKKA_00851 O-Antigen ligase
MKERLKAFFSADSVFTLIFALFFLTSTFKKPLTQVLLIVLMVFLFLRCYFQASLKETFGIN
HLKTMPPFKWLTALFLGVFLSIFPNMFMNHDSQTFRYNLFALNMSLTACGALCLLFASRL
RIKLDQKILFYSMVANFINGLLSLVQKIYFNMMPRAQGFSTVKEYVVLVSVSILGCYIYV
LYSQNQKEKLFFTLSVFGFLVVLVSATRSATIAFVITFLILSCFILIYAKKSLKPLGYMV
VVSILSALYVGSNALEKRGAEQSRVQVQSFEEDLKRYAKKDADSSIGWRLEWKEALT
VLRRLRPFPGMAASEKQRLLEEILSLSKSYRAKDLILCYERYDNQIIHVLATRGIIGFLIW
LFFLLAVVKIFWSGIKQNSLISFFILTTLTFTYLI FGIGFDPDFDFFITGSFFVGMTMMAVF
VKKDKSAF
>PROKKA_00852 3-dehydroquinate dehydratase
MKILVIQGPNNMLGHRDPRLYGMVTLTQIHEIMQTFVKQGNLEVELEFFQTNFEGEIID
KIQESVGSDEYEGIIINPGAFSHTSIAIADAIMLAGKPVIEVHLTNIQAREEFRKNSYTGA
ACGGVIMGFGPLGYNMALMAMVNILAEMKAFQEAQKNNPNPNPINNQK
>PROKKA_00853 Aminopeptidase
MKGLERESHFTLNENAMFFECAYSCDNALFLQLDDRSFFITDSRYTQEAKESI QPKNGVL
AEVIESSDLAQSAIDLIIKSSVKKIFFDPNQVNLQTYKRLDSAIGNKVILEGVPSYHRQK
RIIKNEHEIQLLKKSQALNVEAFENFAEYVKKIFDEKESLSERYLQHKVKDFLTKEGVYD
LSFEPILALNANASKPHALPSAKDFLKAHESILLDMGIKYERYCSDRTRTAFFDPKDFVF
KREQSFKDKERQKIYDIVKEAQEKAISGIRAGMTGKEVDGLARGVISDYGYGQYFTHSTG
HGIGLDIHELPISSRSGTILEEGMVFSVEPGIYIPGFFGVRIEDLVVIKNSRAELL
>PROKKA_00854 2-amino-4-hydroxy-6-hydroxymethyl dihydropteridine
pyrophosphokinase
MREILTNRFFPSLFKKRLDFSNRVVLGLGSNLKNPLEILKNCFLYFKNHSGKIGKIFSSPI
YINPPFGYTKQPNFYNATIILKTSLSLRHFFALVFYIERFRGRQRKCDFKDAPRTLDDIDI
IAFNQVILRQNDLTLPHPKWGERDSVLVPLALQQILFKKGEW
>PROKKA_00855 flagellar biosynthesis regulator FlhF
MKFYTYSGETAEEALKIAQSHHGVDTLVFKTQEIRKKTLTSSGLYEIVVAVEEKSENKPS
KAPPYPRKFV
>PROKKA_00856 Flagellar biosynthesis protein FlhF
MYDEELNEEDVVMQLSSTVEEMRKLAVSSSQRNYSFSKNKTLLEKDAPLEDAPLGANKQ
DALLQALKDEANHHKEREKREVKQEEEEIKDINLQLSKIRDSLKLIQNMFWDEKNPNSVNI
PQEF AEIYKLAKQSGMKPSHLDEIMQLSLELMLPLRMRENSVTIKRYFREVLKRMILCCPE
DLNLRQKRILMLVGPTGVGKTTTLAKLAARYSRMLAKKYKVGIIITLDNYRIGALEQLSWY
ANKMKMSIEAVIDAKDFAKEIEALEYCDFILVDTTGHSQYDKEKIALGLKEFIDGGYNIDV
SLVLSVTTKYEDMKDIYDSFGVLGIDTLIFTKLDESRLGNLFSLVHESQKPISYLSVGQ
EVPMDLKVATNEYLVDCMLDGFSPNPKEQA
>PROKKA_00857 Flagellum site-determining protein YlxH
MNNQASRLDNLMNLKNPKSFFDNKGNTKFIAITSGKGGVGKSNISANLAYSLEYKKGYKVG
VFDADIGLANLDVIFGVKTHKNILHALKGEAKLQEIICEIEPGLCLIPGDSGEEILKYIS
GAEALDQFVDEEGVLSSLDYIVVDTGAGIGATTQAFNLASDCVVIVTTPDPSAITDAYAC
IKINSKNKDELFLIANMVAQPKEGRATYERLFKVAKNNIASLELHYLGAIENSSLLKRYV
RERKILRKIAPNDLFSQSIDQIAGLLVSKLETGALEIPKEGLKSFFKRLNLNLG
>PROKKA_00858 hypothetical protein
MKVQNFIFHSVVVGFFLGLVFSVLKFNPEPESILLWTVLSTLGGYLIALLFASIFIACDLD
DICLFDKKGTEESLLRFNHEFKNREKEVASILEYIRSDFDDGK
>PROKKA_00859 RNA polymerase sigma factor FlhA
MPKEIQNIETSEKNIEKVLNAYDEQQHHHQDALAIQYLPVAVRAMAFRLKERLPSSIDFND
LVSIGTEELIKLARRYESALNDSFWGYAKTRVNGAMLDYLRSLDVISRSNRKLIKSIDIE
ITKHLNEHGKEPSDAYLAEALGENIEKIKEAKTASDIYALVPIDEQFNAIEQDEITKKIE
AEELLEHVQKALNQMSEREQILIQLYYFEELNLSEIKEILGITESRISQIIKEVIKKVRK
SLGVDHG
>PROKKA_00860 Flagellar motor switch protein FliM

MADILSQEEIDALLEVVNDENVDIQNVQKKDIIPQRSVTLYDFKRPNRVSKEQLRSFRSIIH
DKMARNLSSQVSSIMRSIVEIQLHSVDQMTYGEFLMSLPSPTSFNVFSMKPMGGTGVLEI
NPSIAFPIDRLLLGGKGSAYDQNREFSDIELNLLDTILRQVMQILKEVWSPVVEMYPTID
AKESSANVVQIVAQNEISIMVVLEIIIGHSRGMMNICYPVISIESILSKMGSRDLMLSET
NSKKS RNKELQALLSGVSDMIVFLGAVELSLKEMLDLDVGDITRLNKIANDEVSVYVHK
KKALFSERGVSRV

>PROKKA_00861 hypothetical protein

MCIKKKRYLASVGFQGYRKTIIQIKEVIYSEKERTKEILEMLEEQRRGKVG DIMKIEEE

>PROKKA_00862 Flagellar motor switch protein Flin

MQDFIKIFIQEVVSTLEGLVGKAPSVGLEKEISSSDESFLKLSTPYARVKISAIEKEES
SIELLAPVDLVLTALSDMLLGEGASKEEMDNDLDAFKEMASNIFGAIATSLKSQELLPK
LNFTTTNAEIAKELPKKEDYAKAMVFSFKMEAIKESQIILLTTEAFEDQFKKTHQEEKEE
ATKSTIEEIKTHDASLENIEIRNISMLLDVKNVVRIGQKKMILKDVVSMDIGSVVELD
QLVNDPLEILVDDKVIAGEVVIVDGNFGIQTIDIGTKKERLEQLKN

>PROKKA_00863 hypothetical protein

MAVFGEGLGSLGHLFKKTQELEILHEYLKEIMQKGSKANQRVNLNDSNTEFQVPLGHGIFS
IEQSYCLEHAKSESEKGFESHKQYVDFQLIVKGVGAKVVDINQATIKTPYDEKRD LIVY
EPVSEASFLRLDAGMLAIFFENDAHALRFYGESFEKYREEPIFKAVVKAPKGLIKLKL

>PROKKA_00864 hypothetical protein

MKLVSLVIVFCCFLGAVELPGIYQTQEFLYMKSSFVEFFEHNGKFYAYGISDVDGSKAKK
DKLNP NPKLRNRSDKGVVFLSDLIKVGKRSYKGGKAYNFYDGKTYVVRVTQNSNGDLEFT
SSYDKWGYVGKTF TWKRLSDEEIKNLKLKRFNLDEV LKTLKDSPI

>PROKKA_00865 Ferric uptake regulation protein

MKRLETLESILERLRMSIKKNGLKNSKQREEVSVLYRSGTHLSPEEITHSIRQKDKN TS
ISSVYRILNFLEKENFICVLET SKSGRRYEIAAKEHHDHIIICLHCGKII EFADPEIENRQ
NEVVKKYQAKLISHDMKMFVWCKECQESEC

>PROKKA_00866 Replication-associated recombination protein A

MSLTSLNPKSLEDFLGQEHVKGDAPLFKALQSKHFPFAFFYGPPGVGKTS LAQIIARS
LERPILLFNATDFKLEDLRLKLKNYQNTLLKPVV FIDETHRLNKTQQEFLLPIMEKDHAL
ILGASTQDPNYSLSHAIRSRSFIFELTPLKKSDDLKCAKALTLLKKQIEPGA KTYLLNS
SAGDARALLNLLDLSAKVEDPITLETQLSLRPHSLNDGSYSDDTHYNLASALIKSLRGSD
ENASIYYLARLIAGGENPEFIARRLVIFASEDIGNANPNALNLASSCLFSVKQIGYPEAR
IILSQCVIYLACSPKSNTAYRAINQALDCVQKGS LYPKHLHPNAKDYLYPHDYNGYVK
QDYLEKPLDLVSSQGIGFEKTLLEWLDKIRN

>PROKKA_00867 Putative heat shock protein HspR

MCDYDEPLYLISVVAKILGVHPQTLRQYEKEGLIEPSRTDGKMRLYSQRDMDKIKTILRL
TRDMGVNLAGVDIILRLKEKLDLNLNKLQDALHKH SKNNKTPTKNLNTPTNFYELIL
FKK

>PROKKA_00868 Curved DNA-binding protein

MSKS LYQTLNVSENASQDEIKKSYRRLARQYHPDLNKTKEAEKFKEINAAYEILSDEEK
RRQYDQFGDNMFGGQNFSDFARSRGPS EDLDDLSSIFGKGGFSQRF SQNSQGFSGFNFS
NFAPENLDMTTTLNVSVLDTLLGNKKQVSINNETFSLKIPIGVEEGEKIRVRNKGKTGR T
GRGDLLEIHIIEDEIYRREKDDITQIFDLPLKTALFGGKIEIATWHKTLTLTIPNTKA
MQKFRIKEKGIKNRKTS HVGDLYLQARLILPKTETLSNELKALLEKEL

>PROKKA_00869 hypothetical protein

MQYKKNKKRYYHLALGILFCNGLSLKALEIAVKPFGYLGLLYNQGTQKNPHSYVGALARL
GVDFSYSNGWSFGIGAIGAWNIYNKQRLANLYISLGNFFGNPNNVKPYLSAGDVSDAYLQ
YTNQRFKIALGRFNTDFVDFDWIGGNIQGISVAFKQNSMRYFGIFMDSMLYNGHQINKEQ
GNRIATSLNALASYDPVSKRLYVGGEV FVLGTEYRHENLKVVPFILTDTRLPLSTQNVLV
QVGGKLEYDASLAKGFTSHTLVHGM YQYGNTDSATSVKNAGLFLIDQTFKYKIFNFGTG F
YIVPARNNKGYLWTFNDR TKFYGRGINAPGVPAIYFANSSISGYVFLGLKTKRVRLDAMV
AFGDYQEYSLMSSFRVW TYRSLSFDMGGGYVYAYNSKATRKS LGNSSFVFFGKFLF

>PROKKA_00870 DNA polymerase I

MEKYPLKKGARIQGEHCFEYEADDIISFYKKKDPNNYVIASMDKDILYSNRGSHFNLKTN
AFFNVSQKEAHFFAYYQCVVGDKGDNIGVKIGG FNYKXFLNEDAKEHELWXQIIQAFK
IKEDLSDSEAKEKALLNMRLVNMHQMTRHGVIKLWEPEFKKTFPPKKTQKPDFKRIS

>PROKKA_00871 DNA-binding transcriptional regulator BasR

MKILIIEDDLALARSISHNLHDLGHFCEIIS SISEENKEPYDVILVSSKVCTKGRCEHFV

RYNSKQIIIMMASHVNEDGVNKPIQAGARDYILKPFKMDLLRKIQYHKAYQEMTARLGF
YENYLDIFIHAELPLPKDFSYPFPFIHTPSQELANAYLLQYAKERQMDFSFFSLKDTTWK
ELYKNKDKLERPPFYIMHLEELKKDEQLKLELARSPIVLSYTHKEPLEFPKIVSIECGN
KPLSLFSNHTTFLSIQEYEKEAIRHFSSTCTDTELANKLGISRKSLWEKRRKYNLPRK

>PROKKA_00872 Bifunctional enzyme IspD/IspF

MSLIRVNGEAFKLSLESLEEDLFETKETLEILVKQTSVLLAAGESRRFSQTIKKQWLRS
NHTPLWLSVYESFKKALDFKEIILVVSGLDYIYIQRHYPEIKLVKGGASRQESVRNALKI
IDSAYTLTSDVARGLANMEALKSLFLTLQQTSHYCIAPYLPDYDTAIYYNEALDREAIKL
IQTPQLSHTKALQSALNQGDFKDESSAILQAFPNVSVYIEGSKNLHKLTSDDLKHFALF
FNPADFTFIGMGFDTHAFVKNKPMVLGGVLDCEFGKLAHSDGDALLHAVIDAILGAIKG
GDIGEWFPDNDPKYKNASSKELLKIVLDFSQSIGFELFEMGATIFSEIPKITPYKPAILE
NLSQLLGLEKSQISLKATTMEKMGFIGKQEGLLVQAHVSMRYKQKL

>PROKKA_00873 Periplasmic serine endoprotease DegP precursor

MKKTTLFISLALALSLNAGNIQIQNMPKVKERSVPSKDDTIYSYHDSIKDSIKAVVNIST
EKKIKNNFIGGGVFNDPFFQQFFGDLGGMIPKERMERALGSGVIISKDGYIVTNNHVIDG
ADKIKVTIPGSNKEYSATLVGTDSESDLAVIRITKDNLPITKFSDSNDISVGDVLFVFAIGN
PFGVGESVTQGIVSALNKSIGIGINSYENFIQTDASINPGNSGGALIDSRGGLVGINTAII
SKTGGNHGIGFAIPSNMVKDIVTQLIKTGKIERGYLGVLQDLSGDLQNSYDNKEGAVVI
SVEKDSPAKKAGILVWDLITEVNGKKVKNTNELRNLIGSMLPNQRVTLKVIRDKKERTFT
LTLAERKNPNKKETISAQNGAQGQLNGLQVEDLTQKTKRSMRLSDDVQGVLSQVNENSP
AEQAGFRQGNIIITKIEEIEVKSVADEFNHALEKYKGKPKRFLVLDLNQGYRIILVK

>PROKKA_00874 Amino-acid permease RocE

MEHQNITQKQGELKRDMMKRHLLMIAFGGAIGTGLFVGTGGNIANAGPLGTLLIAYGFGGL
VVCIMLSLSELASVYPTTGSFGDYAAKFIGPGTGYMVFWMYWLGWVITVALEYIAIGML
MQRWFVNIPIHYWILCIALVFLNFFSVKIFAEGEFFFSLIKVLAVIAFIGIGVIGIYY
QIYSHGFSSIFDNEHFHGDKGFFPNGSAAVFSAMLAVIFAFTGTEVIGVAVGETKNASEVM
PKAIKATLWRIVFFFLGSVFVISIFLPMNDSSITQSPFVSVLERINLPFIGMGIPYVADI
MNAVIITAMFSTANSGLYGASRMIIYGLSKQKMFFKVFSQLNKQGTPTYAMLFSLSFSLIG
LLVQIYAKENIVEALINVISFTVIVVWSVISQYSFRKQYLKAGHSLEDLPYKAPFLPF
LQLIGITGCAIGVIGSAMDKDQRIGMILTIVFAVICYIGYYFTQKANENNNKKDLI

>PROKKA_00875 CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase

MKVLKLLPNFLTILRIVLSLFLFLNTRTYFSFLTTPFHINMISSLVFLFAALTDLLDG
YIARSYKAKSRFGEIFDPLADKILILSAFLGLVYLDRVNAWIPFVILGREFFISGLRVLA
ANGKKDIPVNALGKYKTVSQVVAIGALLANLTYSYVLVAIAVFLTLYSIDYTIKYYKS

>PROKKA_00876 hypothetical protein

MQNIHIIHQNKELQFIKKCLLGYFFAPLCGAVLLVLFLVSSGAKSFQISNLFNQLAYIV
LLSLFLCTLGFIAGAIGFYRLSKITRHLSTFFENFAFSFLAVILCALLSYLIPNASNALSL
IGNGVSIFYLHKLYRELSLYTQERFFLSGFRLLLSFMLALLGVLAQALVLIIFLTIAVIL
MCVALSFLGRAFLNFSQVFLKA

>PROKKA_00877 Enoyl-[acyl-carrier-protein] reductase [NADPH] FabL

MNHSNHMKNTLVISGATRGIGKAILYRFAQSGVNIAFTYNNRNVEEANKIIEDEVEQKYSI
KAKAYPLNVLEPEQYTELFKQIDADFDRVDFFISNAIIYGRSVVGGFAPFMRLPKKGLNN
IYTATVLAFFVGAQEAARMQKIGGGAIVSLSTGNLVYMPNYAGHGNSKNAVETMVKYA
AVDLGEFNIRVNAVSGPIDTDALKAFFDYVEIKEKVEEQSPLKRMGNPNDLAAGAYFLC
DETQSGWLTGQTIVVDGGTTFK

>PROKKA_00878 4-hydroxy-tetrahydrodipicolinate synthase

MQFHSSSALITPFKKDLSVDEAAYESLIKRQIFQGMDACVPVGTGTSATLTHKEHMRCI
EIAIETCKNTKTPSNSRMKVLAVGSGNATSESLSLAKFAXKIGADAILCVSPYYNRPTQQ
GLFEHYKTIAQSVEIPVMYLDVPSRTGVSEIVPTALKLFREVPNIKAIKEASGSLKRVTE
LHYEKFDFKIFSGEDSLNHSIMFSGCGVISVTGNLMPNLISQMVNCALKLEYQQALETQ
NKLFLYLHQALFVETNPIPIKAMAMHLAALIENPSYRLPLVAPSKETIKLLEKTLOQYEVIA

>PROKKA_00879 Peptidase M16 inactive domain protein

MKHFSVKRLLGLSSVLLVTLGASMHAQSYLPKHESVTLKNGLQVVSVPLENKTGVIEVDV
LYKVGSRNEVMGKSGIAHMLEHLNFKSTKNLKAGEFDKIVKRFGGVSNASTSFDITRYFI
KTSQANLDKSLELFAETMGSLNLKEDEFLEPERQVVAEERRWRDNTSPIGMLYFRFFNTAY
VYHPYHWTPIGFMDDIQNWTLDIKKFHSLYYQPKNAIILVVGDVNSQKVFELSCKHFES
LKNLDGKAIPTPYMKEPKQDGARTAVVHKDGVHLEWVALGYKVPAPFKHKDQVALDALSL
LGEGKSSWLQSELVDKKRLASQAFSHNMQLQDESVFLFIAGGNPNVKAELQKEIVALLE

KLKKGEITQAELDKLKINQKADFISNLESSSDVAGLFADYLVQNDIQGLTDYQRQFLDLK
VSDLVRVANEYFKDTQSTTVFLKP

>PROKKA_00880 Dihydroorotate dehydrogenase (quinone)

MLYSLKLYLFSLDAEDVHEKVCKILRTLSSSPFLCNLIDSQWGYKNPKLENEILGLNFP
NPLGLAAGFDKNASMLRALIAFGFGYLEAGTLTNTAQSGNERPRLFRHIEEESLQNAMGF
NNYGAVLGVRSFKRFAPYKTPIGINLGKNKHIEQAHALEDYKAVLNKCLNIGDYTFNLS
SPNTPNLRDLQNKAFVNELFCMAKEMTHKPLFLKIAPDLETDDMLGIVNSAIGAGAHGII
ATNTTIDKSLVFAPKEMGGLSGKCLTKKSREIFKELAKAFFNKSVLVSVGGISDAKEAYE
RIKMGASLLQIYSAFIYNGPNLCQNILKDLVKLLQKDGFLSVKEAIGADLR

>PROKKA_00881 Polyphosphate kinase

MNRFFNRELSWLAFNTRVLNEAKDESLPLLERLKFLAIYDTNLDEFYMRVAGLKQLYEH
KIASKGIDGASPEEQLEKIKHYLAHEIEERELEFQKIQALLFKKGKLCITPYNELNLEQKA
KAKTYFKEQLYALVLPFKLDSSHTFPPLANLTFALFARIKDKETQTISYALIKLPSFIFR
FVELEKGLFVLAEEIVEAHLEELFLEHEILDMAFRVTCADIAITEDEAHDYADLMSKS
LRKRNQGEIVRLQTQKGSQELLKTLASLSFSQTHSYKKHKLTMHVYKSAIMNLNLGDLW
ELVNHSDFKTLKSPNFTPPIHPHFNENDLFKSIEKQDLLLLFHPYESFEPVIDLIEQAASD
PATLSIKMTLYRVGKHSPIVKALIEAASKIQVSVLVELKARFDEESNLHWAKALERAGAL
VVGVFVKLVHAKMLVITKKTNDQLRHFTHLSTGNYNPLSAKVYTDVSFFSAKNEIANDI
IKLFHSLLTSSATNSALETFLMAPKQIKPKIIELIQNEMNHKQEGYIILKANALVDSEII
EWLYQASQKGVKIDLIIRGICCLKPQVKGLSENIRVYSIVGKYLEHARIYYFKHENIYFS
SADLMPRNLERRVELLVPATNPKIANKLLHILEVQLKDTLKRVELNSKGCYSKVSNPNDP
LNSQDYFEKQALKTF

>PROKKA_00883 Type I restriction modification DNA specificity domain protein

MTKLDDLKNSVIAINQNNYTKEYNYKKVCYLDTDNITNNKINAFKIDLTKEKLPSRAK
RKCSINSIIYSSVRPNQRHFGIIEIPKNFLVSTAFIVIDVIDLEKLDPNLYYYYITQDE
IIHYLQRIAECGTSSYPSITPLDFLNIKIKLYPLETQQKIARTLSVLDQKIENNHKINEL
LHTLAYKIIEYYYFKYKPKNAKLEQIIIEENPKSNIMVKNAQKTQDKYPFFTSGDNILSYPK
AIIIDGRNCFNLNTGGNAGIKFYVGKASYSTDTCICANEFSDYLYLLLSSIKNHINQSFFQ
GTSCLKHLQKNLLKKYPIYMPSAHEIKKFNQIMMPLLTLSINTRTSKKLEQIRDFLLPL
LKQQVKPQ

>PROKKA_00884 putative type I restriction enzymeP M protein

MPNNALLQIKQDTLSLIDDLKVICTSFGLGNDGNEYKIITQCFLYKFLCDKFEFFFTETKF
PNKTIQDYKDFNEEEKEIFFTDLYDNKLPKLAYDELLSYLFEKHFNDNDLHLKLDAIFNR
ISSNNAELFNTKSTDKTTIALFESVSQYINEESKRANFTKALLDKLKNFNFKQAFNLQON
QQGYDFFAPIFEYLLKDYNNAGGGKYAEYYTPLSIASIIAKLLVSEPTQSVKIYDPSAGT
GTLLMALAHQIGTDSCITLYAQDISQKSLRMLKLNLIINLNLTHSLRYAIEGNTLINPYHSK
ECHGKMDFIVSNPPFKLDFSNEHAEISQNKNDFFLGVPNIPKNDKSKMPIYTLFFQHCLN
MLSNKGKGAIIIVPTGFISAKSGVENKIIIRHLVNERLVYGVVCMPSQVFANTGTNVSIIEFF
QKTPSAKEVILIDASKLGEYTENKNKKTRLRTSDIDLILETFQNKTKADFCALVSFDE
ITEKNYSLNPGQYFTIEDTSEKISQEEFENLMQQYSSSELTSLFDESQNLQQEILETLGNL
NHD

>PROKKA_00885 Type I restriction enzyme EcoR124II R protein

MFASLNVLKEQLKHYGTNPDKPLKGIWHWTQSGSKTALTYHLTKLIRDFFSPLGKKTIFY
FIVDRLDLLEQAKNEFLKRGKCAHEAENKEDLSQKLKSSSVFENPQGNDEIIVNIIQKFK
APNEEKAPNEDPSNSAPKEIVSKTELQESIQRNLQRFVFIIDEAHRSYDPKGCIFYANLI
ECDKTAIKIALTGTPLEDNAQDKATKNTFGNYLHTYSYTESIKDKHTLKLQLESIETSY
KEKLQEIYRLLQESITIEDTEVKKETIFNDEKYINAMLFYIIRDLLNFRQLHDNECLKAM
VVCFSKQAKLANCLFNEVQEKVLQENPNLRILNKLKSSLILHDEQEVKEKIHSFKHEDT
DIVFVFNMLLTGFDLPSLRLYIHRELKDHNLQALARVNRSYNNMSFGYLIDFVGIQEN
FDKTTDDYLKELNRFNQSGAHSDSNIKIDFADRETLEKDIKNAYDDLFDYPIDDIEGMTS
AIVSMSAMNELLKISHAINTLKINYNLIRASSDEKILSLKEKMDIEKINKISSMLHQKAK
YLHALKNINEPKNPNDLMILEDLIALLDKIEFKESKELHFKEKEEITAKYKQAKEILEK
IPDKQDKEIQKFYKDFSKLLQTPATSQNFEEISHSYDAIISQLKQHKQETHLLNKYDND
RAYVITHKRLHNRLEENISKGIFTLLSALKKALDERIFKRQEILNEEDTLKTAIKIELR
DAFKENPSLKDLOKEEFIAQTLFNELTQNHQGNLHA

>PROKKA_00886 hypothetical protein

MPYNEITRVQIPALMHLAKLGYDFIPTNSKENKPNLDTDNILIDRFTQAFERLNPSTKNA
QEMLAEMKKRLNYDDLKGSFYEYLLKSENQIIDFDNPNNNLYEMMAELPYKSFRPDITLF

INGLPLVNIEVKQPYAEKGIKEEKDRHIKRYENPENKVFYNL

>PROKKA_00887 hypothetical protein

MPFLKALĒSFDAFFLEKEISKRFRDNLIFFKSYNPNLNFALNTPFKNYQLLFEKNHFNLL
HTPTNALSYPKNQMIETAFNMA SNPLNNPRYSLDNNHLSLHYLKSQNNPKLPLTLKATHA
ISNFLNNHQTPCSLEKFLPPTMIYGVLDGLFLAILQAQNYRFHSLYLFEENLDFKISCY
FARYEDLITKGAKLFIQGFNPNELKMDFLKRPVTHSFLKLEIMPYKSAFN SRMRENIQS
YYKQALRGWGSFEDELLGLKNTLKNLPLYQTLTKPKKINAPICVVGNGPSLDLLLDLFLK
ENEDNFIIIFSCGTALKPLKTHGIKVDFQIEVERIDYLKEVLEKAPLEDTP LIGANMLNPN
AFNLAKEALMFMRGGSACAYISPLSVEYAAPFVGNAGVALAGLMSDEIYLCALDCAYIKG
FKKHAQNSYYENEKEIDTSSLISVEGNFKGYETFSDSLFLLSKERIEEALHYYQPKKVYN
LSYGAKIKHAVSLNHSQVKLKQINKQEA IARIKSMFSPKNNHAKDLNNLQKNLMSFKEDF
FTHLNTPCCKTQEIFEWVDSL SGCQTASAKTPTIGILFEGSIAHILQSVLIVSLHLNEN
ELTRFIKFSQNALKQFLKEACLL LQMQLKQP

>PROKKA_00888 Tellurite resistance protein TerB

MEIILLIIVA AVVLFYFYNTLKEYLKNPLNPKTKTEEYDLKNDPYLLAQSSPLDKFKQTQI
GAYMRL LKFLDIQKNALDNALRTLFIHELEQSLNSEQQNLAKELLNEPVDKKENFESLCQ
EIADHTHGEYTKRLKLVEFLMLLAYADGILDSKEKELFLDVGAFLQIDNQDFNELYDNFE
RFNAIEIPMSLEEAKNLFEIQANTTKQDLEEKALNLSAPYYHKMNDNKRYSEQDFISLKK
IALASQLLENGLKDS

>PROKKA_00889 hypothetical protein

MNAFLKLĀLASLMGGLWYAFNGEGSEIVAIGIFVLILFVFFIRPVSFQDPEKREEYIERL
KKNHERKMILQDKQKEEQMRLYQAKKERESRQKQDLKEQMKKYS

>PROKKA_00890 hypothetical protein

MSVLKLHVKVFRFETNKDYNPAYESYFLEYQEDQYLLDLLKQLKGVSYNENIALKINQIA
VFEDAKVSDLVAFFSKEWVLDPLSKRYALKDLMIDEKAVLKNYEDFFKQVPYITKGEKEE
LEKFIQINFINPQTNPKYLG DGFFLYVKWLMKRYPTERDRLLEMISQPESGVMNFLSVAH
YLYKNDDNIDHEIYELQEILTNSKIKPWKDFSKNLLSLFQYNSNPPKTPNPPKTCALFNA
YAKHLDAQSLLKS AKLYLEKMGGKIIDL PFCYDGGYYGKIISTHDFLTASAYNLALAKAN
GVSLIFCEEDAYLNILHAKEVLDNNPEIINSVNEK LKKYQLVYEKDIEIVYLNEWVNEFL
AWELKSPFDAFLGA EFSRIKRS DHFFNKIHLKAPHFLESFQNYAPLLEVNEVSGLLQCAH
LRYLGIDLGA DFLITHSLGLFHAFENLSLKASKIYKRDNDNTPTLFLS QIALMAMGEKDK
KALGLDTHYHKVTFI

>PROKKA_00891 hypothetical protein

MRRELRLĒNNKHCIEYLQFLSKNHLSFNLLCERDAIDFSPKLPKEIHEKFGSLVLFILAG
YTLES LMIDAQSVQFEAGFGPNNIGSVVQVKLP GIIQILIKEKNENAVLFNRCD SLEVVFH
KEEPPFQEPKKDERESKEWLD SKEALFSNSKNRAILENLHKS

>PROKKA_00892 Oligoendopeptidase F, plasmid

MKEQEWDĒSALFENKESAEEFLKTLQTEAQEFESTYQNNLKNLDATGFANALKHYENLSE
KISRVMAYAQLLFAKNTKEAKFY SQCEMACANIQQHLLFFEIEFKNLDAKKQLAFIKKCK
NHAFYLN NLIERKKHTLNLDEEKIALALSPVGVGAFSRLFDEHFSSLKIPFEEQNLSEEE
ILALLHNPKRKIRKKSQKAFSKALEKS RPLLT YILNMVRKDLLIETRLRKYDKKESFRHI
DNQISQESVDSMIEIVNANFSLVHRYYHQKAQILGHKLKDYDRYAPLS DENTTMTYSQAL
EEVLNTLKA FSPEFYKIASKAIKEGWVD SHPKDFKQGGAFSHGGVPSAXPYVLLNYTGNR
RDAFTIAHEFGHMIHQELSKKQGV LNMDTPLTTAETASVFSEMLFFEHLKKGLKQDELLF
MLAGKLEDIFSTLFRQVVM TNFERRIHEIDEELDTKDFDRIWFEEENQRMFEKSVKLTKNY
HLWWSYIPHFIHSPFYCYAYSYGQLLTALYGLYKKS DAKFEVKTYTEFLSLGGS KSPKE
LVSMFGFDIDSKEFWEIGMQEVRHLL EEFERLLACEEN

>PROKKA_00893 Glutathione-regulated potassium-efflux system protein KefC

MENSTLYĒVIAGLWLAVFGGIFLKKLDMPVIIGYICTGTVLA AFFKINDFDLLSDIGEFG
IVFLMFMI GIEFNFDKLSIKQEV LVFGLLQVILCALIAFLVGYFILGLSPIFSLVLGMG
LSLSSTAIVLKFFEDSKQLSTPMGKS AVGILIFQDIAAIPMLLILTILGSKDSNVNLLIL
KTFISAGIILV LLLLP GKKGANLILEQAKDTHLPEIFIGTILVIVCSAAGLSHFFGFSMS
LGAFIAGMAISKSHYKINVQEEFAQLKNLFLALFFITIGMQINISFFIDKIFAVIFSLIL
VMGFKTAIIYALLRFRD SKTAIKTALS LAQIGEF SFVIFLNSGSHQLFNLQEKKGILGF
LHQKNILSIAQNDIHQFLILMVVFSMLATPFILKYLE SIAQFILHQKSQENEPAKK

>PROKKA_00894 Helicobacter outer membrane protein

MIKRIACĒLSLSTSLALAGEVNGFFMGAGYQQGRYGPYNSNYSDWRHGNDLYGLNFKLGF
VGFANKWFGARVYGFLDWFN TSGTEHTKTNLLTYGGGGDLIVNLIPLDKFALGLIGGVQL

AGNTWMFPYDVNQTRFQFLWNLGGRMRVGDRSAFEAGVKFPMVNOGSKDVGLIRYYSWYV
DYVFTF

>PROKKA_00895 Molybdate-binding periplasmic protein precursor
MDKWDTPTKPPQTFKIQIFKIQIGIITNSNHLIKGSSMKNPFAFAFLIVFFSSTLLAQD
LKIAAANLTRLKTLVKEFQKEHPKDAISISFNSSGKLYAQITQNAFPDLFISADITRP
KKLYDEKITPFKEEVYAKGVLVLWSENLMDSLEILKDPKIKRIAMANPKLAPYGKASME
VLENLKLASSLSKIIYGASISQAHQFVATKNAQIGFGALSLMDKKDKLSYFIIDKALY
NPIEQALIITKNGANNPLAKVFKDFLFSFKARAIFKEYGYIVD

>PROKKA_00896 Sulfate transport system permease protein CysW
MDHEFLITMRLSFSLSALITTLILLPIGIFLGYFLSLKRNLTSLTETLVYMPVLPPSVL
GFYLLLIFSPSSFLGAFLQDALNVKLVFSFQGLILGSVIFSLPFMVSPIKSALISLPTSL
KEASYSLGKGEYYTLFFVLLPNIKPSVLMAIITTFHTTIGEFVVMMLGGDILGETRVAS
IAIFNEAEALNYSKAHQYALTTLISFSLLFVTLFLNKKQSSFL

>PROKKA_00897 Spermidine/putrescine import ATP-binding protein PotA
MIKARFKRRLGSRGAFDLNIDLEIKEAEVVALLGESGAGKSTILRILAGLEAVSSGYIE
VNRSVWLDLTQKKIFLKPQQRKIGFVFQDYALFPHLNVYQNI AFAPKDKNKIHEVRLMR
LENLSQQKISQLSGGQAQRVALARALIAAKNLLLLDEPLNALDNALKNEVQQGLLDFIKR
ENLSVLLVSHNPNEITKLAQTFLLNNGVIDPNQENRFLFSNRLLIKPLFEDENYCHYEVI
PQTISLPKDCLNPTFKLDFIQNKKF

>PROKKA_00898 Glutamate--tRNA ligase
MSLIVTRFAPSPTGYLHIGGLRTAIFNYLFARANQGKFFLRIEDTDLSRNSIEAANAIIIE
AFKWVGLEYDGEILYQSKRFEIYKEYIQKLLDEGKAYCYMSKDELDALEEQKARKETP
RYDNRYRDFKGTTPPKGIEPVVRIKVPQNEIIVFNDGVKGEVKVNTNELDDFI IARSDGVP
TYNFVVTIDDALMGITDVIRGDDHLSNTPKQIVLYKALNFKIPNFFHVPMILNEEGQKLS
KRHGATNVMDYQERGYLKEALVNFLARLGWSYQDKEIFSMQELLECFDPKDLNSSPSCFS
WHKLNWLNAYHLKNQSTQELLKLLKPFSSDLSHLNLTLQDLRLDLTKERSQTLKELALK
IDEVLIAPVEYEEKVFCKLNQALVMPLLEKFKLELNTTNFNDENALENAMHQIIEEEKIK
AGSFMQPLRLALLGKGGGIGLKEALFILGKAESIKRIEEFLKN

>PROKKA_00899 hypothetical protein
MQFPKTLFSLSLFLSYCIAEENGAYASVGFSEYSISHAVEHNNPFXNQERIQIDF

>PROKKA_00900 Helicobacter outer membrane protein
MPLSTITPFXIKNASKSISNAQNKIYKLNQVKNEITNMQNTFNYINNALKNNNSKLTPTM
QAEQYYLQSTLQNIKIVTLSSGGVASNPQLAQALEKMQEPIITNPLEFEENLKNLEVQFAQ
SQNRMLSSLSQIAAISNSLNALDPSSYSKNISSMYGVLSVGYKHFFTKKNQGFYYL
FYDYGTYNFGFVGNFDFGLGKMNNHLYGLGIDYLYNFIDNAKKHSSVGFYVGFALAGSSW
VGSGLGMWVSQTDFFINXLXXYQAKMHTSFFQIPLNFGVRVNVNRHNGFEMGLKIPLAVN
SFYETHGKGLNTSLFFKRLVFNVSIVYSF

>PROKKA_00901 Modification methylase VspI
MPSNALLIEEIALRVNVSHSSVHNWIKTNLLEKLEIDHKIYVKTSSFLDFCRNHLGKNKL
NKYANKSLKGAHNHQELILKYLQILENSSDLEKLGSYEEELSNTTRNLEGIYYTPNKIV
EQLFTLPKDFDASQAFCDPAVGSGNFIMHALKLGFKVENIYGYDTDAFAVALTKKRIKE
RYHLDCPNIMQKDFLNLKHTPQFDCIFTNPPWGKKYNQNKENFKQRFNLSQSLDSASLF
FIASLNYLKENAHLGLLPECLNIDAFSKMREVALKFQIRSLIDFDKPFKNLMTKAVGL
VLKKTTPKNQKISCFYQNKLFKRSPSSFFNPNKKIFNIHCSNKENKILDHLFSIPHTTLK
NNAHFALGIVTGNNKEKLHPKQEKNTIPIFRGSDILKDGLKAPSQFINADLKDCQQVAPL
SLYQAREKIVYKFISSKLFFYDNEQRLFLNSANMFVLKENFPINANALKELLNSDLMQF
IFESLFKTHKILRKDLECLPLFAQFINNSFDEKFYLNKLGIEKKDPKHFTIRKNHAHRLS
FGFRG

>PROKKA_00902 hypothetical protein
MHIACLLALGDNLITISLLKEIASKQQQPLKILGTRTLTKIAKLECEKHFEIIPLFENV
PAFYDLKKQGVFLAMKDFLWLLKAIKKHKIKHLVLEKQDFRSASLAKFIPITTPNKEIKN
VYQNRQELFSQIYGYVFDNPLYPMSVKNPKKILINPFTRENNRNISLEHLKIVLKLKPF
CVTLDDFEERYAFLKDEVAHYRAKTSLEEVKNLILESPLYIGGDSFLIHLAYLKKNYFI
FFYRDNDDEFMPPKNENFLKAHKSHSIEQDLAKKFRHLGLL

>PROKKA_00903 GTP-binding protein TypA/BipA
MKNIRNIAVIAHVDHGKTTLVGLLSQSGTFSEREKVDERVMSNDLERERGITILSKNT
AIYYKDTKINIIDTPGHADFGGEVERVLKMGVDGVLVVDAQEGVMPQTKFVVKALSFGI
CPIVVVNKIDKPAAEPDRVVEVDFDLFVAMGASDKQLDFPVVYAAARDGYAMKSLDDEKK

NLEPLFETILEHVPSPSGSVDEPLQMQIFTLDDYDNYVGKIGIARVFNNGSVKKNESVLLMK
 SDGSKENGRIITKLIGFLGLARTEIENAYAGDIVAIAGFNAMDVGDSSVDPANPMPLDPMH
 LEEPTMSVYFAVNDSPLAGLEGKHVTANKLKDRLLKEMQTNIAMKCEEMGEGKFKVSGRG
 ELQITILAENLRREGFEFSISRPEV I I KEENGVKCEPFEHLVIDTPQDFSGAIIERLGKR
 KAEMKAMNPMSDGYTRLEFEIPARGLIGYRSEFLTDTKGEVGMNHSFLEFRPFSGSVESR
 KNGALISMENGEATAFSLFNIQERGTLFINPQTKVYVGMVIGEHSRDNDLDVNPISKSHL
 TNMRASGSDDAIKLTTPRTMVLERALIEWIEEDEILEVTPLNLRIRKKILDPNMRKRAKK
 >PROKKA_00904 D12 class N6 adenine-specific DNA methyltransferase
 MHANLFNQASKKDVFLHNLSNNGRYKRYIKAPLRYGGGKSLAVGLIVECIPNGVRRMI
 SPFIGGGSVEIACAAELGLEVLGFDIFDILVNIFYQVLLKDKQALYNHLFSLEPNQETYST
 IKQELKAHYKKECVLGPLILARDYYFNFNLS
 >PROKKA_00905 D12 class N6 adenine-specific DNA methyltransferase
 MSKIYTDKERYLNALLKIKGFNAPSLKVECSSFEVLLAYLNDFFYLDTPYVLENSKMFK
 GIYPMRNFPIHNGFKHEVLAHMLKRHKGPFILSYNDCELVRNAYKDFKILEPSWQYTMG
 QGETRMGKNRLERGDNNHVKQSHELLIIKE
 >PROKKA_00906 hypothetical protein
 MHISEVKTAFKIADVEYVKDSTKLNFNYLKDLDKDENNQSLSQNILTQNVARVYLIVVDGE
 IKKIGGSQADGGIKNTLNIYKDGGVKGRPSVRSFGVWYFLYHTILTGAKIEFYMIYQPNF
 ETQIKGLFGFCAIKDASISYKLLQACLTDIRNNNNNDALPEWNEQEQQKD
 >PROKKA_00907 Modification methylase BspRI
 MYKVADIFCGAGGLSYGFSKHPYFELIWANDIDKDAILSYQANHKEAQTILCDIVQLHCH
 NLPCVSINILLGGPPCQSYSTLGKRKMDEKANLFKEYLRLLDLVKPKIFVFENVVGLMSM
 QKGQLFKQICNAFKERDYILEHAAILNALDYGVPQIRERVILVGALKSFKQKFHFPKPIKT
 HFSLKDALGDLPPIQSGENGDALGYLKNADNVFLEFVRNSKELSEHSSPKNNEKLIKIMQ
 TLKDGQSKDDLPESLRPKSGYINTYAKMWWEKPAPTITRNFTSPSSSRCIHPRDSRALSI
 RERARLQSFDPNYKFCGSASAKRLQIGNAVPPLLSVALAHAVFDLFRGKNV
 >PROKKA_00908 hypothetical protein
 MFDNNDFFKGYRNLLGFNSQNAFKEFLGAKDIQPCVDFNDLNALKKRLIEIFSAINSIYCF
 KYNEYELECEFFKDSIERVFSKIVDTHIIYKLNNQGRPEEVCFSWMRGFLVAEFFKDFIA
 CLFGAQKETIKFFGGDNFESIESFKRSPKADFLLDNHLLELVQSDFGINDIKEHKVLEA
 KRRLITDKIPTIVVHFDLFNGQVACVEISKIKDNDLNLWITRQQMEGQSVFNISQNFDDYK
 ITEIPNKPLS
 >PROKKA_00909 Catalase-related peroxidase precursor
 MKKIGLSLCLVFSGLFLKAHEVSAEEIADIFYKLNAKEPKMKINHTKGFCAGVFLPNLQ
 AKKDLDPVLLNEKEIPASVRYSLGGVAMDDKSKVRGMALKLESQNASWTMVMLNTEINFA
 KNPNEFAQFFEMRIPKNGKVDEARIKKLYEEVPSYRNFAAYTKTIGISSSVANTPYYSVH
 AFRFKDKKEKLLHARWKFPKEGIKYLNPQELKQKDSNYLLSSSQHLKNKPIEYQMYLV
 FANQNDATNDTTALWKGKHKELLIGTLKVEKYEGMGCNKDVYFPADLPKGVEAPTDPFLFQ
 IRNEVYGITFSRRQ
 >PROKKA_00910 Putative outer membrane protein
 MKLKKRKVAAVLLKREFTLPLLFTTGSGLGAVTYEVHGDFINFAKVGFNHSPINPVKGIYPT
 ETFVNLTGKLEGSVHLGRGWTVNLLGGVLGGQVYDSTRYDRWAKDFTPPSYWDKTSCTGDS
 MSLCMNATKMWQQSGPGGIINPRGIGWEYMGEWNGLFPNYYPANAYLPGGSRRYEYKAN
 LTYDSDRVHVMVMGRFDVTEQEOMDWVYQLFQGFYGTFLTNKMKFLLFSSWGRGIADGQW
 LFPYIREKPGWGIHKAGIIYRPTKNLMIHPYVYLIPEVGTLPGAKIEYDTNPEFNGMGMRN
 RTTFYVLYDYRWNNAEYGRYAPARYNTWDPFLDNGKWRGLQGPGGATLFLRHHIDINNYF
 VVGAYLNLGNPNMNLGTWGNIIAVDGIEQWVGSIIYSLGFAGIDNITDADAFTEYVKGKG
 KHGKFSWSLYQRFTTAPRALEYGIGMYLDYQFSKHVKAGLKLWLEFQIRAGYNPGTGFL
 GPNGQPLNLNGLFESSAFAQGPQDMGGIKKSITQDRSHLMTHISYSF
 >PROKKA_00911 Putative outer membrane protein
 MEIKKYFSYSLFFLLFFSLFSLKLQAYKFNMISIVGKVSSYTKFGFNNQRYQPSKDIYPTG
 SYTSLGELNLSMGLYKGLRAEVGAMMAALPYDSTAYQGNIPNGQPGSRTDPFGAGIFW
 QYIGWYAGHSGLNVQKPRLAMVHNAFLSYNYKKDKFSFGVKGGRYDAEEYDWFTSYTQGV
 EGFVKYKDTRLRVMYSARASASSDWFYFGRYYSKGALMVADLKYEKDNLKINPYFYA
 IFQRMYPAGINITYDTNPNFNDKGRFVGTGTFVGGFFPIFATPANQNDIILFQQVPLGKSGQ
 TYFFRTRFYYNKWQFGGSVYKNIGNANGDIGIYGDPLGYNIWTNSIYDAEINNIVGADVI
 NGFLYVGSQYRGFSWKILGRWTDSPRADERSLALFLSYFSNKYNIRMDLKLEYGNITKK
 GYCIGYCGMYVPVDPNGPGTQPLTHNVYSRSHIMFNITYGFRIY

>PROKKA_00912 hypothetical protein

MKLPKALNEATAGAALKYHLKRALERSHTISEFSKNLELSAQNAKFSNNTLKIIEELNNG
IKQASEEIKEKATKYEKALQELQKIDESKLTKEQQQVLKVFKGELDNGEIKKIDLNDLYI
LEQGSRNAGAKKILIKHYGEENTGGLTNDLINMSEVIKNGSVLLESFERLKNGFRYAYE
WENNGVKLRRLVDDLDNGNKIFDFYSDRNFKDFRDARPPSKAKDSRPQPTTLNEPNPTP
KPLNSQEDLLKTSENLNETTPKPTNLSPLEQANAELAKLQREQEQEFLKAKEQETK
RKEALKKKLEHERD

>PROKKA_00913 hypothetical protein

MAQLQGVGLGTPKLTPEPNPNTNALKTQEPLSPIEQANADNAKNNVKNA

>PROKKA_00914 hypothetical protein

MKSIKKPHLDHTQALSSYVSGSLGTLKKCFYGLEIFSNPLIDFTTKRSNKS RDKK

>PROKKA_00915 NYN domain protein

MKANTIIILVDWENFRDRIKQTKCVNYNIALDVIVTIRAFLLDDEWISRIYFYTTTPPFDFE
HALWDKRNDTLQNEKNPGTEMKIFTSSDIEEILQSSEATKWEKIYSDVENFQHDLASLDQ
VELRLGRTKLNAIRVEFDGSRALLEQKQVDMMLGLDIQRIAFKKIADRILIFSKDIDL
PALKLARDEGLRVDIADLSNRLSLLSQDLKYNSDKVRKLSSNEVKDKLFSIRENLTKTNW
A

>PROKKA_00916 Crossover junction endodeoxyribonuclease RuvC

MRILGIDPGSRKCGYAIISHASNKLSLITAGFINITTTTLQEQILDLEALDCLLDRIYEV
NEVAIEDIFFGYNPKSVIKLAQFRGALSILKILERIGNFSEYTPLQVKKALTGNGKAAKEQ
VAFMVKRLNITSEIKPLDISDAIAVAITHAQRLKPH

>PROKKA_00917 TonB-dependent heme receptor A precursor

MFLRSYPKLRALCLPLLTETCYSEERTLNKVTTQAKRIFTYNNEFKVTSKELDQRQSNE
VKDLFRTNPDVNVGGGSSVMGQKIYVRGVEDRLLRVTVDGAAQNGNIYHHQGNTVIDPGML
KSVEVTKGAANASAGPGAIAAGVIKMETKGAADFIPRGKNYAASGAVSFYTNFGDRETFRS
AYKSAHFDIIAYYTHQNIFYRSGATAMKNLFNPTQADKEPGTPSEQNNALIKMNGYLS
RDTLTFSWNMTRDNATRPLRSNAIGLAYPCEAPFSPDSSQGCPNVLDSETRYMYHSINSA
NNFSLQYKREAGNSFGDPRDLFTLYTSIRNAQFDPLFDPNGVYARFPTSQATAWEKENYP
CVEGAYCTPSFSDVDKPPSSQPRNLFNNTGLNLKVAHVIDEATDSLFEYGFNYQNLVSFD
ARIPKSELYRPNQVYTDDKGQKQIACSLVDNPNNDPTLCQRGKANGNIYGGYVQANYSPH
KIITFGAGVRWDAYTLYDKDWNHRYTQGFSPSAALVLSPIEPLSLKITYSQVTRGVMPGD
GVYMRQNDLRYAKNIKPEVGSNAEFNIDYSSQYFSGRAAFYQALDNFISQYAQSLIVTN
LSQAIRIYGYEVGGTFKYKGVSLNVGISRTWPTTRGYLMADSYELAASTGNVFI IKLDYT
IPKTGINLAWLSRFVTGLDYCGFDIYLPDYGMAEKPKTPTDLAKCGSQLGLVHMHKPGYG
VSNFYINWSPKTKSRWKGLLLSAVFNNVFNFYVDQTSPIYVMSPDMPGTD AIKRAIAEPG
FNARFEVAYKW

>PROKKA_00918 Catalase

MVNKDVKQTAFGAPVWDDNNVITAGPRGPVLLQSTWFLKLAADFDRERIPERVVHAKGS
GAYGTFTVTKDITKYTKAKIFSKVGKKTECFRSTVAGERGSADAVRDPGRGFAMKYYTE
EGNWDLVGNNTPVFFIRDAIKFPDFIHTQKRDPQTNLPNHDMVWDFWSNVPESLYQVTWV
MSDRGIPKSFRHMDGFGSHTFSLINAKGERFWVKFHFETMQGVKHLTNEEA AEVRKYDPD
SNQRDLFNAIARGDFPKWKL SIQVMPEEDAKKYRFHFPDVTKIWYLDYPLMEVGIVELN
KNPENYFAEVEQAAFS PANVVP GIGYSPDRMLQGR LFSYGDTHRYRLGVNYPQIPVNKPR
CPFHSSSRDGYMQNGY YGSLQNYTPSSLPGYKEDKSARDPKFNLAHIEKEFEVWNWDYRA
EDSDYYTQPGDYRSLPADEKERLHDTIGESLAHVTHKEIVDKQLEHFKKADPKYAEGVK
KALEKHQKMMKDMHGKDMHHTKKKK

>PROKKA_00919 hypothetical protein

MKRRDFIKTTALGATGAVLGAQILQAEESKGSVAKYKIEAQYSIDFDSA EHTSLFIPMPS
VVASNVHLQGNHASYESMLNFGVPYLQVDFLKSAQKKQVHLSYEIASYQLNERL FETSDF
VAMGRYERDDASVVS IANQLKGSTPKESVRNFYAFIKHDMPKRQKALEGKENLPKRESLP
WFATISQESMFVSLCHACGIKSAEVQGLKLGQNSVVKNA PRVEVYLKDSFLAFDFQNNHK
EVFIPLNRHKDMQLDSALLATFGDAFALVDGRDLGNYESKLF EKRVSYTIV

>PROKKA_00920 hypothetical protein

MREFFKKLGTEYASKLFLVYWLWRWMLSALVMLPFMEIFYFYFNFLWLNLFLGQTIGAVIF
FKLDKLIFS

>PROKKA_00921 hypothetical protein

MQDLPPCPKCNDAYTYHDGAQLICSSCLYEWNEANEDEELIVKDCHNNLLQNGDSVILI
KDLKVKGSSSLVLKKGTKIKNIKLVNSDHNVDCKIEGQSLSLKSEFLKKA

>PROKKA_00922 CDP-diacylglycerol pyrophosphatase
MKKAGFLFLAVMAIVAMILNAKDPNVLRKIVFEKCLPNYEKSNPSPCIEVKPDAGYVVL
KDINGPLQYLLMPTTHISGIESPLLLDPSTPNFFYLSWQARDFMSKKYGPIDSAISLT
INSKKGRSQNHFIHISCISLDAHKQLDNNLKKSTAVGRHYQVA

>PROKKA_00923 CDP-diacylglycerol pyrophosphatase
MNGHKYLARRVTESELAQKSPFVMLNKEAPNAHKRMGDYGLAVVQQSDNSFVLLATQFNP
LTLNRASAEIQQDHECAILR

>PROKKA_00924 Flagellar hook protein FlgE
MLRSLWSGVNGMQAHQIALDIESNNIANVNNTGFKYSRASFVDMLSQVKLIATAPYKNGL
AGQNDFSVGLGVGVDAATTKIFSQGNIQNTDVKTDLAIQGDGFFIISPDRGITRNFTRDGE
FLFDSQGS�VTTGGLVVGWVRNGSDTGNGKSDTDALKVDNTGPLENIRIDPGMVMPARA
SNRISMRANLNAGRHADQTAAIFALDSSAKTPSDGINPVYDSGTNLAQVAEDMGSLYNED
GDALLLNENQGIWVSYSKAKMVKDILPSAENSTLELNGVKISFTNDSAVSRTSSLVAAKN
AINAVKSQTGIEAYLDGKQLRLSENTNELDGDEKLKNIVVTQAGTGAFANFLDGDKDVTAF
KYSYTHSISPNAIDIGQFRTTEDLRALIQHDANIVKDPSLADNYQDSAASIGVSVNQYGMF
EINNKNKNVIKENLNI FVSGYSSDSVTNNVLFKNAMKGLNTASLIEGGASASSSKFTHA
THATSIDVIDSLGTKHAMRIEFYRSGGAENFRVIVPEPGELVGGSAARPNVFEGGRLHF
NNDGSLAGMNPPLLQFDPKNGADAPQRINLAFGSSGSFDGLTSVDKISETYAIEQNGYQA
GDLMDFVRFDSDGVLLGAFSNGRTLALAAQVALANFANDAGLQALGGNVFSQTGNSGQALIG
AANTGRRGSISGSKLESSNVDLRSRLTNLIVVQRGFQANSKAVTTSQILNTLLNLKQ

>PROKKA_00925 Hydrogenase/urease nickel incorporation protein HypA
MHEYSVVSSLIACEEHAKKNQAHKIERVVVGIGERSAMDKSLFVSAFETFREESLVCKD
AILDIVDEKVELECKDCSHVFKPNALDYGVCEKCHSKNIVITQGNEMRLLSLEMLAE

>PROKKA_00926 hypothetical protein
MQEELNAYQQEIEDTRGVLKKIRLELQVQVEILRKKKSVLKGKQVEICQKKSEKENSRLN
KEAQNTEEDVIFPKALEEEVEVFTSDNQVIMAKPSKRVFDEGIYLYQYRSVLRENRLKNNHL
SKKDFENSLKIELRDLHKEIKLYQVQNNLLKDK

>PROKKA_00927 Lipid-A-disaccharide synthase
MPTILVSÄLEASSNIHLEELRRNLPKDYRFIGVFEGKNALYSPREFSVMGFRDVIGRLGF
LLKAHKEMVQLAKQADMVLLMDSSSFNIPLAKKIKKQDPHKKIMYYILPQVWAWKKWRTK
SLEKYCDFLGAILPFEVSYYQKKAQYVGHPLLDEIKYYKKDIKGETLVFMPGSRKSEIAK
MFPLFVKVAQILEQNEGFKRRVLVVPSSFKGDLKALYGEDIKLFEISYDAHKSLEAEF
AFICSGTATLEAALIGTPFVLAYRAKTMDFLIARMLVNLHYIGLANIFYNALNDETPLGLG
ESQLHPELIQHFLSVEGLLKVYEEMDRERYFKESLRLREYLKHGSARKVANEMAFLNLT

>PROKKA_00928 Transcription elongation factor GreA
MNKEPMSMHGYNKICAELKQLKEVERPNIVKEIDIARGHGDLENAYHAAKEKQRFIEA
RIVDLSEIVANAQVIDPSVLAHNKVVSFGSTIKILNLDNDKEFSYTIVGSVESAPAKGLIS
FGSPIAKSLIGKSKGDTVSIQLPNGESDFEILDIIYYKEICFDEN

>PROKKA_00929 Deoxyuridine 5'-triphosphate nucleotidohydrolase
MKIKIQKIHPNALIPKYQTEGSSGFDLHAVEEVTIKPHSVGLVKIGICLSLEVGYELQVR
TRSGLALNHQVMVLNSPGTVNDNDYRGEIKVILANLSDKDFKVQVGDRIAQGVVQKTYKAE
FIECERLDETSRGSFGFGSTGVSKA

>PROKKA_00930 hypothetical protein
MNIKENLEQVRNEFKSDEKLLEGAFRLEKFFKRYKWVLLFIVVAFIAYLGDTKLQDYKHE
QTRERITQIYNEVLESNNIALQKRLKEVAPELYDLYQFARASERNANEFKKLSQSSNE
IVKTFAQYSYASLSRDKNLLEKSPVLKEMSALQEVNLLYEENSKDAIKKAHQSLSTIPLS
SSLYAII SVLKHYGMLEDIQQNPSKPTNLKKETIQGTH

>PROKKA_00931 Plasminogen-binding protein PgbB
MMNKPFLILLIALIVFSGCNMRKYFKPAKHQVKGEAYFPNHLQESIVSSNRYGAILKNGA
VIGDKGLTQLRIGKNFNYESSFLNESQGFFILAQDCLNKIDKKT SKNKVAKSEETELKLK
GVEAEVQDKVCHQVELISNNPNASQQSIVIPLETFALSASVKGNLLAVVLADNSANLYDI
TSQKLLFSEKGSPTTINSLMAMPIFMDTVVVFPMLDGRLLVVDYVHGNTPIRNIVISS
DKFFNNITYLIVDGNNMIASSTGKRILSVVSGQEFNYDGDIVDLLYDKGTLYVLTLDGQIL
QMDKSLRELNSVKLPFASLNTIVLNHNKLYSLEKRGYVIEVDLNDNDFDSYNVYKTPTIGSF
KFFSSNRLDKGVFYDKNRVYYDRYYLDYNDFKPKLYPVVEKSASKKSQKGEKGNAPIYLO
ERHKAKEKPLEENKVKPRNSGFEEEEEVKTRRPEPIRDQNNATQQGEKQNAKPKENKASKEE
NKPNSKEEKRRLLKEEKKAKAEQRRAREFEQRRAREHQRERDEKELEERRKALEMNKK

>PROKKA_00932 Type III pantothenate kinase

MPARQSFTDLKNLVLCDIGNTRIHFQNYQLFSSTKEDLKRLGIQKEIFYISVNEENEKA
LLNCYPNAKNIAGFFHLETDYVGLGIDRQMACLALNNGVVVDAGSAITIDLVEGKHLGG
CILPGLAQYIHAYKKSAKILEQPFKALDSLEVLPKNTRDAVNYGMILSVISCIQHAKDQ
KIYLCGGDAKYLSAFLPHSVCKERLVFDGMEIALKKAGILECK

>PROKKA_00933 hypothetical protein

MQMMQNLSFLGMFLAALSMSLGHCVGMCGGIVSAFSQIRFSKVTSSFSYQLTCHALYNVGR
ISTYMLLGAITAGLGHSLSVSMGFRGVLLMSMGVVLILLALLGSKAEKLSFQIPFISFLM
KKTLSQSNILGLYFLGVNLGFLPCMMVYSFLASVILSHSAFMGAMLGLSFGLGTSVPLFL
MGIFLSKISISYRKFFNLLSKGLMGVFGLYVLYMGIMLINHQTPHAMHHQNSTTQHDHKG
AHSHEH

>PROKKA_00934 D-glycero-alpha-D-manno-heptose-1,7-bisphosphate 7-phosphatase

MNTNKALFLDRDGIINIDKGYVSQKEDFEFQKGIFELLNHAKSLGYKLLITNQSGINRG
YYTLKDFENLTEYLQESLFKELGFNLGDIYFCRHAPEENCACRKPKPFLILQAAKEHQIC
LERSFMIGDKESDMLAGLNAKVNNLLTTNFKTSHSWIQCKDLKEMMELIK

>PROKKA_00935 ADP-L-glycero-D-manno-heptose-6-epimerase

MRYIDDGLENQTILITGGAGFVGSNLAIFYQENHPKAKVVILDKFRSNTLFSNNRPSSLG
HFKNLIGFKGEVIVADINNPLDLRRLKHLFDYLFHQAAVSDTTMLNQELVMKTNYQAF
NLLIARSKKAKVIYASSAGVYGNTKAPNVVGSNESPENVYGFSKLCMDEFVRFHSNDNI
QVGLRYFNVYGPREFYKEKTASMLVQLALQALAFKEVKLFEFGEQLRDFVYIEDVIQANV
KAMKAQKSGVYNVGYSAQARSYNEIVSILKEHLGDFKVSYIKNPYAFFQKHTQAHIEPAIL
DLDTPLYDLESIGIKDYLPPIHAIFKEQCA

>PROKKA_00936 Bifunctional protein HldE

MKKILVIGDLIADYYLWGKSERLSPEAPVPVLEVKKESKNLGGAANVANNLISLKAKVFL
CGVVGDDLEGKHFISALKTRGIDTSGVLIDKTRCTTLKTRIIAQNQIARVDKEIKDPLN
ADLRKKLLDFFTEKIQEIDGVILSDYNKGVLDFELTQKMITLANQHKLILCDPKGKDYS
KYSHASLITPNRTELEHALHLKLDSEANLSKALQILKETYHIAMPLVTLSEQGIAFLEKG
ELVNCPTIAKEVYDVTGAGDTVIASLTLSLLESKSLKDACEFANAAAAVVVGKMGALAS
LEEIALILNQTHPKILPLEKLLTLEHNQQKIVFTNGCFDILLHKGHASYLQAKALGDIL
IVGLNSDNSIKRLKGDKRPVSEKDRAFLASLSCVDYVVVFEEETPIKLIQALRPDILV
KGADYLNKEVIGSEFAKETRLMEFEEGYSTSAIIEKIKRTHND

>PROKKA_00937 Phosphoheptose isomerase 1

MIDDLIKKEFLAHKEALEKSLESLSQEALKQSVHLLIETLENQGKILICGNGGSASDAQHF
AAELTGRYKLERKGLSAISLSTDTALTAIANDYGYEEVFARQVEALGVKNDVLIGISTS
GNSKNVLKAYEKAKDLGVKTLSLAGRDGGKMKPLSDMALIVPSDDTPRIQEMHILMIHIL
CDCIERHFARKN

>PROKKA_00938 GMP reductase

MKVFDYEDVQLIPNKCIVNSRSECDTTVILGKHAFKMPIVPANMQTIINEPIAEFLAENG
YFYIMHRFDGAARIPFVKMKERQWISSISVGVKKEEYLLIEELAKQGLAPDYITIDIAH
GHSNSVIEMIQHIKTHLPETFVIAGNVGTPEAVRELENAGADATKVIGPGKVCITKIKT
GFGTGGWQLAALRWCAKAARKPIIADGGIRTHGDIAXSIRFGATVMVIGSLFAGHEESSG
ETKIENGVAKEYFGSASEFQKGEKKNIEGKKIWIQHKGLLKDITLVEMHQDLQSSISYAG
GRDLEAIRKVDYVIVKNSIFNGDAI

>PROKKA_00939 putative ABC transporter ATP-binding protein YheS

MLQTINLTQRYATKKLFENVNIKLDKNKRYGLIGANGAGKSTFLKILSKNIDCSSGEVII
TSGMRMGVLGQDQYAFEDLSLKDAVLIGNKRLYDAIKEKERLYTEGDLSDDKVNARLGEL
ETICVEEDPMYECVVVEKILEDLGIPSSKHNDLMKTLPSDDKFILLAQVLFPPKDILL
LDEPTNNLDLNAIEWLENNLKRHEGTMVVISHDRHFLNAVCTHILDLDLDFHSVREFSGNYD
DWYIASTLMIKQQEAERNKKLKEKEELEKFIARFSANASKAKQATSQKQLDKLDIQSLA
VSSRRDPSIIFKPKRTIGNEALECENISKSVDGQIILNQVSLKVMPKDKIALIGPNGVGK
STLCKILVEELKPDITGVVKGATVSKGYFPQNVSEEISGEETLYQWLFNFNKKIESAEVR
NALGRMLFNNGEEQEKCVNALSGGEKHRMVLSKLMLEGGNFLVLDEPTNHLDEAIIALGE
ALFKFDGTVICVSHDRELIDAYANRIIELVSPKASIIDFKGSYEEYLASK

>PROKKA_00940 hypothetical protein

MKPRDIGIIQSVLEITGPIKVTEIYHKAKELFEKGEITNMFDGGETPHQSVSSSIYAL
NKGEELPFKKVQENPTLIALKGAAKEPVLNAQKPSVKIAHERDLHPFLTYMAINNENLKC
YTKTIFHEESLKSPPKMDRWLYPDMVGVRFLHAESNENLIAFSKKFDTLPVKLVSFELK
KEISVHNCRECYFQAISNSSWANEGYLVGRHIDTHNPQLMDLLKRLHASFGIGVIDLRD
EDKSAILLNAKYKEKIDYTVASELSAKNEKFSGFLKSVVDYDPKTQHRYKDEFDEVKKKE

KLYPNPSLSF

>PROKKA_00941 phosphatidylglycerophosphatase B

MVFDRTISVREKKAATLGIVGIIFFILFGIVISGVAFQKEWVQQDLFFIDLIRNPAPI
QKSTWLSFVFFSTWFAQSKLTTPIALLLGLWFGFQKRIALGVWFFFSILLGEFTLKSLLK
LVARPRPVTNGELVFAGHGSFSPSGHALASALFYGSLALLLCYSNANNRTKTIIAVVLLFW
IFLMAYDRVYLG VHYPSDVLGGFLLGIAWSCCSLALYLGLKRPYKAA

>PROKKA_00942 HaeIII restriction endonuclease

MSDKSNNQGRAYEYAWCLALEQKLSVLKKVIVDKQNGFNTCYRAYESLEKSLQERYLESA
KQGVLLLLDCEPLLSEVIGSSQNEITLSLQKDKLGEIGDIRDILYFDRFCIGLSIKHNH
DAVKHSRLSKDLDFGGKWLGVRSQNYKDTIKPLFERLENAKKEGMLWRDFPNKEQEIYA
PLLQAFKKEVLRIDENKKNKVPQKMVEYLLGKYDFYKAILLEREQKTKLEAYHFNNTLNR
SMKNKPKRIIPLSKLPTRMIYFDFKPKSFNTLELVLNEGWSFSLRIHNASSKVEPSLKFD
IKLLSIPVSIASFVIVGF

>PROKKA_00943 Modification methylase HaeIII

MNLLSLFAGAGGLDLGFEKAGFKIVVANEYDKNITPTYRLNHKNQTQLEKDIKNLQTSEI
NFSVDGIIIGPPCQSWSEAGNLKGIDDARGQLFYEYLRLLKELKPKFFLAENVRGMLAQR
HKTSVKNILNAPKECGYEVNTHLVNAKDYGVAQERLRVFIYIGFRRDLKVNYIFPKGSTHL
KKLTLKDVIWDLKDSVVCALAKNKRNPAINNHEYFIGSYSPIFMSRNRVKNWDEQAFTI
QASGRQCQLHPQAPKMAKFGKNDCRFIENYQHLYRRLSVRECARIQGFDDDFYFVYENLN
DAYKMIGNAVPVSLAKEIAISIIYKVLH

>PROKKA_00944 putative type I restriction enzyme P M protein

MENPNNNQAFLERNELHNTIWKVANELRGSDVGWDFKQYVLGILFYRISENMTHYINK
EERKHDPSTFYAKLSDEEAERAREELIEEKGFPIPPSALFCNALKNAPNNEDLNVTLQNI
FNEIEKSSLGAPSEENVKGLFADLDVNSNKLGSSSHKIRVEKLT KILEAIGGMQLGDYQQS
GIDVFGDAYEYLMAMYASNAGKSGGEFFTPQEVSELLAKIALHNQESVNVKYDPCCGSGS
LLLQFSKVLGDKNVSKGYFGQEINLTTYNLCRINMFLHDINYSKFHIALGDTLLDPKHED
DEPFDAIVSNPPYSTKWVGDNPLLINDERFSPAGVLAPKNAADLAFTMHMSYLSNSGT
AAIVEFPGVLYRGNAEAKIREYLVKENVIDCVIALPDNLFFGTSIATCILVLKKNKKDDT
TLFIDASKEFVKEGKKNKLKERNREKILQTYIERKEVKHFSSLANIEKIKENDYNLSVNR
YVEQEDTKEIIDIKALNSEISQIVEKQSALRSSLESIIKELEA

>PROKKA_00945 Putative type-1 restriction enzyme specificity protein MPN_089

MHKIERLLQTLAPKGVEFKTLEEVFEIKNGYTPSKNNLEFWKNGTIPWFRMEDLRENGRI
LKDSIQHITPKALKGKKLFPKNSIIISTTATIGEHALLIVDSLANKRFTFLSKKANCNLA
LDMKFFFYQCFLLEGEWCKNNINVSFGFASVDMTAFKKYKFPIPPLEVQQEIVKILDAFTL
NTELKARKKQYEEYQNMLLDFKDIKQNHKDAKEKLAQKTYPKRLKTLQTLAPKGVEFRK
LGDIGEFYGGVLGKNKKSFSQGNKFYVPYINVFNNPQLDLNALESVQIGDKEKQNTIQLG
DVLFTGSSSENLEDCAMSCVVTQKIEKDIYLSNFCFGFRFFDENLNFNPSFLKHFLRDYNFR
KNISKVANGVTRFNVSKQLLSKITIPIPPLEIQQEIVKILDQFSILTDDL LAGIPAEIKA
RKKQYEEYREKLLTFKPLQNK

>PROKKA_00946 Type I restriction enzyme EcoR124II R protein

MNYETIAESNESTVVAEFHSSNEKKSSYSEAELEMAFIALLEKQGYEFKQIHNEKELKD
NLKEQLEKLNNHHTPKEDWTLYSQFIANKNDDYKAKTKKIQEDPIFNLTLENGKTKNIK
IIDKKNIHRNALQVIHQYSNKGKQYQNRDYVSILVNGPLPLVHVELKKRGVAIREAFNQIK
RYKRDSFSAEDGLFEFVQIFVISNGTSSKYYSNTTRIAQLEKNHKADTFEFTNYWADSKN
HNIEDLMDFAQAFFAKRSLLNVLTRYCVFTSEEVLVMPYQIVAAERILEKIKTAQNSK
TYGKNQNGGYIWHHTGSGKTLTSFKSATLAKELEGVSKVLFVVDRKDLDYQTMKEYDKFQ
KDCANSNTSTKILKQQLDFNAKIIITTIQKLDKFVKSHLKGHAIFNEEVVIFDECHRS
QLGSMHQAITKAFKKYHLFGFTGTPIFAQNCDKNNPLGTTEQKFGKCLHQYTIIDAIRDK
NVLPFRVEYHNTIKAKEGIIIDNKVRVDEKNALLDTRRIKEITKCILERFNQATKNKKFN
SILACSSIEALKKYYQAFKEEKHDLKIATIFSYSANEEDTLEDENNESACRLDKSSRDF
LEGAIADYNGMFGVSFDTSDQKFQSYKDLSSQVKERKIDLLMVNMFLTGFDAATRLNTL
WVDKNLKYHGLIQAFSRANRILDSVKTHGNIVCFRDLEQDLNDALMLFGNKDAQSVVLLR
KYEDYLGKGYTDNHKEYEGYGLIKRLLSEFPLKEPIVSESQKKDFIKLFGKILKLENILN
SFENFKKDDHINPRDFQDYQSKYLDIFYDAMRPEKGDKEEINDDLIFEIELIKQVEVNID
YIHLNIEEFAKEHGEVIEQGVKTKIEPIINSSIELRNKKDLIMGFIDKYNKDQEVHAYFQD
YIHQKREEEFQNIIEENRLNEEKAYSFMQHAFKGEISFNGTKFPEIIEEKPSMFSQNSR
YQEVKEKVAASLSRFFHRFCDLTSAIFKKNGVKKDEVNEK

>PROKKA_00947 Hydroxyethylthiazole kinase

MKELRQKRPLVHNITNYVAAQFVANGLLALGASPLMSDAIDEMQDLAKISDALAINIGTL
NDRAILCAKEAIKHYKALNKPIVLDPVGCSASALRYDTSLELLKSGGISVLRGNAAELGS
LVGISCESKGLDSNDANTPVEIIKRAAQKYSVIAVMTGKTDYVSDGEKVL SITGGSEYLA
VITGAGCLHAAACASFLGLKKDPLDSMVQLCVLYKQAAFNAQKKVLENNGSNGSFLFYFL
DALSLPIELEN SLIKEEL

>PROKKA_00948 Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase

MKIYPQVL SIAGSDSGGAGIQADLKAFQTLGVFGTSVITCITAQNTQGVHGVYPLSVES
VKAQILAIRDDFSIKAFKMGALCENNQIIECVADTLETCDFGLCVLDPVMVAKNGALLLEE
EAILSLKKRLLPKTNLLTPNLPEVYALTGVQVRDDKSASKAMGVLRDLGVKNAVIGGHT
ERFQGEFSNDWVFL EDAEFILNAKRFTKNTHGTGCTLSSLIVGLLAQGLDLKNAISKAK
ELLTII IQNPLNIGHGHGPLNLWSIKKLV

>PROKKA_00949 Thiamine-phosphate synthase

MFDANCLKLMFVAGSQDFYHIKGDKNDRINALLETLELALQSQITAFQFRQKGDALQDP
IEIKQLALE CQKLCQKYGVPIVNDVQLALEL KADGVHVGQEDMAIEEVITLCKRFLFI
GLSVNTLEQALKARHLDAVAYFGVGPIFPTPSKKDAKEVVGINLLKKIHDSGVKKPLIAI
GGITMHNASKLREYGGIAVISAITQAKDKALAEKLLNNA

>PROKKA_00950 hypothetical protein

MLEALNALNQLNALH SKNATHHFNAALPILLKVLEKQDKDLFLLQVGNRIIPTKSEQELK
INQPYFATMQRDRLGDIVLKNLVPAPKILDALDDLPVLEMKKIKEILNAKDNTPLKEYKE
LLSEKLIHAKNSQEFLNTANMLLSLQSQVLSFVIENERKKAFLQVKAKKQSVDFYALYPN
LGEIGGVIYLKEKEKQLFLKTTLQRTKEVLKEAQNTLLGFSSVEIVCEKTPMLFAFEERL
LDTLG

>PROKKA_00951 Coenzyme A biosynthesis bifunctional protein CoaBC

MNFLEDLFYPLRLL ENKRVLLLVSGSIAAYKSLELTRLLFKSGASIQVVM SKGAKKFIKP
LSFEALSHHKVLHDRNEK WYYNHQNALHHNHIACAINADLLIFAPLSANSLSKIAHALAD
NIVSATFLACTSPKILAPSMNTNMLNSPIIQSHLRLKDTNHIILDTQNRL LACDTKGDG
AMAEPLEILFKAAQTLLKDAYFENREVIIMGGASVEKIDSVRTISNLSSGIQASALALAL
YFKGARVTLIASNFPTPLPKGIKSVLVSDTASYENALNXAANNLQKHALKPLL FNLA AIS
DYVPKTSFNYK LKKSEIGETLNIECVQNKDLLASINPNQFVKIGFKAEDNQQNAIKNAQN
LLKPLEDNKGKCSMAALNLIKDSRPFSGLENELWLFSHQKTQKIPSMNKLEASF KILDFI
KDNAL

>PROKKA_00952 UDP-N-acetylglucosamine 4,6-dehydratase (inverting)

MPNHQNMLDNQTILITGGTGSFGKRFVRKVLDTTNAKKIIVYSRDELKQSEMAMEFNDPR
MRFFIGDVRDLERLNYALEGV DICIHAAALKHVPIAEYNPLECIKTNIMGASNVINACLK
NAISQVIALSTDKAANPINLYGATKLCSDKLFVSANNFKGSSQTQFSVVR YGNVVGSRGS
VVPFFKKLVQNKASEIPITDIRMTRFWITLDEGV SFVLKSLKRMHGGEIFVPKIPSMKMT
DLAKALAPNIPTKIIGIRPGEKLHEVMIPKDESHLALEFEDFFIIQPTISFQTPKDYTLT
RLHEKGQKVAPDFEYSSHTNNQWLEPDDLKLL

>PROKKA_00953 long-chain fatty acid outer membrane transporter

MKNFSPLCCFKKLK RHLIALSLPLLSYANGFKIQEQLNGTALGSAYVAGARGADASFY
NPANMGFTNDWGENRSEFEMTTT VINIPAFSFKVPTTNQGLYSVTS LQIDKSQQNILGII
NTIGLGNILKALGNTAATNGLSQA INRVQGLMNL TNQKVVTLASNPDTQIVNGWTGTTNF
VLPKFFYKTRTHNGFTFGGSFTAPSGLG MKWNGKGGEFLHDVFIMMVELAPSMSYTVNKR
FSVGVGRLRGLYATGSFNNTVYVPLEGASVLSAEQILNLPNNVFADQVPSNMMTLLGNIGY
QPALNCQKAGGDMNDQSCQEFYNG LKKIMGYSLIKASANLYGTTQVVQKSNGQGVSGGY
RVGSSLRVFDHGMFSVVYNSSVT FNMKGGLVAITELGPSLGSVLTKGSLNIN VSLPQTLS
LAYAHQFFKDHLRIEGVFERTFWSQGNKFLVTPDFANATYKGLSGTVASLDSETLKKMVG
LANFKSVMMNGAGWRDTNTFRLGV TYMGKSLRLMGAIDYDQAPSPQDAIGIPDSNGYTVA
FGTKYNFRGFDLG VAGSFTFKSNRSSLYQSPTIGQLRIFSASLG YRW

>PROKKA_00954 hypothetical protein

MTLFFVSCSKHPFSKQTPKTREQIRQEEANKKREETLNALRQFR LIYINTPVFRFYDYGT
IKTDKDHNI EVTLYKLSQRVGDIYMTKRNICFSQKCSAKWIAARDLFGKVS YGDLFDDIV
LGRDIFKGLGKRHLTPEYVIQR FQKSGEIILYERKNDLISFQNL TQKIAIRIEPYEPSLQ
DLEDNENADSELQ

>PROKKA_00955 hypothetical protein

MRLHTAFFGINSLLVATLLISGCSLFKKRNTNAQLIPPSANGLQAPIYPTNFTPRKSVQ
PLPSPRLENNDQPIISSNPTNAIPNTPI LTPNNVIELNAVGMGVAPESTISPSQALALAK
RAAIVDGYRQLGEKMYGIRVNAQD TVKDMVLQNSVIKTRVNALIRNAEITETIYKDGLCQ

VSMECLKLDGRIWYRILSGARG

>PROKKA_00957 DNA-binding protein HU

MNKAEFIDLVKEAGKYNSKREAAEAIASFTLAVQTALSKGESVELIGFGKFETAEQKGKE
GKVPGSDKTYKTEDKRVPKFKPGKILKQKVEEGK

>PROKKA_00958 GTPase Der

MNTSHKTLKTIAILGQPNVGKSSLFNRLARERIAITSDFAGTTRDINKRKIALNGHEVEL
LDTGGMAKDALLSKEIKALNLKAAQMSDLILYVVDGKSIPSEDLKLFREVFKNPNCFL
VINKIDNDKEKERAYAFSSFGIPKSFNISVSHNRGISALIDAVLSALDLNKIEQDLAD
ILESLENNAPKEETKEEEIIQVGIIIGRVNVGKSSLLNALTKKERSLVSSVAGTTIDPIDE
TILIGDQKICFVDTAGIRHRGKILGIEKYALERTQKALEKSHIALLVLDVSAPFVELDEK
ISSLADKHSLGIILILNKWDIRYAPYEEIMATLKRKFRFLEYAPVITTSCLKARHIDEIK
HKIIEVYECFSKRIPTSLLNSVINQATQKHPLPSDGGKLVKVVYATQFATKPPQISLIMN
RPKALHFSYKRYLINTLRKEFNFLGTPLILNAKDKKSTQQN

>PROKKA_00959 hypothetical protein

MFLVKKIGVVVVVLIGFLACSQERFIQIAKKAQEENDGSKRPSYVDSYEVFSETIFLQ
NMVYQPTTEERDSFAQLTKDGNDSFNPETSVILLNEPSDNNEKNPPLYPNESDNNTNNANN
DTKNPFLYKPKKEAKNPKLIEYSQQDFYPLKGDGDIIMSKEGDQWLIEIKSKALKRFLKDQ
NDKDRQIQTFTFNDTKTQIAQFKGKISSYVYTTNDSDSLRLPFYESFLEKKSDNVYTIE
NKALDTMEISKQCMVLKKHSTDKLD SQHKAISIDLDFKKERFKSDTELFLECQS

>PROKKA_00960 Spermidine synthase

MWITQEITPYLRKEYTIEAKLLDVRSEHNILEIFKSKDFGEIAMLNRQLLFKNFLHIESE
LLAHMGGCTKKELKEVLIVDGFDELAHQLFKYDTRIDFVQADEKILDNFISFFPHFHEV
KNNKNFTHAKQLLDLDIKKYDLILCLQEPDIHRIDGLKRMLKEDGVFISVAKHPLLEHVS
MQNALKNLGDFFAVAMPFVAPLRILSNKGYIYASFKTHPLKDLMTPKIEALKSVRYYNED
IHRAAFALPKNLQEVFKDNIKS

>PROKKA_00961 Dephospho-CoA kinase

MILKNAIALTGGIGTGKSTTLKILESQGYKILDADKIAHQLLQEHRLEIAQHFGSDILEK
GILNRKKLGAIVFQDPNKLKWLDFLHPLIRECMLKKACELEKNHQAYFLDIPLFFFEVGG
KKRYPVSRVVLIIYAPRVLQIERLLERDKLKEAEILQRLACQMDIEQKRAMSDYIIDNSSS
LKDLNKQVERFLKTLL

>PROKKA_00962 Glutamyl-tRNA(Gln) amidotransferase subunit A

MITLKQALSLSQDELETLKNEIDAKVRASDLNAYIKAPSLNGASAKGVPILIKDNISVKG
WEITCSSKILEGYVAPYHASAIENLHQNSMAGFGLSNMDEFAMGSTTESSCYGITKNPRD
KNRVPGGSSSGSAAAVAGGLAVAALGSDTGGSIRQPASYCGCVGLKPTYGRVSRYGLIAY
CSSFDQIGPITQNVEDASILFDAISGYDSKDSTSANLKPTQTFKNLNRDKRFKIAILRDH
IKDASNEVQLAYENTIKALKEMGHEIVEKKMLDSHYQISIIYIIISMAEASSNLARFDGVR
YGRRAENVKDLKELYLKSRSSEFGDEVKRRIMLGNFVLSSGYDAYYLKAQQMRLMIKEQ
YNKIFEEVDLIFTVPVAPTS AHLFNHASPLEMYLSDIYTIGANLSGLPALSLPVAKDPLG
LPIGMQFIAKAFDEQSLLDVSYALEQELDLKLD

>PROKKA_00963 Inosine-5'-monophosphate dehydrogenase

MRILQRALTFEDVLMVPRKSSVLPKDVSLKSRLTKNIGLNIPFISAAMDTVTEHKTAIAM
ARLGIGIGIVHKNMDIQTVKEIAKVKKSESGVINDPIFIQAHRTLADAKAIADNYKISGV
PVVDDKGLLIGILTNRDVRFETDLSKKVGDVMTKMPLVTAHVGISLDEASDLMHKKHIEK
LPIVDKDNVLKGLITIKDIQKRIEYPEANKDDFGRLRVGAAIGVGQLDRAEMLVKAGVDA
LVLD SAHGHSANILHTLEEIKKSLVVDVIVGNVVTKEATSDLISAGADAVKVGIGPGSIC
TTRIVAGVGMPQVSAIDNCVEVASKFDIPVIADGGIRYSGDVAKALALGASSVMIGSLLA
GTEESPGDFMIYQGRQYKSYRGMGSIGAMTKGSSDRYFQEGVASEKLVPEGIEGRVPYRG
KVS DMIFQLVGGVRSSMGYQGAKNILELYQNAEFVEITSAGLKESHVHGVDITKEAPNYY
G

>PROKKA_00964 ATP synthase subunit a

MEHRVFTIANFFSSNHDFITGFFVVLTAVLMFLISLGASRKMQMVP MGLQNVYESIIIGAI
LSVAKDII GEELARKYFPLAGTIALYVFFSNMIGIIPGFESPTASWSFTLV LALIVFFYY
HFEGIRVQGGFFKYFAHFAGPVKWLAPFMFPFIEIISHFSRIVSLSFRLFGNIKGDDMFLLI
MLLLVPWAIPVAPFMVLF FFMGILQAFVFMILTYVYLAGAVLTDEGH

>PROKKA_00965 RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)

MKNIIYVGNLVYSTTSEQVKELFSQFGKVFNVKLIYDRETKKPKGFGFVEMQEEGVSEAIA
KLDNTDFMGRITIRVTEANPKKS

>PROKKA_00966 Lipooligosaccharide biosynthesis protein lex-1

MRVFIIISLNQKVCDDQFDLVFRDTTTTLLNNINATRHKAIQIFDAIYSKTFEGGLHPLVKKHL
HPYFITQNIKDMGITTNLISEVSKFYALKYHAKFMSLGELGASHYSLWQKCIELNEA
ICILEDITLKEDEFKEGLDFLEKHIQELGYVRLMHLLYDTSVKSEPLSHKNHEIQERVGI
IKAYSHGVGTQGYVITPKIAKVLKHSRKWVVPVDTIMDATFIHGVKNLVLQPFVIADDE
QISTIAKKEEPYSPKIALMRELHFKYLKYWQFV

>PROKKA_00967 Thioredoxin reductase

MIDCAIIIGGGPAGLSAGLYATRGGVKNVLFEEKGMPGGQITGSSEIENYPGVKEVVSGLD
FMQPWQEQCFRFGKHEMTAVQRVSKKDSHFVILAEDGKTFEAKSVIIATGGSPKRTGIK
GESEYWGKGVSTCATCDGFFYKNKEVAVLGGGDTAVEEAIYLANICKKVYLIHRRDGFRC
APITLEHAKNNDKIEFLTPYVVEEIKGDASGVSSLSIKNTATNETRELVPVPGFFIFVGYD
VNNAVLRKQEDGSMCLKCDEYGSIVVDFSMKTNVQGLFAAGDIRIFAPKQVVCAASDGATA
ALSVISYLEHH

>PROKKA_00968 Thioredoxin

MSHYIELTEENFESTIKKGVALVDFWAPWCGPCKMLSPVIDELASEYEGKAKICKVNTDE
QEELSAKFGIRSIPITLLFTKDGVEVHQLVGVQTKVALKEQLNKLGG

>PROKKA_00969 hypothetical protein

MRFLNNKHREKGLKAECEACGFLKTLGFEMVERNFFSQFGEIDIIALKKGVLHFIEVKSG
ENFDPIYAITPSKLKMIKTIRCYLSQKDPNSDFCIDALIVKNGKFELLENITF

>PROKKA_00970 Homoserine dehydrogenase

MKKRLNIGLVGLGCVGSAVAKILQENQEI IKDRAGVEIKIKKAVVRDMKKHKGYSFEISD
DLESLEDEEIDIIIVELMGVEAPYLLAKKTLAKQKAFVTANKAMLAYHRYELEQTAKNT
PIGFASVCGGPIIKALKDGLSANHILSFKGILNGTSNYILSQMFKNQASFKDALKDAQ
HLGYAELNPEFDIKGIDAAHKLILASLAYGIDAKLEEILIEGIEKIEPDDMEFAKEFGY
SIKLLGIAKKHQDCIELRVHPSMIKNECMLSKVDGVMNAISVIGDKVGETLYYGAGAGGE
PTASAVISDIIIEIARKKSSMLGFETPQKLPLKPKEEIQCAYYARLLVSDEKGVFSQISA
ILAKNDISLNNVLQKEIPHNSKAKILFSTHTTNEKSMLNALKELNQLQSVLDTPKMIRLE
N

>PROKKA_00971 UvrABC system protein C

MADLLSSLKNLPNSSGVYQYFDKNRQLLYIGKAKNLKKRIKSYFSIRNNEITPNHRTSLR
IQMMVKQIAFLETILVENEQDALILENSLIKQLKPKYNILLRDDKTYPIYMDFSTDFPI
PLITRKILKQPGVKYFGPFTSGAKDILDLSYELLPLVQKKNCIKDKKACMFYQIERCKAP
CEDKITKEEYSKIAKECLEMIENKDRLIKELELKMERLSNNLRFEEALYRDRIAKIQKI
APFTCMDLAKLYDLDFAFYGASNKAVLVKMFMRGGKIISSAFEKIHSLNGFDTDEAMKQ
AIINHYQSHLPLIPEQILLNACSNEALKEQLDFISHQYSKKIALSIPKKGDKLALIEIAM
KNAQEIFSQEKISNEDLILEEARSLFKLECMFYRVEIFDTSHHSSSQCVGGMVYENNAF
QKDSYRRYHLKGSDEYTQMSSELLTRRALDFAKEPPPNLWVIDGGRVQLNIALEILKSSGS
FVEVIAISKEKRDSKAYRAKGGAKDIIHTISNTFKLLPSDKRLQWVQKLRLDESHRYAINF
HRSTKLKNMQIALLKEKGIGEASVKKLLDYFGSFEAIEKASEQEKNAVLLKKRI

>PROKKA_00972 hypothetical protein

MEQNIFSLLIQKKSYYKLETLLKLKLKVFMPLSLQENLLFIFIKDSKLLFAFKDLWASK
EFNQRFKEISRFNLNTQGHAYGFDGLNGLEILGYVPKDSLEKANFYAPIKKQARFFRPSA
LGLFHNPIKDARLHECFEKARALIHYSRFFEE

>PROKKA_00973 Choline transport ATP-binding protein OpuBA

MKEIVTIENVSFNYHNRAIFKDFNLISIQEGDFLCVLGESGSGKSTLLGLILGLLKPSLGS
VKIFNETLSNNAFLRQKIGYIAQGNLSLPHLNALQNMFTCLNLQGINKQTAQKEAKALAL
KMGLDESLMDKFPNELSGGQAQRVGIIIRGIIHRPELILLDEPFSSALDSLNRKNLQDLIKE
IHQNSCATFIMVTHDESEAQKLATKTLEIKALKQEQ

>PROKKA_00974 Glycine betaine/carnitine/choline-binding protein OpuCC precursor

MPNMGVFKQLIKELYEWLLHSVDVITQHILIAMALKISVVKYLIKEFHDRFIYFIDLIAQH
FIIIVALLSSLLVLVFGVLIGVFVFNYSKARAFLLPVNFLYTIPSLALFALFIPVIGVGLK
NALLVLVLYGLLPVHSTYNALKEAREEVIAAIGLGCNPKELFFKVRFLAIPQILVGL
RIAVVMLVAMAGIGALIGAGGLGQAFRGLNTQNTTLLVAGSFIIALFSVLADQFVSVFQ
HENALQRLFSQNAQKQKRRVYNLAVFLFLLLASTLWLIPRSAIEEKPLVVATKPSSEQ
YILGEILSLLLEKHHPIKRAFGIGGGTMNIHPALIRGDFDLYVEYTGTAWVNTLKNPLT
QKVDVFETIKKRYEKEFNLLWVGLLGFNNTYSLSISKEDAQKYAIETFSDLAFHSPNFDG
AEFDFFEREDAFKGLIKAYRFHFRSLHEMDINLRYKSFESHKINALDVFTTDAQIKELDL
KVLKDDKGFFPNYQAGIVIRKEIVKKYPEVLEILEKLDSKISDEKMQLDNYQVEALKKSP
KIVAKDFLESGL

>PROKKA_00975 hypothetical protein
MNRMNKNYLLIFLLLASLIAREKDASSNLFDLIDKGINREQELKEQEQKTRLKLAQSPLV
ALEIVPQETPYLEWQGARESYLKVSAVVESVILKIDINQGRSCSLYPTPKSVSLVRNQ
SVAYEILCENQPLWIEVSTNLGKRTFQF

>PROKKA_00976 Motility protein B
MAKKNKPTECPAGEKWAVPYADFLSLLLALFIALYAISAVNKSKEALKTEFIKIFNYAP
KPEAMQPVVVIPPDGKEEQMASESSKPASQNTETKATIARKGEGSVLEQIDQGSVLKL
PSSLLFENATSDAINQDMMLYIERIAKIIQKLPKRVHINVRGFTDNTPLNKTRFKSHYEL
AANRAYRVMKVLIQYGVDPNQLSFSSYGSTNPIAPNDSLENRMKNNRVEIFFSTDANDLS
KIHSILDEEFNPHKQQE

>PROKKA_00977 Chemotaxis protein PomA
MDLSTILGLVLAVASISLGDILEDGNPLHIIHLSSVIIIVPTSLFAAMTGTHARYVKAAY
KEIKIVFLNPKINLNETIKNLVELATLARKDGVLSLEGRVAQIEDDFTRNGLSMIIDGKD
LKSVKESLEISIEEMEEYHGAHYWETAGETAPTMGVLGAVMGLMLALQKLDNPAEMAA
GIAGAFTATVTGIMCSYAIFGPFHGKHLAKSKDIIKEKTVLLEGILGIANGENPRDLLEN
LLNYIAPGEPKKSQFEG

>PROKKA_00978 putative adenylyltransferase/sulfurtransferase MoeZ
MLSRLEKERYLRHIMLEDVGEQGQLKLLKSSVLVIGAGGLGSAVLMYLCAAGIGKIGIVD
FDVVDVSNLQRQIIHSQDFLNQPKTSSAKARLKQLNAGIEVETFEERFKANNALSLIEPY
DFIIDATDNFNAKFLINDACVLAQKPYSHAGVLKYRGQSMVLPBSTCLACVFDKPPKKG
LNPISGLFGVLPGVLGCIQASECLKYFLGFETLLINTLLIADIKTMDFKKIQAPKNPECR
VCGTHKITHLQDYEI

>PROKKA_00979 putative metallo-hydrolase
MEILRREC GAVEENAYIVKLSSGIDFIIDPGFSSSEWVLENKPNKAILITHGHYDHVWD
SAQLSKALKDTPYAPKDDVFMLENDIFHLGMPVFSNFSVPCNKGCTTLEIENTTIKYW
HFPGHTPGCSIIEIEGVIFSGDFIFYSIGRYDFPYSNEKDMKESLLRFQNLDFPKDIEI
YPGHGDKTSFFAEREHSKIWVSMA

>PROKKA_00980 hypothetical protein
MTRIALLLLVKMSLNERKIVRSFGNYMQEWLYGKKGYRKALIGPKGDFYTSVSLSKFF
GGAVAFYIIKLLLEEKLFLPLKIVEIGSHHGHFLSDIASFLNALSVMGMEKCEFVSCEPL
KELQKLQRTIFKQATQLDLMICDLKDLDFKGHENAFVVSNELFDAFACEIIKDNKMLFIT
HDHKGWVGIDEPTKELLKNLNLKEGCTPLFLEAFIKDLLEKLNEASSWVFLSFDYGDEV
ERKDMHLRAFKNHQVLDKFDILNNLASLYQKSDLTYDVNFSVLVRFLEKHHAKFSFFKTQ
ANALLDMGLMGLLEMFSSKSVGYERYLKEAAKIKPLISPGGLGERFKALEFVKKNKM

>PROKKA_00981 hypothetical protein
MKQFKPIDGIKGGFLKIFLSCKSLLIQRSLEFYLSDCLSPMEVCDFVLSDDETLEINKPL
CFIEECLRKPFQKQSVREDVKNFYRALKESEKSCEECEEMKISKEQKIQLLEEYQTQKLC
QIISQ

>PROKKA_00982 Ribosomal RNA small subunit methyltransferase D
MPNHQPVKFKIIGGACKGLGLNLPNISSTRPTKAIVRESFFNTLQAEINGAHFIEVFSG
SASMGLEALSARGAKSAVFEQNKSAKYTLLENISLFGNRLKKEIEIQTFLLDAFKLLPTL
CLKNGVLNILYFDPFETSGFLGIYEKCFQALERLLKRSNPNKLLAVFEHESVHEMPKSL
ATLAIKQKKFGKTTLYFQ

>PROKKA_00983 flagellar basal body-associated protein FliL
MAEEQENTAQQPQKSKALLFVIIIGSVLVMLLVGVIIIMLLMGNKEESKENASKNTQEVQ
ANPMANKNQEAKEGSNIQQYLVLGPLYAIDAPFAVNLVSQNGRRYLKASISLELSNEKLL
NEVKVKDTAIKDTIIEILSSKSVEEVVTNKGKKNLKDDEIKSHLNSFLIDGFIKNVFFTD
FIQ

>PROKKA_00984 Holo-[acyl-carrier-protein] synthase
MIGIDIVSIAIEKCVKRFEMRFLERFLSPSEIVLCKDKSSSIAGFFALKEACSKALQVG
IGKELSFLDMRISKSPKNAPLITLSKEKMDYFNIQSLASISHDAGFAIAVVMISS

>PROKKA_00985 Fe(3+) dicitrate transport protein FecA precursor
MNGYLRVKTPYFLALVLTFWAFNSLVGAKDKHHFLKKVTTTEQKFSSSAPLSWQSEEVH
NSTSSRTVISNKKLTGNLNIENALQNVPGIQIRDATGTGVLPKISVRGFGGGNGHSN
TNMILVNGIPIYGAPYSNIELAIFPVTFFQSVDRIDVIKGGTSVQYGPNTFGGVNVITKE
IPKEWENQAAERITFWGRSSNGNFVDPKEKGKPLAQTLGNQMLFNAYGRTAGMLGKYIGV
SMQGNWINGQGFRQNSPTKVQNYLLDAIYKINATNTFKAYYQYYQYNSYHPGTLAQDYA
YNRFINERPNDQDGRAKRFIVYQNYFGDPDRKVGGDFKFTYFTHMSRDFGFSNQYQS

VYMSSQNKILPFGKGGEISATNPNCGLYSYSDTNSPCWQFYDNIRRFVVNAFEPKLNLA
VNTGKVKQTFNIGMRFLTEDLYRRSTTRKNPSVPNNGSGFDAGTSLNNFNNTAVYVSDEI
NFNNGMLTITPGLRYTFLNYEKKDAPPFKAGQTGKTTKERYNQWNPVNVGYKPIKDWL
FYFNYQKSYIPPQFSNIGSFVGTSTDYFQIFNVMEGGSRYFNNQVSFNANYFVIFANHY
TGRYGDNREPVNARSQGVLELYYTPIRGLNFHAAYTFIDAVITSHTMVTNPANPKGPKK
DIFGKKLPFVSPHQFILDASYTYAKTTIGLSSFFYSRAYSVDLNTVPFTEYAPTINKGAI
TTKTAGMTPYYWVWNLQISSTLWERKNQSVNASLQINNIFNMKYWFSGIGTSPNGKEAAP
PRSITAYVSYNF

>PROKKA_00986 hypothetical protein

MLDNIPITIQSKKIKTSLNITPSLEVILKMPDSCPQARANAFLKEQETWLKKTFLAMQ
EKHSLRLVRLNTYKDKILVFDEVKNANDYTLTDLKKILKTYLEQKLPLISQKMQTSYTHF
SIRNNAKVLGSCSYHNRLSFALLLVCAQKEAIDYVIIHELAHTIHKNHSKNFWRCVQIFC
PNYRALRERLKQNIIFYAQLLKTLP

>PROKKA_00987 Lipooligosaccharide biosynthesis protein lex-1
MRVFIIHLSPKTCHNFSKETHITPLLESCLKLQGISYEIFDAIYSKSSPTQLHPLILEHL
HPSFVIEDLLAFCENEKHPPCALKNFFHAIKHCGKRMGFGELGCYASHYSLWQKCIELNE
AICILEDIVLKERFKESLEFCRQHINELGYIRLMHLEENVAKQKTPVKGVSQLNFKDG
IGTQGYVLAPKAAQKLLKYSKEWVMPIDCVMDRHYWHGVKNYVLEEFACDGMNAQNS
NTEKQRPKKLPLSIRIGRFLHKSALQQWNILKSFFPNHIIK

>PROKKA_00988 Riboflavin biosynthesis protein RibBA

MILRRVTEALEAYKNGEMLIVMDDRENEGDLVLGIFSTPEKINFMATHARGLICVSL
TKDLANKFELPPMVSVNDSNHETAFTVSDAKEARTGISAFERHLTIELLCKDTPKPSDF
VRPGHIFPLIAKDGGVLARTGHTASVDLCKLAGLKPVSVICEIMKEDGSMARRGDKFLS
DFALKHNLKTLVSDILISYRLNESLLKMFCQEEREFLKHQTKCYTFLDHQQKNHYAFKF
KGAKTHDLAPLVRFHPIKEDDFDLTTDAFEVFFKALEYLKHEGGYLIIFMNTSEQNNIVK
DFGIGALVLKNLGKDFRLLSSCEDRQYKALSGFGLKLVTISL

>PROKKA_00989 hypothetical protein

MLISLKTFLKILLKIFLKTFFQKIWVVFVIIWGLGCSFLNANSVQLEETLRRNPKNLIWQH
FKKKFKKSNTIPYAPNSRWKYLGAISIGILGVSLVIGIVGLYLMPESTNWDKEKFGVKSW
FENVRMGPKLDNDSFIFNEILHPYFGAMYMQPRMAGFSWMAAFFSFITSTLFWYEGLE
AFVEVPSWQDLVITPLLGSILGEGFYQLTRYIQRNEGKLFGLSLFLGRLAIALMDPIGFVI
RDLGLGEALGIYNKHEIRSNLSPNGLNLTYKF

>PROKKA_00990 GTP cyclohydrolase-2

MKRLEVSÑQAKLPTQGFECIQCFREKGSNGSKDHLVIFTNFSQNPLVRLHSECLTGDA
LGSQKCDCCGALQMALERISKEGGLVIYLRQEGRGIGLFNKVNAYALQDKGYDTIQANEM
IGFKDDERDYSVAGEILEYYRIKKMRLLTNNPKKIAALEKYAEVTRESLIVCANEHNOGY
LEVKKLKMGHLL

>PROKKA_00991 This family protein

MVEVRFFGPIKEENFFIKANDLKLRAILQEKEGLKEWLGVCALNDHLIDNLDAPLKD
GDVISLLPPVCGG

>PROKKA_00992 Molybdopterin synthase catalytic subunit

MLKIIQGALDTEKLLKDYQEEARMKNFGAFCVFGVIREENNVOGLSFDIYEALLKTWFE
KWRHKAKDLGVVLKMAHSVGDVLIGQSSFLCVLMGKNRKNALYENFIEDFKHSAPIWK
YDLIHNKRIYAKERSHPLKGSGLLA

>PROKKA_00993 Molybdopterin adenyltransferase

MQTIHIGVLSASDRASKGIYEDLSGKAIQEVLSYLLNPLEFHYEIVADERDLIEKSLIK
MCDEYQCDLVVTTGGTGPALRDITPEATEKVCQKMLPGFGELMRMTSLKYVPTAILSRQS
AGIRNKSIIINLPGPKSIRECLEAVFPVPCVDLILGNMQVNEKNIQAFRPKQ

>PROKKA_00994 Cyclic pyranopterin monophosphate synthase accessory protein

MPLTHLNEENQPKMVDIGDKETTERIALASGRISMNKEAYDAIINHGVKKGLVLQTALIA
GIMGAKKTSILPMCHPIMLNGVIDIDILEEKETCSFKLYARVKTQAKTGVEALMSVSV
GLLTIYDMVKAIDKSMTISGVMLEHKS GGS GSDYNAKK

>PROKKA_00995 Neuraminyllactose-binding hemagglutinin precursor (NLBH)

MKANNHFKDFAWKKCLLGASVAALLVGCSPHIIETNEVALKLNYPASEKVQALDEKILL
LRPAFQYSDNIAKEYENKFKNQTAIKVEQILQNQGYKVINVDSSDKDDFSFAQKKEGYLA
VAMNGEIVLRPDPKRTIQKKSEPGLLFSTGLDKMEGVLI PAGFVKVTILEPMSGESLDSF
TMDLSELDIQEKFLKTHSSHSGGLVSTMVKGTDNSNDAIKSALNKIFANIMQEIDKKLT
QKNLESYQKDAKELKGKRRR

>PROKKA_00996 Helicobacter outer membrane protein
MNKLLKKGFLLAFRLVYLRADDLVITYTIKKEDLGFQRFLAKKCLRGAHSPCFTKPKKP
KRKLFNIDKSSHYGYTSVQMSWLQSRERFENHSKYRDIPFAEVSLIYGYKQFFPKKERY
GFRFYISLDYAYGFFLKNKGVLGDSLRESSQIPKSYREKLQRKETFINAIFYGAGADFLY
KRAFGLTILGMNFVGETWIFYETKIFKKWAKDPLSVYHPYMFQVMLNIGYRYRFSRYKNWA
IELGARIPFLTNDYFKTPLYTLHFKNISVYLTSTYDF

>PROKKA_00997 Trigger factor
MNLEVKKIDTANARLSAKPSIENLEKRYDKIAQKIAQVKIDGFRRGKVPLSLVKTRYQA
QIEQDAQEEMIQEVLNKAFKELGIENKDLIGSPNLTKFEKKDTHFEIEADIGLKPTIVLD
KIKECVPSVGVEIPNEEKIDERLKQLAKDYAKFVDVDAQRKAQNDDKLTIDFEGFINNAP
FEGGKAENFSLILGSKQMLEDFEKALLDMQAGEEKEFPLTFPSKYHAEHLAGEALFKVK
LHQIQAREVLEINDELAKIVLANEENATLKLKERVEGQLFLENKARLYNEELKEKLIEN
LDEKIVFDLPKTIIEQEMDLLFRNALYSMQAEEVKSLSQESQEKAKEKRESFRNDATKSVK
ITFIIDALAKEEKIGVHDNEVFQTLTYEAMMTGQNPESLIEQYRKNMMLAAVKMAMIEDR
VLAYLLDKNLPKEQQEILEKMRPNAQKTQVG

>PROKKA_00998 ATP-dependent Clp protease proteolytic subunit
MGYIPYVIENTDRGERSYDIYSRLKDRIVLLSGEINDSVASSIVAQLLFLEAEDPEKDI
GLYINSPGGVITSGLSIYDTMNFIRPDVSTICIGQAASMGAFLLSCGAKGKRFSPLPHSRI
MIHQPLGGAQGGASDIEIISNEILRLKGLMNSILAQNSGQSLEQIAKDTDRDFYMSAKEA
KEYGLIDKVLQKNVK

>PROKKA_00999 Peptide deformylase 1
MALLEIIHYPISKILRTISKEVVSFDAKLHQQLDDMHETMIASEGIGLAAIQVGLPLRMLI
INLPREDGVQHKEDCLEIINPKWIETGGTIMYKEGCLSVPGFYEEVERFEKVKIEYQNR
AEVKVLEASELLAVAIQHEIDHLNGVLFVDKLSILKRKKFEKELKELQKKQKHK

>PROKKA_01000 Competence protein ComM
MINTIFCATMQRGVAEIVAVEATFTRALPAFVISGLANSSIQEARQRVQSALQNNDFTFP
PLKITINLSPSDDLKSGSHFDLPALIALQKQELAFKEWFAGFELGLDGKIKPNPNIFP
MLLDIAIKHPHAKVIAPKANEELFSLIPNLQCFVVEHFKEALEILQNPFIKADTHTKLP
FKTIELNDKEYYFSDAYALDFKEVKQGAFAKEAALIASAGFHNLIILEGSPGCGKSMIINR
MRYILPPLSLNEILEATKLRLISEQDSAYYPLRSFRNPHQSASKSSILGSSSLKEPKPGE
IALAHNGMLFFDELPHFKKDILEALREPLENNKLVSRSRVHSEKIEYETSFLFVGAQNPCLC
GNLLSATKACRCQDREITQYKNRLSEPFLLDRIDLFFVQMEEGNYKDTPSHSWTSKEMHQV
LLAFKQKQLRKQSAFNGKLNNEEQIERFCPLNAEAQKLLLEQAVERNLSMRSVNKVKKVAR
TIADLNACENIEKSHMLKALSFRKIS

>PROKKA_01001 Cadmium, zinc and cobalt-transporting ATPase
MQEYHIHNLDCPDCASKLERDLNKLVDYVKKQAQINFSTSKLFLDTSDFEKVKAFIKQNEPH
LSLSFKEVAEKPLSFTPLIVTIIIVFLGAILILHLNPSPFIEKAMFFVLALVYLVSGKDVI
LGAFRGLRKGQFFDENALMLIATIAAFCVEAYEESVSIMVFYSAGEFLQKLAISRKKS
KALVDVAPNLAYLKKGDALVSVAPEDLRVNDIVVVKVGEKVPVDGVVVKGESLLDERALS
GESMPVNVSENSKVLGGSNLKAVLEIKVEKMYRDSSIAKVVDLVQQATNEKSETEKFIT
KFSRYYTSPSVLFIALMIAVLPPLFSMGSFDEWIYRGLVALMVSCPCALVISVPLGYFGGV
GAASRGILMKGVHLEVLTQAKSIAFDKTGTLTKGVFKVIDIVPQNGHSKEEVLHYASC
SOLLSTHPIALSIQKACEEMLKDDKHQHDIKNYEEVSGMGVKAQCHADLIAGNEKMLDQ
FNIAHSPSPENGITIVHVAFNQTYVGYIVISDEIKDDAIECLRDLKAQGIENFCILSGDRK
SATESIAQTLGCEYHASLLPEEKTSVFKTFKERYKAPAFVVDGINDAPTASADVGIGM
GKGSELSKQSADIVITNDSLSSLVKVLAIAKKTKSIIWQNILFALGIKAVFIVLGLMGVA
SLWEAVFGDVGVTLALANSRAMRA

>PROKKA_01002 Putative type-1 restriction enzyme specificity protein MPN_089
MHKIERLLQTLAPKGVEFKTLEEVFEIKNGYTSPSKNNLEFWKNGTIPWFRMEDLRENGRI
LKDSIQHITPKALKGKKLFPKNSIIISTATIGEHALIIVDSLQNRFTFLSKKANCNLA
LDMKFFFYQCFLLGWCKNNINVSFGFASVDMTAFKKYKFPPIPPLEVQQEIVKILDAFTL
NTELKARKKQYEEYQNMLLDFKDIKQNHKDAKMSTKPYPKRLKTLQLTLPKGVFRKLK
EVLEYDQPNKYCVMSKEFDKSYPTPVLTAGKTFILGYTNEKDNIYQASKSSPVIIIFDDFT
TATQWVDFPFVKSSAMKILLPKNPTINIRFIFFYMQTIPYNISGEHTRQWISRYSQITI
PIPPLEIQQEIVKILDQFLALTDDLGAIPAEIEARKKQYEEYREKLLTFKPLTPHKEVK
KC

>PROKKA_01003 Putative outer membrane protein
MNYKVASAKNIATLLFLFSSQSEALDLGKIAKIKAGAESFSKVGFNKPINTNKGLYPTE

TFMTIMAYMQVDFTELLPKSATANGHHLDGSLGGWGGAVIYDSTKDFINEVTGKTYGAMA
WNYVGYWGGLVGQKPWASCGLATGNLTQGQYDKMTQAEMTQLSNQEALEASTCAKGYADH
TRNYVIYNAYLRNYKDI FEIRGGRYESPADYMSGYNQGLDMTLNLGNFKFWWFSSFGRG
FAYNEWLYNFYSPKTYTLKNGQ TINPGVHAFYIIWNYKGWSIQPFVYFSPFNEYDPNFTI
TYDSNPTFTGLGFRSQTDVTVLNPFFYAKRFWDTYQFGMPAGKNAHSLMIKQKFEWNEYNF
GFGIYKSFGNANWMIGYHGNRLGFDFTNTVYANTLNSLSYMMDANAFTVFAFGGGVHRK
FLWGLLGRLTYGPRANEQVLSLNLGYKFTKNFSADIKFEYYNVLMHQGYKMGWNGPKLDS
QPATDQDRSHIFTEIVWKL

>PROKKA_01004 Lipoprotein-releasing system transmembrane protein Lole
MPNRSLIFFLIKRYLRFDKSQPFISITALLAFFGVAVGVMVLIVAMAIMNGMSKEFEKKL
FVMNYPLTLYTISPYGISEEVVDLEKKFPNLLFSPYLQTQSLIKSAHSMNGGVVFGVDF
SKERRINGVLNDALKNTNENDLSKNPFNLIVGKSLRYSNLNLDLNQKADLFFTELEPTGLT
LSPIMKRFTIKGDFDSGLKSYDMSYMYASLQAISAIRRLPLGLYDGVHVYSKTPMKDIEI
LRNALKTINHGGIGIEGWQQNGNFFSAMELEKRALFIVLMLIILMASLNI ISSLLMVVM
NRRKEIALLFSGSSQKEIQKTFYLGNIIGLGGVILGVVLAFLSMYLLSVFPIISLPAD
VYGINTLPLDLSLMDFTLTLLIGSVIIIVALSSYPSKKASIIDALSVLRNE

>PROKKA_01005 preprotein translocase subunit SecA
MIKAIIGKIIIGTRNDRWIKQYKKQVLTINALEPTYEKMSDVELQNAFEELKKRVRSTEKD
LQEKTLLEVLPESEFAITREASKRILKMRHFDVQLIGGMVLNDGKIAEMKTGEGKTLVATL
AVALNALKGESVYVVTVNDYLAHRDSKEMEPLYHFLGYSVGTITASVRDDDERLEIYSKD
IVYGTNNEFGFDYLRDNMKYSLEHKVQKSHAFAIVDEVDSILIDEARTPLIISGPVDRRM
ENYNKADEVAKSMQVETDFTIDEKNRAILITEEGIKKAENLFGVDNLYKIENAALSHHL
QALKANYLFFIDKDYIVANNEVVIVDEFTGRLSEGRRFSEGLHQALEAKESVSIKEESQT
LADITFQNYFRMFSLSGMTGTAQTEATEFLEIYNLEVVS IPTNLAIKRKDLNDLIYKSE
KEKFDVAILKIKELHDKGQPVLVGTASIEKSETLHALLKKERIPHTVLNAKQHTKEAEII
KDAGLKGAVTIATNMAGRGVDIKLTDEIKELGGLYIIIGTERHESRRIDNQLRGRSGRQGD
PGTSQFYLSLEDNLLRIFGSDRIKGVMEKLGLKDGHEHIESKLVTRAVENAQKKVENLHFE
SRKHLLEYDDVANEQRKSVYKFRDELLDANYDISAKIAENREYALNQIFSKLKAFDHQNL
SEEELLGLKNILKEDFNHVSLEDLKKA APIENFVAEKLKSDYENKMKVLDSEQRSRIER
IVYLQILDNAWREHLYTMDNLKTGINLRGYNQKDPLVEYKKESYNLFLEFIEDIKMEAIAIK
TFSKIQFENEXDSSDAERYLDNFSEEREHESVTYRHEETLDEDLNVAMKAFAKTPKRNEP
CPCQSGKKYKDCCAKSGPKKGLFAK

>PROKKA_01006 Outer-membrane lipoprotein carrier protein precursor
MKAFLKILMVLI FIFISVACAKNPPTLSKEEEVLQNLQSFSAHFKQVLKNEKPLVYYGVLKA
KAPNWALWVYEKPLKKEIYMNDKEVVIYEPNLFQATITPLKDKTDFFTILKRLKKQNDGS
FKTTINKTTYRLVFKDGKPFSLFEFKDEMNNLVTTITFSQAEINPTIANEIFVFKPKDENID
IVRQ

>PROKKA_01008 hypothetical protein
MQKFFSRFRRWALPFYFVSALAAIDIDEVTEAQANSIKLSDQLVSLSDKLLEKAVDRGRN
TDHLKDLSDLHEKIKHLRLILEPKPKDKENPNPLKDHQGSETIEIGEAVKKALGEPVFPQ
DALDAALQIDKQLDSFKQDNLIDVKPLKDVLSQLEAELRKAINFQKQWLNSTKPN

>PROKKA_01009 Putative outer membrane protein
MLKFQKLPLL FVSILYNQSPLLAFDYKFSGVAESFSKVGFNHLSKLSKEGIFPTATFVTA
TIKLQVDSNLLPKNIEKHSKIGVGGILGALAYDSTKTLDQATHQVYGSELFFFIGRW
GYLGNAPWKDSRIESDAHTRNYVLYNSYLFYSYGDKFHLKLGRYLSNMDFMMSGYTQGFEL
DYKINSKIALKWFSFGRALAAGQWIRDWYSPIVTEDGRKDVYGIHAVQLYFSNKHVKA
TPFFYFSPKTYEAPGIKIHIDSNPKFRGLGLRSQT LINVIFPVYAKDLYDVYWRNSKIGE
WGASLLIHQRFDYNEFNFGFGYYQNFNGANARIGWYGNPIPFDIRSNSIYGLVFSNAITA
DSVSGYVFGGGVYRGFLWGILGRYTYATRASERSINLHLGYKWNSFFSIDMNLQYYAVSM
HTGYKVNDLTSPFNKAFKANTQDRSNLMVSVKFFF

>PROKKA_01010 hypothetical protein
MPYALRKRFFKRFALIVSTFCAISLNAKSYLFSPLPPAHQQIIKTEPCSLECLKDLMLQN
QIFSFISQYDNNNQDES LKIYYKDILNKLNPVFIASQTPTKESYEPKIELAILLPKKVVG
RYAISVMNTLLAYLNTRNNDFN IQVFDSD EESPEKLEQTYKEIEKEKFPFVIALLTKEGV
ENLLQNTTISTPTYVPTVNRAQLENQTERSLSERLYFGGIDYKEQLGMLTAFISPNSPVI
EYDDDGLIGERLRQITESLSIEVKHQENISYKQATSFSKNFRKNDAFFKNSILILNTPTT
KSGLILSQIGLLEYKPFKILSTQINFNSSLLLLTQPKDRKNLFIVNALQNNDETLEIYAS
LLESDLRHDWVNYSSAIGLEMFNLTLDPHFKKSFQESLEDNQVRYHNQIYQALGYSFEP

KNESGTTKE

>PROKKA_01011 hypothetical protein

MKKIVVSWCVALAFLRADPAQANKAISNADLIKEIRDLKKIINAQNTTEINQLRKVQEVLS
GQLGDMRKDILSTRDYCISLRPYIYNWR

>PROKKA_01012 Aconitate hydratase 2

MVAKVSYNLTSSKKS FVIMGQMIRALANNAIKKGEMMKDFLEDYKKS VLERESENIPPLP
LSAKQVQAVVEILT KDPTNAAF AKELLIHRVSPGVDEGAKVKA EFLAKLSQKKLE CAYIS
ALEATTLLGTMLGGYNVEPLIMGLESQDKNIAKESAKALKTTLLVYGSFDKIAAMSKTNA
LAKEVLESWANA EWFLNKEPLNECIEACVFKIDGETNTDDLSPASDAFTRSDIPLHAKAM
LKNRIENYEQR IEA IKTKGVPVAYVGDVVG TGSSRKSATNSIMWHFGKDI PFVFNKRSGG
IVIGGVIAPIFFATCEDSGALPIVADV KDLKEGDLIKIYPYKGEITLNDKV VSTFRLEPE
TLLDEVASGRIP LIIGRGLTNKARKFLGLGESEAFKKPSAPKSDAKGYTLAQKIVGHAC
GVKGILPGTYCEPKVTTVG SQDTTGAMTRDEVKELASLKFDAPFVLQSFCHTAAYPKPSD
VSLHATLPSFITQRGGVALHPGDGVIHTWLNRMGLPDTLGTGGDSHTRFPLGISFPAGSG
LVAAFAVTGTMP LNMPESVLVRFKGEMNPGITLRDLVNAIPYYA IKKGLLTVEKKGKINV
FNGRILEIEGLPD IKMEQAFELSDASAERSAAACVRLNKEPMVEYLK SNIKLIDEMIAS
GYEDKETLKKRKDAMQAWVDNPVLL EPDSNAQYAAVIEIDVAEITEPILACPNDPDDVAT
LSEVLADTTGKRPHAIDEVFIGSMTNIGHFRAFGEIVKNAPPSQARLWVPPSKMDEQE
LINEGYAIFGAAGARTEVPGCSLCMGNQARVRDNAVVFSTSTRNF DNRMGRGAKVYLG
AELGAACALLGRIPTKEEYMN LVSEKLESQKDKIYRYMNFNL MENFRL

>PROKKA_01013 Chorismate dehydratase

MRFGKIDYLNMLPFDFVFIKSYPTPCYFKQFLRLKKTPSKLNQSF LFRRIDAGFISSIAG
YSFALHSHSLGIVAYKEVLSVLVVGTKNAFDKESASSNALSQALGLKGEVLIGNKALQFY
YSNPKKDFIDLAALWYEKKRLPFVFGRLCYHKNKDFYKRLSLAFKHQKTKIPYYILKEAA
LKTNLKRQDILHYLQKIYYTLGKKEQSGLKAFYRELLFKRIQKPKRF

>PROKKA_01014 Uridylate kinase

MQAKIKNKRVLVKFSGEALAGDNQFGIDIHVLDHIAKEIKSLVENDIEVGIVIGGGNIIR
GVSAAQGGIIRRTSGDYMGLATVINAVAMQEALEHIGLDTRVQSAIEIKEICESYIYRK
AIRHLEKGRVVI FGAGTGNPFFTTDTAATLRAIEIGSDLI IKATKVDGIYDKDPNKFDA
KKLDTLSYNDALIGDIEVMDDTAISLAKDNKLP IVVCNMFKKGNLLQVIKHQQGVFSMVK

>PROKKA_01015 DNA-directed RNA polymerase subunit omega

MKKERTESLVAQALKNIGNDRYMLDNLV FARVKQLNAGAKTLVNMDPKRHKLV DIAIREI
AEGKIDIDRIDERN

>PROKKA_01016 Bifunctional (p)ppGpp synthase/hydrolase RelA

MNEIDKSVDIGFLRILDVIKKVTT PKGGIEILRTLIDFTP KIENALNLA AKSHKGQYRKS
GEPYIVHPICVASLVAF CGGDEAMVCAALLHDVVEDTPCKIETIEREFGRDVANLVDALT
KITEIRKEELGVSSQDPRMVVSALTFRKILISAIQDPRALVVKISDRLHNMLTLDALPHD
KQVRISKETLAVYAPIASRLGMSSIKNELEDKSFYIYPEEYKNIKEYLHKNKQSLLLKL
NAFASKLEKKLFDSGFSHSDFKLVTRVKRPYSIYLMQ RKGAVNIDEILDLLAIRILLKN
PIDCYKVLGIIHLNFKPIVSRFDYIALPKENGYKTIHTTIFDESSVYEVQIRTFDMHMG
AEYGNSAHWKYKAGGVNEDHHEGMRWLQNFKYHDS DLKNDPKEFYELAKNDLYREDIVV
FSPHGDYTLTPVGAIALDFAYMVHSDLGDKATEAYINSKKALLNQELRSGDVVKIVKGDK
IIPRFIWMQDKTSKAKNHLRIQRNRNLKEIDTKSMINILATFFGRSVFEDVGLKDYKNF
EERLTD CGVETTLTEAMKSFENLAKLTEE IENKVFSLKEDAILEYQEMSLWTRGLRYLGF
KTNVLNFLT PNRQWQCKELEHFSVCSSNALEIKQVLLNDCCYPKYGDEIIAIVTDLKDPK
AVVHHKFCKKAIAEVDAPAMVYIEWHKRDR TIYKMMFYLGEKKS VLAGLLTFLNRNECN
IVGVSYLGYKDKYSSHCEVSFEIATDKADWIRALINRKYQDRIVELSSLD DAYES

>PROKKA_01017 Tyrosine--tRNA ligase

MEQKIAI ALKEIKRGANEIIGLEYIEKLVRKYETNERFIVKAGFDPTAPDLHLGHTVLI
QKLALLQQYGARVKFLIGDFTAMIGDPTGKNETRKPLNREQVLEN AKTYEEQIYKILDEK
HTEVCFNSTWLDALGAKMIELCAKFSVARMLEDDFAKRHKENRPI SIVEFLYPLLQGY
DSVAMDADIELGGNDQKFNLLVGRFLQRAYGLNKEQSVITMPLLEGLDGVQKMSKSLGNY
VGITEEPNAMFGKIMSVSDDL MWRYYTLLSTKTLEEIEDLKHGILNQTLHPKAVKEKLAG
EIVARYYDNDQAIKAKEQFSKVFSANLLPEILSESDFDEGVGILDVLKQIGFCPSTSQAR
RDIQGGGVKINQEVIKDES YRFVKGNYVIQLGKKRFMKLNIN

>PROKKA_01018 Dihydroorotate dehydrogenase

MVSTLKLPLKIGKHTIKFPIFQGGMGVGISWDELAGNVAKEGALGVISAVGTGYK NMRFV
ERIVAKKPF EALNFYSKKALNEIFANARKICGNKPLGANILYAINDYGRVLRDSCEAGAN

IIITGAGLPTNMPEFAKDFSDVALIPIIISAKALKILCKRWSDRYKRIPDAFIVEGPLSG
GHQGFYEDCFKEEFRLNLVPKVVEASKEWGNIPIIAAGGIWDRKDIDTMSLGASGVQ
MATRFLGTKECDAYADLLPTLKKEDILLIKSPVGYPARAINTGVIKRIEENAPKIA
VSNCVAPCNRGEEAKKVGCIADGLGRSYLGNREEGLYFTGANGYRVDKIISVHEL
TEG

>PROKKA_01019 N-acetylmuramoyl-L-alanine amidase AmiC precursor

MLVRLGVVACLFWLHFACATTLKITNIVPFGSSSVKILFNQEVKKFKEVPLKNFKSYLEL
EAVLTIPKKHYQFSKQSSITIAQFSPKLARVVISYAPKMTYEIKIFKDKLYVSIVEKKPL
VRHQMAPKPLKHHALKHHTPKPAPKSIKKEAKEKTPIKHAHSHKHAHSPLNERSAKKEILK
KEAENESKNQVFAEKNDAFIKTRKKHKKIVLDAGHGGKDCGAMSANLVCEKDIVLEV
KFLHKLKRGYSVLLTRDKDIYIDLVRTELNRKSADLFISVHANSIPKRSTSNAGI
ETYFLSTARSERARKVAEQENKDDVNLMDYFSKSLLNSLNTQRLIVSNKLADVQYGM
QSIRKNYPDVVDGGVREGPFWVLGALMPSILIEIGYNASHAIESKRIQSKPYQKILAKGI
ADGIDSFFSKND

>PROKKA_01020 hypothetical protein

MNFYQKIYTHKVVFSLLFLLFNFVETLLFSHFSDDFSQLEFFLFENHVYDFIIKLDYLG
LIGVSLIYLLVLILKPFTLTRQKCACVGILCLSFYAWNFPKHSILVLYLFYFALLATLL
WRFLGASMKQSFLPSMNICIVWVFASLSQSFRLSVSDCVDFSLFTLALILLILVLIYHK
RLFGLYEYANTLILIVGLSVVVLCSMFIIQTKEYYGMRLGFYFLGLLGWLLEYIHNTLR
LEHQI

>PROKKA_01021 Flagellar biosynthetic protein FlhB

MAEEKTELPsAKKIQKAREEGNVPKSMEEVGVGLGLLAGLMSIFIFFIWWVDGFSEMYRH
VLKDFSLDFSKEVQELFNQLAKDTFLLLLPILIIILMVVAFLFNVLQFGWLFAPKVI
EPKFSKINPINGVKNLFSKLLDGLSLITLKVFLAFFLGFFIFSLFLGELNHAALLNLQ
GQLLWFKSKALLLISSLLFLLFVLAFLIDLMIKRRQYTNSLKMTKQEVKDEYKQQEG
NPEIKAKIRQMMVKNATNKMMEIPKANVVVTNPTHYAVALKFDEEHPVPVVVAKGT
DYLAIRIKGIAREHDIEIIENKTLARELYRDVKLNATIPPEELFEAVAIVFAQVAKLE
QERQKQKIIKPL

>PROKKA_01022 putative molybdenum cofactor guanylyltransferase

MKNPIIDNIPCVLLAGGKSSRFIINNIQINKALMPLKSYSSLLLEYQYSRLLKLFK
KVIIS TKKSYELNAPYLLEKESDLFSPLFGIYN AFLTLQTPYIFFIPIDTPLVSF
ESIKALCGIK NFSVIYAKSPTKEHYLISLWHQNSLNALCYALKTHNYRLSDLVKNT
SSVAIHFNQEEEFLLNLNTLKDYE LAVQILKKRANG

>PROKKA_01023 Cyclic pyranopterin monophosphate synthase

MLVDSFNRVIDYIRVSVTKQCNFRQYCMPTPLDFFDDEELLPLDNVLEFLKIAID
EGVKKIRITGGEPLLRKGLDEFIAKLHAYNKEVALVLSTNGFLLKKMAKDLKDAGL
SRVNVSLDSLKSDRVLKISQKDALKNALEGIEESLKVGLKLNLTVMKSVNDDEILEL
LEYAKNKH IQIRYIEFMENHAKSLVKGLKEEILDIAQYQIIAAEKPKQGSSKIYTL
ENGYQFGI IAPHSDGFCQSCNRIRLASDGKICPCLYYQDAIDAKEAIMNKDTKNI
KRLKQSIINKPEKNMWNDKNSETPTAFYYTGG

>PROKKA_01024 hypothetical protein

MKKDREKKQKYSNITDATIMGSTSEENVLYASANREHLSVLDRLLEEISKRKINPN
YIKQINQQAGYSAEIKEQAHVNAHNILAGKRGRVRQYDDLSEKKAQVKKLFPNYAT
PSKNHELVDYVSVDEKGDVIPGTAVQSKFVGKNGEECFKKLLSKDYKKYFENGAKM
KIARNHYGDFQRAVNTRIKSLESQIAKQKGLGNFQKAAHLEEKLQCKTIKAHTRPA
RVAKGEAIEARLNP NLSTAKDVTRVSHQAGMNAQTGALIGGGVSLTNVYECIANGK
DPIKAFKHTTIATLKG GALSYGSTFASSSLGGWMQSSANKVIOQLGKGSVPAMIVSV
VATNATILGRYFSGKIDEA ELLKQLGKANTTLVSSGAMAFAGQALIPVVGALIGGF
VGAVLSETCLNSLLKAREEAK LARQRRIEIEKECREIIKLLKAYQNRKFVEFEQYF
HETTKFFNQSFDELWRASYTGADLAIGVNNKSREWLGQKALFDNSKEGWELITSNKN
IRM

>PROKKA_01025 hypothetical protein

MEKNDRYAIVTNVIEPLEKGGSFNQDRDREKFAQAARTHGIEDSVIEEVIDITQT
ISLIHLHEDRLDASDLPREQKKAMRAELQKSIDENLEALKSIRNSILVRDRLPMWS
LLNENWLQEI LKNKNNI

>PROKKA_01026 hypothetical protein

MPLPFILGGLALAAAGYGFKKIDALDADCEADEFIKKAENLKEEATKRLESAQSD
CKLAFARFGEKKLRVLSHSPNFLDHFHRLKRSTITIKGINMQDIQRQVSDARKIT
NQLNANGIDGDSVLGVFAGFGLSGISSFTTGAVLGGLVASGLAGMAVLGGLVAGP
ALAILGAISADE MEKKRDDAKAYLSQVEAAVKKADVIVDNLQAVRKMVDLFTTRQ
ITKIDALFFSLSQDAIAT

MKKHHYDTSRYNQKEKDQLCVTVSTLFSLSAFLKVSIMDEHQKLNKKVQKALNLMRDQVN
AIESAQKSGHYDVAMIQSRQKDFENL
>PROKKA_01027 Flagellin N-methylase
MKGSQEKPRPCTSCGLCKNITGIVELIGFDAGNGVCKFLDLETNLCKIYESRPLICR
VDEAHKKLYPHIPLKEFYAKNAEVCNALQEANHMDKSFRVILKK
>PROKKA_01028 Signal recognition particle receptor FtsY
MFNFFKKIVNLIKGEETKEETKEKKRQSVSKEELEELIGFDIQYDLIESLLKHLGDLVT
PKQLEVALLRFRVRESYDKTRLKTITTKPLVHLIVGVNGAGKTTTIAKLAKLSLKQHKK
TLLGAGDTFRAAAVKQLQLWGEKLNQVISAKEGSDPSSLAYNTIESAIAKNIDEVFIDT
AGRLHNQTNLKNELSKIARTCSKVLKDAPFYKFLILDGTQGSGLTQAKIFHETLALDGV
IMTKLDGTSKGGAILSVLYELKLPILYLGMGEKEDDLIAFDEERFIADLVDVAVFEQ
>PROKKA_01029 hypothetical protein
MRIKAYFLRFIALVLIVLLGFSACKNSQKPQDSQNNTTQQDSPKTYTAMDINNQEYTIMG
DLDSLNISSPDNTPTLVLVSALDNSLKDYAPSFNLIKKTFKDRLRVLILLNKPYSSEVK
DFNTHSQADFMILNPKDTALFDHLKYDSLHDSFNMLLYHKHQLIKMYQGIVPIEMLQFDI
SHLKD
>PROKKA_01030 Ribonuclease Y
MSSGLIYISLEVLVACLITALVMYYVMKKIYYARGQATLKSASAKALMEFQAKSFVEAE
EMRMKSQECKLQQQYENKNLQLQTHFEKKEVHLKHLEAQHKEFVRDEKRYLEKEKKELEK
ERQILEQEKENFKKQRVVCKEAQAKALDAMLNMYATKDEIKSMILEQLEEELEAQKSAL
IRRYEKEAKEEGKKSYAILAEATARFAGNYATENLTTRIALPCSDYIGRVIGKDGKNIE
AFKKVSGVDIEFSEDSSSELCLSSFNLYRRKVASETLKILIEDGRIQPNRIEEVYHRVARN
LEKELLSEGESVVLELELGAMEDELKILIGKMYRSSFQGNALQHSKEVALLAGLIAEQ
GGDKKLARRAGILHDIGKALTQELGRDHVNLGVEVCKRHKEDPVVINAIYAHHGHEEILS
VECASVCAADVLSAGRPGARIKSDEEYAKRMQALEEIALEFDGVEKAYAMESGRELRVIV
KSNQVRDNQVPIIARKIAKKIEESAQYVGEVGVQVVRESRFTTATLKQ
>PROKKA_01031 Multidrug export protein MepA
MLVVALSVFFVGKISHHHIVALGVGLQFLMLFYGINLTYGTNAILSRVLVGARDFTQIN
HAFSSIFIGAFVICLGVLLVSYFLIEPFLNWMQLQNPSRQLTQDYLEVLVIALPSVFLKN
VLVSALASFSDTLTPFIVKIIMVIACIFLNQALIFGDFGFKEMGIVGSALANVVVSYLEL
LALGVWIIQIKKIPLKFKITFHFSFLKTMFRVWGPAGFERLLSLFSLILLSKFVASYGDKV
LAGMQIGIRVETFSFMPGFGFMIAAMVLTGQNLGANKPKIATEYAHILKISMGLMGVLG
VILVLFakeFASLFSQDEEVLEVARSYLIAVGLSQAPLIGYFVLDGVFRGAGISKVSLYI
NTLSLWGLRIMPIYLLLIYHFKEVIFVVIASETFLRSFIYYKVFSKGIWKRCGKA
>PROKKA_01032 Na⁺/H⁺ antiporter family protein
MLENSSIWSNPAFVAIICMCVLSLLRLNVMLSMISATLIAGLMGGLGLTESFNAMIDGMK
GNLNIALSYILLGALAVAIKSNLIKVALSKLIGLMNYKRSTFCFLIAFIACFSQNLVPV
HIAFIPILIPLLHLMNRLELDRRAVACALTFLQAPYLVLVPVGFGLIFQTTILEQLKAN
GVSTTIAQIAGVMWIAGLAMVVGFLAVLTLYKKPRHYKEKSFDIENYASLQLNYHDYLT
FIGIIVAFVIQLATDSMPLAFLALAIILLGRGIKFKETDSLMDDSVKMAFIAFVMLVA
SGFGEVLQKVHAIEGLVNAITSVVQGKLLGAFLMLVGLFITMGIGTSFGTIPIIAVYFV
PLCAKLGFESIESTILLIGIAAALGDAGSPASDSTMGPCTGLNADNQHNHIYDTCVPTFLV
YNLPLIVFGVLGALLG
>PROKKA_01033 N-carbamoyl-D-amino acid hydrolase
MIYASVLQHAYCGSRKKTIEHTANLLEQALKKHPKTNLVVLQELNPYGYFCQSENPKFFD
LGEYFEEDKAFFSALAQKFQVVLVASLFEKRAKGLYHNSAVVFEKDGSIAGVYRKMHIPD
DPGFYEKFYFTPGDLGFEPITSVGKLGLMVCWDQWYPEAARIMALKGAIEILIYPSAIGF
LEEDSNEEKKRQONAWETIQRGHAIANGLPLIATNRVGVELDPGAIKGGITFFGSSFVV
GALGEFLAKASDKEEILYAEIDLERTEEVRRMWPFLRDRRIDFYNDLLKRYI
>PROKKA_01034 hypothetical protein
MKDYEDELEDFFFFEELEGFEEDEEYGDYKNVYDDDDYEDYNSDYEEE
>PROKKA_01035 tRNA threonylcarbamoyladenosine dehydratase
MSEPIDRFTRIRWLFKNDFEKIRRQRVLICVGGVGGFALDALYRVGIGQITIIDKDVFD
VTNQNRQIGSERIGESKVLVLQDLYKGIQALNLHINEAFLNSFNFRDYDYILDCMDDLPI
KTSIAIKCQNFAYGKFISSMGSAKRLNPKHIQVGSVWESYGDKFGRKFRDFLKKRRFKGD
FKVVFSPVPHCIELGSFNAVTFGLQIASEVVQDIK
>PROKKA_01036 hypothetical protein
MNSPNALLDSEFKIALVKKDSKQAFSLIERLSLEQIKNLDLNTLLSLKEMIAQSIELLEKE

KEELQSQMHKAKKIQKFLS

>PROKKA_01037 Flagellar protein Flis

MQYANAYQAYQHNRVSVESPAKLIEMLYEGILRFSSQAKRCIENEDIEKKIYYINRVTDI
FTELLNILDYEKGGEVAVYLTGLYTHQIKVLTQANVENDASKIDLVLNVARGLLEAWREI
HSDELT

>PROKKA_01038 Flagellar hook-associated protein 2

MAIGSLSSGLGSKVLNYDVIDKLKDADEKALIAPLDKKMEQNVEKQKALVEIKTLLSAL
KGPVKTLSDYSTYISRKSNVTGDALSASVGAGVPIQDIKVDVQNLAQGDINELGAKFSSR
DDIFSQVDTTLKFYTQNKDYAVDIKAGMTLGDVAQSITDATNGEVMGIVMKTGGNDPYQL
MVNTKNTGEDNRIYFGSHLQSTLTNKNALSLGVDGSGKSEVSLNLKGADGNMHEVPIMLE
LPESVSIKQKNTAIQKAMEQALENDPNFKDLIANGDISIDTLHGGSLLIINDRRGGNIEV
KGSKAKELGFLQTTTQESDLLKSSRTIKEGKLEGVISLNGQKLDLKALTKEGNTSEENTD
AIIQAINAKEGLSAFKNAEGKLVINSKTGMLTIKGEDVLGKNSLKDGLSAGMVQSYEAS
QDTLFMSKNLQKASDSEFTYNGVSITRPTNEVNDVISGVNITLEQTTEPNKPAIISVSRD
NQAIIDSLKEFVKAYNELIPKLEDDTRYDADTKIAGIFNGVGDIRAIRSSLNNVFSYSVH
TDNGVESLMKYGLSLDDKGVMSLDEAKLSSVLNSNPKATQDFFYGSDSKDMGGREIHQEG
IFSKFNQVIANLIDGGNAKLKIYEDSLDRDAKSLTKDKENAQELLKTRYNIMAERFAAYD
SQISKANQKFNSVQMMIDQAAAKN

>PROKKA_01039 flagellar protein FlaG

MVNEVQGIGIPTSVQTTPTKEISRTNTTNTIDESKTTIDPEQYKPKLELLSERLNEE
MKRIGTDINFSYNDTIKGLVSVKDANGDKVLREIPSKEAVELMQRMRDVIGIIFDKKG

>PROKKA_01040 Peptidase family M23

MYKLGVFLLATLLSANTQKVSDIAKDIQHKETLLKKTHEEKNQLNSRLSSLGEAIRSKEL
QKAEMERQMVALKKSLEKNRNESLAQEKVLTNYRKSLDHLQKQRSFLQKRVFDTLLEDFL
FSQALKGQNLASSNDVILQVAFENLHQSTLSKMSQLSQEEKDLNTQALKVKSSIQKISSV
IDEQKTREVTLSLQIEQNKLILSMQKDYAIYNQRLTLEKERQNLNALLKRLNIIKQNR
ENEERSVLKKSSQALEVKQVASSYQNINTTSYNGPKTIAPLNDYEVVQKFGPYIDPVYNL
KIFSESITLVSKTPNALVRNVLDGKIVFAKEINMLKKVVIIEHKNIGIRTIYSQLDKIAPT
IKSGMRIQKGYVLGRIDQRLGFEVTMREKHINPLELIARN

>PROKKA_01041 cell division ABC transporter subunit FtsX

MNTLKKHLAFIIPLVALLFSLECVLFINQAIEQKEKKLIEDYSVVLASTQKLSLELLRQN
FSEIIALKEIDPNYSLEPLQKTLGIDGLKELRKNLPFFYSLQLSTFPTQERLENIKEKLL
KIPGVQKVEVFAKTYMQVYDLLSFIKTAVYIFALVVFVLSVLLMFKQVRIWIYQYHERLE
IMDLLGASVSFKNGFLYKIALMDSLIALFLAPMLMLYITSQKGFECTMDTLGIIGDAFVL
NHFLWGLLFSLVVSFVSVLLVAWRTRHV

>PROKKA_01042 Lipoprotein-releasing system ATP-binding protein LolD

MSVIIAANNLCLQYQQNEPVIKHANLRIKRKDFVFISGPSGSGKSTLLRSFYGDLKLFSG
KLEVCNINMNNASKATILDLRKNIGVVFQDYKLIQDYTIEQNIKLPMIICGIKKEECHLQ
LEKLLGHIDLRHKANRYPKELSGGEQQRVAMARAMANCEPILILADEPTGNLDDYSSDKIW
SLLRGMNTQLNATIVVTHKFKNFSAYHRKFYIEDGEVYEYS

>PROKKA_01043 tRNA (guanine-N(7)-)-methyltransferase

MPHFLAKLDFKPLEYPLIEGDFCFHKEFLSLKHPTKSCVYASFKDHIFLLQKIRRASDFL
IKSEKATPLKREVLKQALRIYSQSFEVISHNLQENSKHASQKKALDLEAFEDFIQNNQAP
ILAEIGFGSGRHLIELAKNNPTKTCLGIEIHTPSIAQALKQIELLDLKNLHILQGDGRLI
LESMPNHRCEKIFVHFPPVWNEKKHRRVLSEKFLNEALRVLKPRGFLELRDDSLYFEDS
LKLALKNFQCEIEIKNAQIPVISKEYEARWNKLKKDIYDLRIYSLELDETPFDNHAFDFS
FDTITMSKKS VKTILKTPKIIQEGYFVHVCNIYEDKGDFLVELSMGDFDWPVRLFVLLAE
NQIFYLNKSPLKTLNNHKAHLLLQNILSQKGIG

>PROKKA_01044 Fibronectin type III domain protein

MRSWMKKKYFTLLLQSSVVLAVFIGCSSTRNHTFSALNNQENTDTKLPVVHSIKTINDVS
SVGFEPKIIDTYDIDGFILYRLKKDSKLKRIATIKNPYATHYYDEGLETESSYTYQLAT
YKGDKISKLSEPVLVKTSFINPVESVFASLEYPKSVKVFWSPPHPNPSVSKYIIQRQNKDG
KFLNVGAVKNRFLVEFFDKDLEDGKKYRYQVIAENFMGDKSRPSIIVEGKTKDLPKEIAN
VRVSQNLTRHIELSWDKSTQEDVIAYRIYASNNRNDKYKFAQTNTTSYVDKIEKDNLTR
YYKVVAVDKTHLEGALPKEPAMGETADRPEAPIITKGTIQDSSAFIQWENNPSAKIATYA
VYRFEDTSKTPLRFGNITKNQFVDKDMKVGVAIRYQVVSVDKDGLESHPSKEVRLFLER

>PROKKA_01045 Ribosomal large subunit pseudouridine synthase D

MQKVFIALTDPHKLDEFLAKELQISKQVNLNLIKELVFCQKKEVKKGGLALKEGDAITL

LTPQITPKPLKKELDLEIEVIFEDEDLLVLNKPNNLVVHKAPSVKEPTLVDWLKSKNYEL
 SNLGLKERYGIVHRLDKDTSGGIVIAKNNFTHVHLSEQLKTKMMGRYYIALSTPLKEEK
 MSVECYLTRPNPNRLKMIALKAAKKEKSRYSKSEFISLLASQNGMDLIGAKLFTGRTHQI
 RAHLEYLNRHIIIGDNLGLNQLPKEEIRIMLHAYLIEFKHPRSEQKLRFKVPLLKDMLE
 YLKKVFDKENLDEVLDDEEKILHAFIAK
 >PROKKA_01046 hypothetical protein
 MEHDNSLGGAFNKETGELNIDKLSEYLSYSVVFVHSRVLH
 >PROKKA_01047 hypothetical protein
 MKRELLSFAKNWNIPTIFVFTNTQEKAGDAFVKESQRIIDEEWGFKGFVRAYVRVNSVKY
 LSDDDKEIPIKGLEELVDETKKCLIEARKNKQNHFLLIQKANIQKRKQAMIEECKTIIHV
 ASGAAGAAGLIPIPFSDALAIAPIQAGMIYKMNDAFGMDLDKSVGASLITGLLGVTAVAQ
 VGRTLVNGFLKFIPVVGSVAGGTTAVIITEGIGFAYLKVLEKCFNDETGEVNLPDEVGMI
 TSLFKENYLNLDTIKKLTQ
 >PROKKA_01049 Ribose-phosphate pyrophosphokinase
 MKARGFKAKMRGFKIFSGSAHPAFGKEVSKHLGFPLSKAVIGKFSDEINIQISESVRGK
 DIFI IQPTCVPVNDNLMELLVMVDALRRSSANSITAVLPYFGYARQDRKAAPRVPITAKM
 VANLMQEVGIERIITMDLHAGQIQGFFDVPVDNLYGSIVFRDYIRSKALKNPVIASPDVG
 GVTRARYFANQMGLDLIIVDKRREKANESVMNIIGSAKERDVILVDDMIDTAGTICKAA
 LALKEQGATSVMALGTHAVLSGNAIKRIKESALDEVVVTNSIPLVQKCDKITTLVAPLF
 AEVIRRIYHNESVQSLFT
 >PROKKA_01050 AP-4-A phosphorylase
 MQHLYAPWRESYLKEKNKSCVFCEISQNPTKDSENRVLYRNSDLFVVMNAYPYNPGHLLI
 IPHAHQASVELLDLNTWLNMMNILAPKVLKALYAYGAQGINLGLNLHRNAGAGIPEHLHMH
 LVPRFLGDSNFMSVIAQTRVCGIDLNETYLTNLNLEKELH
 >PROKKA_01051 UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase
 MQSLSWNLNAFRWLFITGLGYIMTLLQWYHYSVFRILTKHHKMRWHGIYFLLPLGVFIL
 SYAFKMPFVDFDFCGVIQMPMLIVWAKRNDKPLVFTPRVKRFFIFLFLILHEILNTEL
 VPLSGISLALGYLCLFIFVLSASLIFEKALSQYQLQTAKDKIASKNLKVIAITGSFGKT
 STKNFLLQILQATFNAHASPKSVNTLLGLANDINQNLDDRSEIYIAEAGARNKGDIKEIT
 HLIEPHLVVVAEVEGQHLEYFKTLENICETKAELLDSCRLEKAFCYSVEKIKPYAPKDSP
 LIDVSSLVKNIQSTLKGTSFEMLLDSVWESFETKVLGEFSAYNIASAILIAKHLGLETER
 IKRLVLELNPIAHRLQLLEVNQKIIIDDSFNGNLKGMLEGIRLASLHKGRKVIVTPGLVE
 SNTESNETLAQKIDGVFDVAIITGELNSKTIASQLKTPQKILLKDKAQLENILQATTIQG
 DLILFANDAPNYI
 >PROKKA_01052 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase
 MAKRSIAYLDSVFDISYTFIDHHSPLNALFLHGWGSSKEVMQQAFOGCFLNYNHLYVDLP
 GFNQSPNDEKVLETKDYANIINLFLKSVGKKAHVVFHGSFGGKVAILCESERIVLLSSAG
 ILEPKPLKVRCKILLAKIFKKLGLNLGFLRSKDAMGLNQAMYETFKKVISEDFSDFKRC
 EKEVLLFWGKDDKATPLSSAQKMQLTKRSALFVLEGDHFFFLNQAKEVEKLVENHYHAK
 S
 >PROKKA_01053 D-alanine--D-alanine ligase
 MEFCVLFGGASFEHEISIVSAIALKEALKDRIKHFIFLDEDDHFFYLIEASNMHISKYFAQI
 KEKKLPPLILTHNGLLKNSFLGAKIIEPLVINLVHGGDGEDGKLASLLEFYRIAFIGPR
 VEASVLSYNKYLTCLYANNLGVKTLGHVLLNEKNRANALDLNMFNFPFIKPNASAGSSLG
 VSVVKEEKELIYALDSAFEYSKEVLIEPFIQGVKEYNLGCKIKKDFCFSYVEEPNKQEF
 LDFKQKYLDFSRAKSPKANLSSALEEQLENFKKLYNDLFDGAIIRCDFFVIENKVYLNE
 INPIPGSLAHYLFDDFKTTLENLAQSLPKTPKIQVKNSYLLQIQKNK
 >PROKKA_01054 Phosphatidylglycerophosphatase A
 MDKFNLRAFCFLTLFFSGYSKAPGTIGSLVALLGLPVLIFSANTLFLGAVFVGLVAITQ
 IDKEEEETKRHDSSYIVIDELVGMWLAMASGLSLAGVLSFIFFRIDITKPSLIGKID
 KEIKGGLGVVADDALAGVLAGLSALLVIHILGFFNIKL
 >PROKKA_01055 Soluble hydrogenase 42 kDa subunit
 MLLFTPGPVAINEEMRSSFSQPMPHHRTKDFEKIFQSVRENKKMTGLEEVLLLNSSGTG
 AMEASVISLCQKELLFVNAGKFGGERFGKIAKAHFIKAHELVEYEDTPAQVDEILNALKAN
 PNIDAFCIQACESSGGLRHPVEKIAQAIAKETNPVNFVIVDAITALGVEPLEITHVDALIG
 GSQKAFMLPPAMSLIALSQKAIDRIEERDVGFYFNLKSELKNQRNNTTSYTAPILHTLGL
 QRYFELVQNLGGFEALYQETKRVALATQKAVLALGLKIFPKSPSLSMTTIVNEHAKELRN
 LLKEYQVQVQFAGGQEPYKDALIRINHMGIIPVYKSAYALNALELALNDLNLRGFDGVANT

TLKQYYGI

>PROKKA_01056 xanthine-guanine phosphoribosyltransferase

MHYSYETFLKDSLELVKQVEQICGVPEALVCVMRGGMTLTHFLSLHWDLREVGINAIISY
DTTHRQNALKIENIPTIKDHLKTILVVDEIVDSGNSLEAVLKVLDKHPDKKFYSASLFQ
KTSAKYKADAFKDAPEWIDFFWEVDLKNLKS

>PROKKA_01057 Ribosomal protein S12 methylthiotransferase RimO

MQVKENKQLCLISLGCSKNLVDSEVMLGKLYNYTLTNDACKADVILINTCGFIESAKQES
IQTILNAAKDKKEGAILIASGCLSERYKDEIKELIPEVDIFTGVGDYDKIDIMIAKKQNO
FSEQVFLSEHYNARIITGSSVHAYVKISEGCNQKCSFCAIPSFKGKLSRELDSILKEVE
DLALKGYKDMTFIAQDSSSFYLDKGQKDGLIQLIRAIKQALKSARILYLYPSSTTLEL
IGAIESSPIFQNYFDMPIQIHISDSMLKKMRRNSSQAHHLKLLDAMKQVKESFIRSTIIVG
HPEENEGEFEELS AFLDEFRDRLNIFAFSAEENTHAYSLEKVPKKIINARIKALNKIAL
KHQNHSEFKALLNKPIKALVENKEGEYFYKARDLRWAPEVDGEILINDSELTTPLKPGHYT
IVPSAFKDNILLAKVLSPF

>PROKKA_01058 hypothetical protein

MEVFQKLEFESLKELESKNNEEWENFTIAFYGETGAGKSTLIECLRLFFKERSKINQQE
RFRRLYAEKNDRGNHALLLELEKFQDGAIIIGDGRSDFTEKTKSYTLKHGNKTFTFLDVP
GIEGDEKKLFSKFLTPHKKPMLFFMLPKSPLLRKKEKKQKKGRLLKSKNNLIRKQRYGRFL
TNRLTTQKLLKMGLLMEAKKA

>PROKKA_01059 hypothetical protein

MGHQIVSAQMAFYGLLSALLPETDFDEKKQKFLKDFKAGELLYQSHFKLLAKFISEELLK
NSRAKIIQSNCNKALKVVEELQKAIEITIEKQIDPWIKETQEHQQEARNLDRSKEKFIS
DLEKSAFSKINQFESDFRIEMHERIERDIGNNECKEIKNELIQRTIELDENIKRRFEEC
EERFDEEIKKDIEQFEERVKDSLATERIGINNFSLNPNFNMDSGIDTMGLLASIGGLVL
LLITPIVGEFVLIAGVGLAFVGIGKAVWGFFDSYKKSQQRKEVDKKLDEVCEKIAEDVR
NQIESGKKGASKKIENLKGALNDLIVCYERM

>PROKKA_01060 hypothetical protein

MNETLEQFKANQKRQENLKKLLDFIHTGEKYGIQIEESLKDKIRNAIKNVADQKLKVAL
LGGFSEGKTSIAAAWIDRLDESMKIAHQESSDAVKIYDIDNEMELVDTPGLYGFKQKNTI
ATK

>PROKKA_01061 hypothetical protein

MSVEAHLILYALNPLNPIKESHKDDLNLWLFRTLNLSSRTIFVISRFDEEADIEDEEDYNK
RFKIKKENVQNRNLNELISLSEEEKESLIIVAVATNPFDLGVEHWLKHQEEFQKLSHIKTL
QDATQKKIKENGGLTIIEEAKKSVIQDVVYRQMPLAKKEQEIKREMEYLNKMIEKRRK
DLQNLNSEISQARINLREFIIRYFSGLIAQVSGTSLETFNDFAMKEIGDKGINIETVRN
AFERETQGIFNEMAKIETGFNADLSLFEKHAGALGKIGINFLNKS GFINATNIKLVRDTI
AAVGKFVGVLDVLKFKPWGAVNLAGSINKGLPLIGLAFEVWDSYNKYQK

>PROKKA_01062 hypothetical protein

MKSSFEEAKEQMIEFINDETFRFKQTCFPSALELEKSFQEIKEAIEKTQECDQEFEKWIQT
GEDFIKGEDFIEVEPERGMN

>PROKKA_01063 hypothetical protein

MFSKICPTLKLNAFRDSLFRKRISSPVQTTIRANMVLDIKNALEGENDPSNKVGKTLDL
VGFKKEYPQDFDELFEILKELIQEYEKNPDEVKQNLTEILK

>PROKKA_01064 hypothetical protein

MNHLILYNPYQQDVIQQHLSVLQEKSQVGF GKIRSKLNDQEKHDSLEEIYKATNEKNF
LQLFLTDYANLFAAKVIKVSKEIDESLIPSYKEKNLEVEDFFIISDLRELVRDFSLLR
DRFLANFIAPNNHTYAIYGNNYVYPLPVKLKEERSYFLGDEKHYSVYKSKEYLTMQENF
MRVFVFGKRLFYLLHPDSINNIIHAELELLQSENDLLNDFTSIIIVKYSKTLEYEIIYLF
VLLKACAKDPSLYDLAYKVQGKSFTLKDFFAKKPNLGSVKFLLKHEKIQCYLEENLNKFI
NYPFSRSLSLIQNIRNEAVHQAPGLNEVEKLRNEILGIEGASLLKGVLT HKETS

>PROKKA_01065 tRNA(Ile)-lysidine synthase

MTARDFKNYLEPLREGKNLLGFSGGLDSVCLFHLLVGENIAFDIALVDYNTQKQRLEIIQ
HAQTLAKTHHKKCHIHYPKPIARNFEMQARKIRYDFFEALTKEHSYKHLILAHHLNDRLE
WFLMQLSKGAGLNTLLGFQAYEKRESYAIRPLLYTPKDTLKTAKDHKFFEDDSNSSLK
FKRNNFRKNYANALMQDYSGKIIQSFKFLDQEKERLYPLMPVSQRHGITTFFQYSQNALFM
VDKILKQKGYVLSFSQKEEIKRHFFSLEIAQKFIIESDKEHVFIALKPPKTL SMPKDFKD
RARRLNIPKRLRPVLYAEFLKQPTHGFLTRFKQSFNL

>PROKKA_01066 tRNA-dihydrouridine synthase C

MDFKNKKWFLAPLAGYTDLPFRRSVVKKFGVDVTTSEMVSSHSLVYAFDKTSKMLEKSPL
EDHFMAQISGSKESVVEAVEKINALHVSIGDFNCGCPAPKVANHGNGSGLLKDLNHLV
KILKVIRENTNKKITSVKVRLGFEEKIPKEIAHALNDAPVDYVVVHGRTRSDRYQKDKID
YESIALMKKILKKPVIANGEIDSVKKA FEVLQITQADGLMIGRAALRAPWIFWQIRNNTT
ELPAVVKKDLVLEHFDKMFVEFYGDRGVIMFRKNLHAYAKGEMQASAFRNCVNTLTEIKSM
RESIEEFFNQEMLQSEVPLWVELNQKSV

>PROKKA_01067 Helicobacter outer membrane protein
MGRIESKKRLKALVFLASLGVWGNATEKTPFFKTKNHIYLGFR LGTGASMR TSMWQQAY
KDNPTCPSSVCYGEKLEARYQGGKNLSYTGQIGDEIAFDKYHILGLRVWGDIEYAKTQLG
QKVGGDTLLSQANYNPSAIKTYDPTSNAQGSNLQKTPNPQNFLFNNGHFMAFVLNVNVF
VNLPIDTLLKLALKTEKMLFFKIGVFGGGGVEYAILWSSQYKNQNTQQEDKFFAAGGGFF
VNFGGSLYIGKRNRFNVGLKIPYYSLSAQSWKNFGSSNVWQQQTIRQNF SVFRNKEVFVS
YAFLF

>PROKKA_01068 Helicobacter outer membrane protein
MIFLQQKH YFKEHFYEKDNFTLSLSLASSLLRAEDNGFFVGAGYQIG EAVQMVKN TGELK
NLNEKYEQLNQYLNQVASLKQSIQNANNIELVNSSLNYLKSFTNNNNYNSTTQSPIFNAVQ
AVITSVLGFWSLYAGNYFTFFVGNKD TKQPANVQGNPPFETIVKNCSGIENCAMDQTTYD
EMKKLAESLQAAQTNSTTKANNLCALSGCTTTESNTPNSTVSNALNLAQQ LMDLIANTK
TAMMWKNIVIAGVSNASGAIDSTGYPTQYAVFNNIKAMIPILQQAVTLSQSNHTLSASLQ
AQATGSQTNP KFAKDIYAF AQNQKQVISYAQDIFNLFDSIPKDQYRYLEKAYLKIANAGS
TPTNPYMAVVNLNQEVQTIQNNVSYYGNRVDAA LSVAKDVYNL KSNQTEIVTTYND AKNL
SEEISKLP HNQVNTKDIVTLPYDKNAPAAGQYNYQINQEQQS QLNQAL AAMSNNPFKKVG
MISSQNNNGAMNGLGVQVG YKQFFGESKRWGLRYYGFFDYNHGYIKSSFFNSSSDVW TYG
GGSDLLVNIINDXITKKEQQALRG SFWRHPTRRDYMA

>PROKKA_01069 Helicobacter outer membrane protein
MGLFGGIQLAGTTWLNSQYVNLTA FNNPYS AKVNASNFQFLFNLGLRTNLATARKKDSEH
SAQHGIELGVKIPTINTNYYSFLGTQLQYRRLYSVYLN YVFAY

>PROKKA_01070 Anaerobic C4-dicarboxylate transporter DcuA
MVDAFFQIIVLLFSFLGARLGGLGVGYAGGLGV LVLCFLGLNPGKIPFDVILIIMAVI
SAISAMQKAGGLDYL VQIAEKILRKHPKQINYLAPSVAYFLTILAGTGH TVFSLIPVIVE
VSQSQNIKPKAPLSLAVVSSQVAITASPVSA AVVFMSGILEPLGANYLTLLMVWIPTTFL
ACMLTAFVMGFTDLKLDSDPHYLERL KARNFSPPTIKEEKETSKSAKLSLWIFIGGVVAI
VFYASAISKNI AFVSPVVLGRDYAIVSFM LSVATLIVIFCKINANEIAHSSVFKSGMQAC
VCVLGVAWLGDTFVSNHIDEIKRYASFLIADY PFLLAVALFLASMLLYSQAATSKALIPS
VITALGISANHTEHLYIIVASFASVSALFVLPT YPTLLGAIAMDNTGTTKMGRYVFDHAF
LIPGVLVVFLSVALGFVAPLVL

>PROKKA_01071 putative L-asparaginase
MRIFLKL LILLFCLKGQVMAQNLPTIALLATGGTIAGSGASASLGSYKSGELGIKELLKA
IPSLNKIARILGEQISNIGSQDMNEEVWFKLAKRAQELLDDSR IQGVVITHGTD TLEESA
YFLNLVLRSTK PVALVGAMRNFTLSADGALNLYNAL SVALDEKSANKGVLVVMDDNIFS
AREAIKTH THTSTFKALNSGAIGSVYYGKTRY YMQPLRKHTIESEFSILELKPPLPKVD
IIYTHAGMTSDLFQASLNSHAKGVVIAGVGNGNV SAGFLKAMQEASEMGVLIVRSSRVGS
GEITSGEIDDKAFITSDNLNPQKARVLLQLALTKTNDKAKIQEMFEEY

>PROKKA_01072 Helicobacter outer membrane protein
MKKTKKTILLSLTLAASLLHAEDNGVFLSVGYQIG EAVQKVKNADKVQKLS DAYEQLSRL
LTNDNGTNSKTS AQAINQAVNNLNERAKTLAGGT TNSPAYQATLLALRSVLGLWNSMGYA
VICGGYTKSPGENNQKNFHYTDENGNGTTINCGGSTKSDGTHSSNGTNTLKADKNVLSI
EQYEKIHESYQILSKALKQAGLAPLNSKGEKLEAHVTTSKYQ QDNQTKTTTSVIDTTNDA
QNLLTQAQTIVNTLKD YCPMLIAKSSSGSSGQATTNTPSWQTAGGGKNSCATFGAEFSAA
SDMINNAQKIVQETQQLSANQPKNITQPHNLNLNTPSSLTALA QKMLKNAQSQA EILKLA
NQVESDFNKLSSGHLKDYIGKCDASAISSANMTMQNQKNNWGN CAGVEETLASLEKSNA
SFSSQTPQINQAETLANAIVQELGHNPFKRVGI ISSQTNNGAMNGLGVQVG YKQFFGEKK
RWGLRYYGFFDYNHAYIKSNFFNSASDVW TYGVGSDLLFNFINDKNTN FLGKNNQISFGL
FGGIALAGTSWLNSQFVNLKTI SNVYSAKVNTANFQFLFNLGLRTNLARPKK KDSHHA AQ
HGMELGVKIPTINTNYYSFLDTKLEYRRLYSVYLN YVFAY

>PROKKA_01073 hypothetical protein
MKKALKILSVSALLFVALNAKDFS KTSDEDLAKMAGV VAPQDIVDYTKELKKRMEKMPED
KRKAFHKQLHEYAVKNTDNMTVADFEARQKAVKEAL KKGNMEDMDDDLGLRSCKHEKMHK

HGGHDKHKGKGDHKKQGKKDHDHNDHGENEK
 >PROKKA_01074 hypothetical protein
 MRRSLAFCLLALLGLQVLGARDFSQLKNEELLKLAGTLPSNEAIDYRMEVSKRLKALSAE
 DAKKFRANFSRIARKNLSKMSEEDFKKMREEVRKELEEKTKGLSAEEIKAKGLNVSVCSG
 DTRKVVWCRAVKKKDEHCSPK
 >PROKKA_01075 Arginine exporter protein ArgO
 MFVVFIEGFGLAISLCAAVGAQSLFIIERGMAVNYVFLICTLCFMCDIVLMSMGVFGVGA
 YFAKNLYLSLSLNLFGAAFTGFYAFLALKTLFQTFKKKQVQTPKKLSLKKTLTLLFTLGVT
 LNPQVYLEMVFLIGASALSFDLAQKFVFLAGTLSAAFSWLLLLCALSLRYGPKLLNNQKI
 FMGVNLFVTAIMGTLSVTLFRDFLALLSQT
 >PROKKA_01076 DNA polymerase III subunit tau
 MQVLALKYRPHFSELVGQESVAKTSLALDNQRLANAYLFSGLRGSGKTSSSRIFARAL
 MCEEGPKAVPCDTCIQCSALNNHHIDIEMDGASNRGIDDVRNLIEQTRYKPSFGRYKI
 FIIDEVHMTTEAFNALLKTLEPPSHVKFLLATTDALKLPATILSRTQHFRFKKIPENS
 VISHLKTILEKEQVSYESSALEKLAHSGQGSRLDITITLLEQAINYCDNAITEGKVAEMLG
 AIDRSVLEDDFFQSLINQDEVRLKERYAILENYETESVLEEMMLFLKAKLLSPDSYSILLI
 ERFFKIIMSSLSLLKEGANASFVLLLLKMKFKEALKKALDDAILELEQSKENAFQPLNQ
 NANAPKQEPKSAEKIEKLEKIESAETPQTPMLSADKDRIFHNLFKQVQTLVYERNYELGAV
 FEKNIRFIDFDSQTKTLTWESLATNKDKELLRERFKIVKGIVDGVFGKGESIKIALKNQN
 KSALEVVKFEPYSPKPKPTTETTAETKEKEIQENDTKEVQETQPKQAPTALQEFMANNS
 NLIEEIKSEFEIKSVELL
 >PROKKA_01077 tRNA threonylcarbamoyladenine biosynthesis protein TsaE
 MRAHLDELDKVAAAILKDDFKGVLLKGVVGSGKTTLVQACLKHLGLDIQATSPTFSLMH
 AYSESVFHYDFYMRDLKACLELGMLECLLEKGIHFVEWGDEKLEKILKKYDLAIKVVEIK
 TESTSRFYTIKIA
 >PROKKA_01078 Lipopolysaccharide export system ATP-binding protein LptB
 MDILKAEHLNKQIKKTKIVSDVSLEVKSSEGVVGLLGPNAGKTTTFYMICGLLEPSGGSV
 YLNDVNLAKYPLHKRSNLGIGYLPQESSIFKELSVENLALAGESTFKNSKESEEKMESE
 LDAFNIQAIRERKGMSSGGERRRVEIARALMKNPKFVLLDEPFAGVDPIAVIDIQKIE
 SLIGLNLGLITDHNVTLSVCHRAYVIKSGTLLASGNANEIYENTLVRKYYLGENFKV
 >PROKKA_01079 RNA polymerase sigma-54 factor
 MAILRANLSLKNKLNATLKGWLPILQSELEDLEEVVKQNALDNPLIKIENKRIKNFSDRF
 SAKKSSDHLENFAIASKSLFETLEAQIIPPLFPTEISQKIAMDIIISGLNNEGYPFEENIEE
 RAKILGVESEVYEKVRKRFSYLNPAIGARDVKESFLFQLESRELDNNELYEEVRKIILN
 LEKHHEFSKDFYFEKALKILKSFKNPPAIEFLEKEIEVPELFIIVEVDEEIIIVRLNDESY
 PTISLEENRFENSGYLKEKLKEAKDLIDALNLRKATIIYKIGLMLLEYQYDFFKGKELRPL
 KLLDLANEFNHSVSTISRANLKYLAECRGVFPKHFHSTALDNSETSNVAVIKDYLLLELI
 KNEDKKEPLSDAKILELIEEFHLLKMRRTITKYRQLLNIIASSSERKRLYLMRA
 >PROKKA_01080 hypothetical protein
 MLIFYSVLEQNLIPFVITKEQKEAYIKALDTRNTESLYQLAKVSQKFELTRIQQQMILNK
 NKP
 >PROKKA_01081 Fic/DOC family protein
 MPQRAFFEVKNYQNMFLFLLNENLKGQSVDSFFIRELHGILMNFLPNKGAFKTTDNTIL
 GASFETTPHFQVPMAMKEWCDNFNYKMKTLQDKEEKLKAILEQHILFERIHPFSDSMVGW
 AEC
 >PROKKA_01082 hypothetical protein
 MNYKELLEFNNDYAMDLTIRTAAHSTAIENNPLILKL
 >PROKKA_01083 5'-nucleotidase
 MYAAHPIKPLKAPKLKTKFLRRVFGASIRRWNDQACPLEFVELDKQAHKAMIAIYLLAKD
 LKDRGNLDLDDLIIKYFCFEFLERLVLTDIKPPIFYAIQQTHSQELASYVAQSLQDEISA
 YFSLEELKEYLSHRPQILETQILES AHFYASKWEFDIIYHFNPNMYGVKEIKDKIDKQLH
 NNEHLFEGFLFGEKEDLKKLMSMFQGLRFQKRWSQTPRPQTSVLGHTLCVAIMGYLLSFD
 LKACKSMRINHFLGGLFHDLPILTRDIITPIKQSVAGLDNCIKEIEKKEMQNKVYSFVS
 LGVQEDLKLYFTENEFKNRYKDKSHKIIFTKDAKELFTLYNSDEYLGVCCELLKVC DHLSA
 FLEAQISLSHGISSNDLIKGAQNLLLELRSQTELLDLGLKLFDRDFK
 >PROKKA_01084 Helicobacter outer membrane protein
 MRKLFIPLLLFSALEANEKNGFFIEAGFETGLLEGAQTQEKRHHTTKNTYATYNYLPTDT
 ILKRAANLFTNAEAIKSLKFSSLSPPVRVLYMYNGQLTIENFLPYNLSNVKLSFTDAQGNV

IDLGVIETIPKHSKIVLPGEAFDSLKEAFDKIGPYFFFFPKIEATSTSISDTNTQRVFETL
NKIKTNLIMKYSNENPSNFNTCPYNNGNTKNDWCQNFTPQTAEFTNMLNMI AVLDSQ
SWGDAILNAPFEFTNKGGGGECDTSKENDCVNPGTNGLVNSQNKSYVLNKQDIVNKFRNK
ADLDVVVLKDSGVVGLGSDITPSNDDGKHYGQLGVVASALDPKKLFGDNLKTINLEDLR
TILHEFSHTKGYGHNGNMTYQVRVPTGQENENGKPKDSGLPYNVCSLYGGSNQPAFPSNYP
NSIYHNCADV PAGFLGVTAAVWQQLINQNALPIDYANLSAQTNYNLNASLNTQDLANSML
GTIQKTFVTSSVTNHYSSASQSFRSPILGVNAKIGYQNYFNDFIGLAYYGIIKYNYSKA
LNQKFQQLSYGGGIDLLDFITTYSNKNSPTGVQTRRNFSSSFQIFGGLRGLYNSYYVLN
KVKGSGNLDVATGLNYRYKHSKYSVGISIPLIQKKASVSVSSGGDYTNSFVFNEGASHFKV
FFNYGWVF

>PROKKA_01085 Adenosyl-chloride synthase

MKKTILALFLSACIGLSSVCASNALILQTD FSLKDGAVSAMKGVAFSVDSNLKIFDLTHE
ISPYNIWEGAYRLYQTASYWPKGSV FVSVDPGVGTKRKS VVLKTKNGQYFVSPDNGTLT
LVAQTLGIDS VREIDEKTNRLKGSEKSYTFHGRDVYAYTGARLASGAITFEQVGPELPPK
VVEIPYQKAKATKGEVKGNIPILDIQYGNVWSNISDKLLNQAKIKLNDTLCVTIFKGSKK
QYEGKMPYVASFGDVLEGQPLVYLNSLLNVSV ALNMDNFAQKHQIKSGADWNIDIKKCAK

>PROKKA_01086 hypothetical protein

MNIKTHSSNEKERFVHI DEDEKKELFAGTANENSHGLSLMNLIGVLVFGGAFLALLAPKI
YLSNNIYYISRKINTLEDQKRL LLEEQQILKNELEKERFKYYIENSENIGDIAF

>PROKKA_01087 Ribosomal RNA small subunit methyltransferase H

MQEIESLHQSVLLQEV LQAFTPLEEGVLIDCTLGLGGH SKALLSQKPHLKLIGIDKDKFA
QEIAKERLKA FEGRYNLLSGGF AKRFKEALEIHGERIKGV LVDLGVSSLQLDDDNRGFNF
HSHTLDMRMDLESELNAQKVINSYPVVALEKIFRDYGEIKEYKKIAHKIAERRAKKPF RD
AKDLSDFLSSFSKNKKIHPATLVFQAVRIEVNNELEELKEFLQCAR NLKGAILCVISFHS
LEDALVKNAFKNYAKNCICDPLSFKCACSNH ALGEILTCKPITPSPEEIKNNRRSRS AK
MRVFQFKP

>PROKKA_01088 Helicobacter outer membrane protein

MEFMKKFVALGLLSAVLSSSLAEGDGVYIGANYQLGQARLNSNIYNTGDCTGSVVGCPP
GLTANKHNPGGTNINWHAKYANGALNGLGNVGYKKFFQFKSFDMTSKWFGFRVYGLFDY
GHATLGKQVYAPNKIQ LDMVSWGVSDDLADIIDKDNASFGIFGGVAIGGNTWKSSAANY
WKEQIIIEAKGPDVCTPTYCNP NAPISTKTSTVAFQVWLNFGVRANIYKHNGVEFGVRVPL
LINKFLSAGPNATNLYYHLKRDYSLYLGNYTF

>PROKKA_01089 UvrABC system protein A

MQHKTIMDKII IQGAREN LKNIFLEIPKNQFVVFTGLSGSGKSTLAFDTLYAEGQRRYL
ESLSSYARQFLDKVGKPNVDKIEGLTPAIAIDQKTTSKNPRSTVGTITEIYDYLRLLFAR
IGEQFCPTCLEPISSMSASDIISQICHLEEDSKIIILAPIIKDKKGSFNDKLESLRLKGY
VRAFDVGVMVRLDEEIH LHKTKKHTIEAVVDRVVINSSENSRIASAVEKALKESYGELEV
EILQDNVPSIRKHYSEHKACFKCKMSFE ELEPLSFSFNSPKGACESCLGLGTFKFSLDISK
ILD PNTPLNQGAIKVIFGYNRSYYAQMFEGFCEYD GIDTALCFNELNKEQQDALLYGNGT
EISFHFKN SPLKRPWKGI IQIAYDMFKEQKDLSDYMSEKTCSSCEGHRLKASSLSVQVAG
LKMA DFLTKEIEEVYHFFNNP THFSYLNEQEKKIAEPILKEILERVFFLYDVGLGYLT LG
RDARTISGGESQRIRIASQIGSGLTGVLVYLDEPSIGLHEKDTLKLINTLRNLQKKGNTL
IVVEHDKETIKHAD FVVDIGPKAGR HGGEVVFGSVKELLQNNHSTALYLN GTKKIERPK
FELPKEKH FLEIKNVNINN IKNLSVQIPLKQLVCITGVSGSGKSSLILQTL LPTAQTL LN
HAKKAQSLNGVEIVGLEHLDKVIYLDQAPIGKT PRSNPATYTGVMDEIRILFAEQKEAKI
LGYSASRFSFNVKGGRC EKCQGDGDIK IEMHFLPDVLVQCDSCGAKYNPQTLEIKVK GK
SIADVLNMSVEEAYEFFAKFPKIAVKLKT LIDVGLGYITLGQNATTL SGGEAQRIKLAK E
LSKKDTGKTLYILDEPTTGLHFEDVNHL LQVLHSLVALGNSMLVIEHNLDI IKNADYIID
MGPDGGDKGGRV IASGTPLEVAQNCEKTQSYTGKFLALELE

>PROKKA_01090 hypothetical protein

MIDNNNTALADLLENGEPTIYYEANSANIRRF LDKGPYNIEWLVRKYLGVGFRSCKVERI
RETEFVIKFRYKKWGS MIDREV KMQFPKYLKDSKKTAVKITSKV FYELLKGT LFGNLLGK
DPIQAIIMRKENNNK

>PROKKA_01091 hypothetical protein

MLANDFMVKKLKNKELHGLLCNKLSKRVD RFGYRVIGVGADKEITKLLFN

>PROKKA_01092 Transcriptional regulatory protein QseF

MKIAIVEDDINMRKSLELFFELQDDLEIVSFKNPKDALAKLDESFDLVITDINMPHMDGL
EFLRLLEGGYESIVITGNATLNKAIDSIRLGVKDFFQKPFKPELLLESYRTKKVLEFQK

KHPLEKPLKKPHKHSFLATSKALEESKRQALKVASTDANVMLLGESGVGKEVFAHFIHQH
SQRSKHPFIAINMSAIPHEHLLSEELFGYQKGAFDTATVPKMGLFESANKGTIFLDEIAEM
PFQLQSKLLRVVQEKEITRLGDNKSVKIDVRFISATNANMKEKIAAKEFREDLFFRLQIV
PITIAPLRERVEEILPIAEIKLKEVCDAYHLGPKSFSKNAAKRLLEYSWHGNVRELLGVV
ERAAILSEGAEIQEKDLFLER

>PROKKA_01093 hypothetical protein

MHALSPLLEDQEFLLISYRLKIVDSRVMGEEYSVSKPIVSRIKTAPYVLDYRCSIITRNPPN
LKDPLLPIKLERFLEIALKKEKERVIDCILKSQVAITHYDHSYKNGTTTTSILDLKALS
VRASLVGDALFLDIFRKEEE

>PROKKA_01094 DNA gyrase subunit A

MQDNLINETKNIVEVGIDSSIEESYLAYSMSVIIIGRALPDARDGLKPVHRRILYAMHEL
LTSKVAYKKSARIVGDVIGKYHPHGDNAVYDALVRMAQDFSMRLELVDGQGNFGSIDGDN
AAAMRYTEARMTKASEEILRDIDKDTIDFVFNYYDDTLKEPDILPSRLPNLLVNGANGIAV
GMATSIPPHRMDEIIDALMHVLENPNAAALDEILEFVKGPDPFTGGIIYGKAGIIIEAYKTG
RGRVKVRKAVHVEKTKNKEIIVLDEMPFQTNKAKLVEQISDLAREKQIEGISEVRDESDR
EGIRVVIELKRDAISEIVLNHLYKLTTMETTFSIILLAIYNKEPKIFTLLELLRLFLNHR
KTIIIRRTIFELEKAKARAHILEGYLIALDNIDEIVQLIKTSPSPEAAKNALMERFSLSE
IQSKAILEMRLQRLTGLERDKIKEEYQNLLELIDDLNGILKSEDRLNGVVKTELLEVKEQ
FSSPRRTEIQESYESIDIEDLIANEPMVVSMSYKGYVKRVDLKAIEKQNRGGKGLSGST
YEDDFIENFFXANTHDILLFITNKGQLYHLKVYKIPEASRIAMGKAIVNLISLAPDEKIM
ATLSTKDFSDERSLAFFTTKNGVVKRTNLSEFGGNSRYSGIRAIVLDKGDELVGAKVVDNN
AKHLLIASHLGIKIFPLEDVREIGRTTRGVGRIKLNNDFVVGAVVISDDSNKLLSVSE
NGLGKQTLAEAYREQSRGGKGVIGMKLQKTGNLVGVISVDDENLDMILTASAKMIRVS
IKDIRETGRNASGVKLINTADKVMYVNSCPKEEEPENLENSPAKFV

>PROKKA_01095 Diacylglycerol kinase

MSDFKVPPKAKGFKRLFKALFYSKDGLKYAWAEESAQRQIIIVLALFCIVLASYLAKDFLE
WGLLILPCFLSVVVELINSSIEKAVDFTSTEFHPLAKKAKDMASSAQLIGLIFWAFIWGR
YLLTLYLK

>PROKKA_01096 hypothetical protein

MVVIGALICVLLGVFIFFTSMSVKKSLSTYLNAYLDQRPHIKGMGIVGAPFECEGFFKIA
CVSKELRFLDSQNSPIVDFKNLSIKLNSLDKSSILSVHSQIKSPILEQSIQKTHQIPL
KDLNALLEKIKPHTLNCSLKFNALDEKTLNDHLKCDLTNAENILAYTFFQDGLMEAQENL
SLKNIFKTLSSSKDAKAEELQDKLRFSAPKLGVSIIQARHLKNLLEAFYHQNKESLGGFFSP
YFSLRSQTPSVSYESALSSLENYFMVLFQSRFKDDTALQQNFKGLLQAFVSMADKQSQI
VLNAQAKDNAKLTFNALDLSLVNFFQSYQISHE

>PROKKA_01097 Acetone carboxylase gamma subunit

MSKYTQEQIKNLVEGNLDWNTVLKMLSMPKDHERFQMYLKVLQDKVDFDDKIVLPLGPHL
FVVQDAQKKWVIKCSGHAFCAPENWKLHANIYVRDTAEKMEEVYPKLLASDTNWQVYR
EYICPDCGILLDVEAPTPWYPVIHDFEPDIEVFYKDWLGIQPPERR

>PROKKA_01098 Acetone carboxylase alpha subunit

MANLLKNGKTLKQARDEILARTEKTGHYNGLKLEFKERDPIGYEKMFSKLRGGIVHARE
TAKRIAASPIVEQEGLCFTLYNAVGDVLTSTGIIIHVGTMGSAIKYMVENNWEDNPGI
NDKDIFTNNDCAIGNVHPCDIMTLVPIFHDEKLIGWVGVTVIDTGSVTPGSMSTGQVQ
RFGDGYMITCRKTGANDESKDWLHESQRSVRTPKYWILDERTRIAGCHMIRDLVMEVIK
EDGIDSYMRFIDEVIEEGRGLISRIKSMTIPGKYRKVAFVDVPYAHKDIGVCSEFAKLD
TIMHSPVEITINKDATWKLDLFDGASRWGWSFNCNQVSFTSGIWMMTQTLIPTSRINDG
AYFATQFRLKKGTWMNPDDRRTGHAYAWHFLVSGWSALWRGLSQAYYSRGYLEEVNSGNA
NTSNWLQGGGINQDGEIHAVNSFETSSCGTGACAIKDGLNHAAAIWNPEGDMGDVEIWEM
AEPLLYLGRNVKANTGGYGYRGNGGFETLRMVWGAHDWTMFFMGNGYMNSDWGMMGGYP
AASGYRFEAHNTDLKNRIKNNASLPLGGDFNPPTDRDYEKHISHASQVKRDKQCITTENC
DNYDLYLNYIKGGPGFGDPIDRLNAILEDLNSKQLLPEYAYKVYGAIVVSQNKDGVWVG
EAKTKARRKEILENRKARSIPVKEWMEQERNAILEKEASKQVKHMYATSFDLSPKFLNDF
KTFWNLPKSWSVKEDELGVFTYGSKYRMDLSKLPDVRTVLLVDEK

>PROKKA_01099 Acetone carboxylase beta subunit

MICLLQSHKNAESERIVRDIALKEIEKLGKNIPVFASVDYYPQRKESHMNTTILEAYAA
EPSRQTLKSVSNRFKEHGAKFDLRVMATHGGTISWKAKELARTIVSGPIGGVIGSKLLGE
TLGYDNIAACSDIGGTSFDMALIVKSNFNIASDPDMARLVLSLPLVAMDSVGAGAGSFVRI
DPHSRSVKLGPDAGYRVGTCWKDSGLDVTDSVTDCHIVLGYLNPDNFLGGLIKLDVDRAK

KHIKEQIADPLGISVEDAAAGVIELLDLELKEYLRSNISAKGYSPSDFVCFSYGGAGPVH
TYGYTEGLGFKDVVVPAAWAGFSAFGCACADFEYRYDKSVDAIPQYSSDKSKIEACKII
QDAWDELTLKVIEEFKINGFSQKDVILRPGYRMQYMGQLNDLEITSPVSKAASVADWEEI
VKEYEKTARVYSESACSPELGFSVTGVMIRGVVATQKPVIPVEKEHGATPPKEAKVGVR
KFYRHKKWVDADVWQMEKLLPGNEVIGPAIVESDATTFFVIPKGFATRLDKHRLFHLKEVK
>PROKKA_01100 Acetone carboxylase beta subunit
MKDAKVQVMGIDAGGTMTDTFFVKENGFSFVVGKAQSNPEDESLAIYNSSQDALSHWQSDV
SKVYPELVTCVYSGTAMLNRRVQRRGMEVGLICNKGFEQMHSMGRALQSYLGYALEERLH
INTHKYDDPLIPLKIRIGVTERTDVKGQVVIPVRQEEVKVAVKELLEAGAKAM
>PROKKA_01101 hypothetical protein
MFFKFILCLLLGVFAWAKEDIPTPLTPSKRYSINLITENDGYINPYIDEEYTAGNQIGFS
TKEFDFSKNKAMKWTSYLGFNFKSPRVTRFGISLAQDMYTPSLENRKLVLHLDNHPYGGY
LRVNLNVYNRHQTFMELFTLSLGTGQDSLAAQTQRLIHKWGHDPQFYGWNTQLKNEFIF
ELHYQLLKKVPLLKTRFFSMELMPGFNVELGNARDYFQLGSLFRAGYNLDADYGVNKVNT
AFDGGMPYSDKFSIYFFVGAFFGRFQPLNIFIQGNSPETRGIANLEYFVYASEIGAAMMWR
SLRVAFTITDISKTFQSQPKHHQIGTLELNFAF
>PROKKA_01102 Short-chain fatty acids transporter
MFLLRHLTSACVFLASKCLPDSFVLVTLTSLFVFLVYCLTGQDASSIISWGNGAWTLL
GFSMQMALILVLGQALASAKLVQKLLKYLASLPKGYTALWLVTFLSLIANWINWGFGLV
ISAIFAKEIAKNVKGVDYRLLIASAYSGFVIWHGGLSGSIPLSVATQNEEDLSKISTGVIE
KAIPISQTIIFSAYNLIIGIILVGLPFLMAMHPKKEEIVEIDPKLLKDEYKETELIDHQ
QDKTIAHFLENSALLSYLLVFLGFGYLGIIYFFKGGGLGLNIVNTIFLFLGILLHKTPLAY
VKAINHSARSVAGILLQFPFYAGIMMMASHSVGGHSLAQMLSLAFTHIANEKTALMTF
LSAGIVNIFIPSGGGQWAIQAPIMLPAGQSLGVDPGVVSMAIAWGDATNMIQPFWALPA
LAIAGLGAKDIMGYCVLTTLIFVGLVVCVGFYFLV
>PROKKA_01103 Succinyl-CoA:3-ketoacid coenzyme A transferase subunit B
MREAIKRAAKELKEGMYVNLGIGLPTLVANEVSGMNIVFQSENGLLGIGAYPLEGGVDA
DLINAGKETVTVVPGASFFNSADSFAMIRGGHIDLAILGMEVSQNGDLANWMI PKKLIK
GMGGAMDLVHGAKKVIVIMEHCNKYGESKVKKECSLPLTGKGVVHQLITDLAVFEFSNNA
MKLVELQEGVSLDQVKEKTEAEFEVHL
>PROKKA_01104 Succinyl-CoA:3-ketoacid coenzyme A transferase subunit A
MNKVITDLKALSGLDGDTILVGGFGLCGIPEYAIIDYIYKKGKIDLIVVSNCGVDDFG
LGILLEKKQIKKIIASYGENKIFESQMLNGEIEVVLTPQGTLAENLRAGGAGIPAYYTP
TGVGTLIAQGKESREFNGKEYILERAITGDYGLIKAYKSDTLGNLVFRKTARNFNPLCAM
ASKICVAEVEEIVPAGELDPDEIHLPGIYVQHIYKGEKFEKRIEKITTRSAK
>PROKKA_01105 Acetyl-CoA acetyltransferase
MNEVVVVAAKRSVAGSFLGSLKSVGAREMGVCVLKDALNASGLKPSDVDSVILGNVLGAG
LGQNIARQIQLDAGIPNDKNAFSVNMCVCGSSMKAIQLAHD SIMLGRDEVVVC GGVENMSA
APYLSFDMREGKRMGNANMIDSMVHDGLWDAFNDYHMGITADNVAQAYHISREEQDNFAL
QSOLKARAAINAGKFQEEITPIEIANKKGVVVFKEDEYPRDTTLES LAKLP AFKKDGSV
TAGNSSGINDGASIIILCSAKKAQKLGLKTMATIKGFGGLGGCSPDIMGICPSIAIKNNLK
NVKMDLNDIDL FELNEAFAAQSLAVLKELELNPNI NVNNGGAIAIGHPIGASGARILVTL
LHEMKRS GHGVGCASLCVGGGQGLSVVVEQK
>PROKKA_01106 putative 3'-5' exonuclease related to the exonuclease domain of
PolB
MLCVFDIETIPSVSLCKEHFQLEENDALKICERSFEKQKEKSGSEFLPLYLHEIISIAAV
IGDDYGQFIKVGNGFQKHENKEDFTSEKELLEDFFKYFNEKQPRLISFNNGRGFDIPLTL
KALKYNLTLD AFYNQENK WENYRVRYSEQFHLDLMDSLSHYGSVRGLNLNGVCSMMNIPG
KFDVSGDLVHAIYYPHLSQKEKKGIIDGYCQSDVLNTYWLFLKYEVLKGALNKEQYLGL
LSDFLKKFPKEKSYSSVFTNALEKEIREFA
>PROKKA_01107 Ferrous iron transport protein B
MKEITIALVGQPNVGKSSLINALSNAHLKVGNFAGVTVDKMEVGLIHKEHQITIIDLPGT
YALNDFTTTEEKVTKDFLEKGQYDLILNVVDSTNLERNLALSQQLD TNKKMLLALNMWDE
AQKEGVKIDTEKLSKELGVVCVPTSARSKEDRFNTELLLN EIVRLYSQNTTNNESIKVPS
QSFKESLKYSQSAQRIACLVI SENQQNASFEHTYKIDKILMHPRYGIFIFLGFMFVIFSL
SFLIGGGVQKALEAGFKFLSDGIKENVANEDLASLVGDGIIGGVGATVSFLPLIVVLYFG
ISLLETTGYMSRVAFLLDGILHKFGLHGKSFIP LITGFGCSVPAYMATRTLQNYNERLIT
LFVIGFMSCSARLP IYVLFVGSFFPSSSAGFVLCIYILGAVVALVMAKLLKLSVFKGQT

ESFIMEMPKYRFPSWRMVYFSIYTKSLSYLKKAGTYILVGAILIWFMSQYPKSDAAMKTY
KQESLLVDKNTALSSEAKEEKLKELKAELDKKNLKSIVGRGGAYLEKVFSPMDFDWRLS
VSLVTGFMMAKEVVVSTLGLVFLSLGDQNEKSDAFRGILRKEVSVPSGIAFIVFVMFYIPCF
AATITFGREAGGIKFVAYLFIFTTVVAYAFSLIAFYATQILV

>PROKKA_01108 Fe(3+) dicitrate transport protein FecA precursor

MKRILVFLAVLSHSAHAVKTHNLERVEASGVANDKEAPLSWRSKEVRNMGSRITVISNKQ
LTKSANQSIEEALQNPVGVHNRNSTGIGAVPSISIRGFGAGGPGHSNTGMILVNGIPIYV
APYVEIGTVIFPVTFQSVDRISVTKGGESVRYGPNAFGGVINIITKGIPTNWESQVSERT
TFWGKSENGGFFNQNSKNIDKSLVNNMLFNTYLRTGGMMNKHFGIQAQVNWLKGQGFRYN
SPTNIQNYLLDSLYQINDSNKITAFFQYYSYFLTDPGSLGIAAYNENRFQNNRPNNDKSG
RAKRWGAVYQNFFGDTDRVGGDFTFSYYGHMSRDFKFDSNYLNVNTNPKLGPVYTDQNY
PGFFIFDHLRRYVMNAFEPNLNLVVNTNKVKQTFNVGMRFMTMDMFIRSDQSTCEKSDII
DGVCHMPPYVLSKKPSNNQEMFNNTAVWLSDKIELFDSKLVITPGLRYTFLNYYNKEPE
KHDFSVWNITKKRQNEWSPALNIGYKPMENWIWYANYRRSFIPPQHTMLGITRTNYNQIF
NEIEVGQRYSYKNLLSFNTNYFVIFANRYAGGYSPQPINARSQGVELELYYAPIRGLQF
HVAYTYIDARITSNADDIAYYFTGIVNKPFDIKGKRLPYVSPNQFIFDMMYTYKHTTFGI
SSYFYSRAYSSMLNQAKSQTVCLPLNPEYTGGLYGCNSVGLLPLYFVLNVQVSSVLWQS
GRHKITGSLQINNLFNMKYFRGIGTSPTGREPAPGRSITAYLNYEF

>PROKKA_01109 Flagellar biosynthetic protein FliP precursor
MEPSLKKAYDTGIKPYMDKKISYTEAFEKSALPFKEFMLKNTREKDLALFFRIRNLPNPK
TPDEVSLSVLIPAFMISELKTAFAQIGFLLYLPFLVIDMVISSILMAMGMMMLPPVMISLP
FKILVFILVDGFNLLTENLVASFKMV

>PROKKA_01110 Flagellar biosynthetic protein FliP precursor
MSADSALPSVNLNAPSDPKQLVTTNLVIALLTLLVLAPSLILVMTSFTRLIVVFSFLR
TALGTQQTPPHSNFSLALFDIDFFHHGT

>PROKKA_01111 Bifunctional protein GlmU

MLSVIILAAGKGTRMRSSLPKTLHTICGEPMLFYILETAFSISDDVHLILHHQQERIKEA
VLERFKGIIIFHTQIVEKYSGTGGAIMQEDKTPIPTKHERVLILNADMPKITKDALTPLLE
SKNNAIGLLHLADPKGYGRVILENHQVKKIVEEKDANDEEKAIQSVNAGVYGFERKFLEK
YLPKLNQNAQKEYYLTDLIALGINKNETIDAIFFLEEECFGLGVNSQTERAKAEEIMLERL
RKNAMD LGVVMQLPSSIYLEKGVSFKEGCVLEQGVRLIGNCLIENARIKAYSVIEESQIV
NSSVGPF AHARPKSVICDSHVGNFVETKNAKLQSAKAGHLSYLGDCIEGKNTNVGAGVIT
CNYDGKKKHQTIIGENVFIGSDSQLVAPINIGSNVLIGSGTTITKDIPSGSLSLSRAPQT
NIENGYFKFFKKS

>PROKKA_01112 hypothetical protein

MSGLKAFSCVVVLCGAMANVAIAGPKIEARGQFGRDIGEGVGAGMGGMIGALGGPW
GTVFAGIGGGIGAYSGAEIGDRVEDFIRGVDREPQAPREPTYDRHFVYDR

>PROKKA_01113 hypothetical protein

MNIKERLSDEYRWAMALVYGGCISITTRIFYDINGSASDPLFDPKYSYYVWLVLGLIAAL
LSNLLFNPKGRSVGYLMIETWQGFPKFKAIFKARFFGAFYDAVLGARLRDFYLLLLLTIP
FIIAMHEISAYCGHPSNLLVEGLVILGFQGFLLKCAKLGW

>PROKKA_01114 Ribonucleoside-diphosphate reductase 1 subunit alpha

MITVVKRNGRIEPLDITKIQKYTKDATDNLEGVSQSELEVDARLQFRDKITTEEIQQTLI
KTAVDKIDIDTPNWSFVASRLFLYDLYHKVSGFTGYRHLKEYFENAEKGRILKGFKEKF
DLEFLNNQIKPERDFQFNYLGIKTLYDRYLLKDANNHPIELPQHMFMSIAMFLAQNEQEP
NKIALEFYEVLSKFAMCATPTLANARTTKHQLSSCYIGSTPDNIEGIFDSYKEMALLSK
YGGGIGWDFSLVRSIGSYIDGHKNASAGTIPFLKIANDVAIAVDQLGTRKGAIAYYLEIW
HIDVMEFIDLKNSGDERRRAHDLFPALWVCDLFLKRVLEDAMWTLFDPYECKDLTELYG
QDFEKRYLEYEKDPKIIKEYINAKDLWKKILMNYFEAGLPFLAFKDNANRCNPNAHAGII
RSSNLCTEIFQNTAPNHYYMQIEYTDGTIEFFEEKQLVTTDSNITKCANKLTSTDILK GK
PIYIATKVAKDQGQTAVCNLASINLSKINTEEDIKRVVPIMVRLLDNVIDLNFYPNRKVKA
TNLQNR AIGLGVMEQAQMLAEHKIAWGSKEHLEKIDALMEQISYHAIDTSANLAKEKGVY
KDFENSEWSKGIFFIDKANNEALKLTEKGLFNHACDWQGLREKVKSNGMNRNGYLMIAIPT
SSISILVGTTQTIEPIYKKKWFEENLSGLIPVVPNLNVETWNFYTSAYDIDAKDLIKAA
AVRQKWIDQGSINVFLRIENASGKTLHDIIYTLAWKLGLKSTYYLRSESPSIDEKSVLDR
SVECFNCQ

>PROKKA_01115 UDP-N-acetyl-2-amino-2-deoxy-D-glucuronate oxidase

MLFAMIGSGGFIAPKHLQAIRDTGHFLDCSFDVHDSVGVLD EYFPQSEFFTNIEDFEKHL

EQSRAMGKEINYLSVCTPTHTHFDYIRFGLKYGMHVICEKPLVLDPGEIQELKDLEVKHQ
KRVFSLPLRLHCDTLALKEKIKSELDKNPSKVFDITLTYISVQGWYFSSWRADVNRSG
GLATQMGVNI FDTLLYLFGGVKDKVINKEEPCVGGILFLEHAKIRWFFSINPEHMGVAK
EKVYHKMILEGEEVNLTQSLDNLYIESYKQILAQGGFGLDEATASIKLAYELRNLSVSEP
NEDSHALCCKNKTDQ

>PROKKA_01116 hypothetical protein

MDIYALYIAIGLFTGILSGIFGIGGLIIVPIMLATGHSFEESIGISILQMALSSFVGSV
LNFKKKSLDFSLGLLIGAGGLIGASFSGFVLKIVSSKILMVIFALLVVYSMIQFVLKPKK
KDFIADTKRYPLQGLKFLIGALTGFFAITLGIGGGMLMVPLMHYFLGYDSKKCVALGLF
FILFSSISGAFSLMYHHIINKEVLLAGAIVGLGSVMGVSIGIKWIMGLLNEKMHKILILG
VYGLSLLIILYKLFF

>PROKKA_01117 Methylated-DNA--protein-cysteine methyltransferase

MVLYHYFYFKTPKSPFLEYLHLCANESHLRLDFDTTNFSHNTPMNTPLKLSVQALERYFL
GQLFEFDAPLDLIGTSFQKQVWSALMTIPYGKTKSYDEIAKLINSPKSCRAVGNNYNNP
ISLIVPCHRVVRKNGALGGYNGGIEVKKWLEFESKILNEQAKNPLIS

>PROKKA_01118 Tyrosine recombinase XerD

MKHPLEELKDPIENLLLWIGRFLRYKCTSLSNSQVKDQNKVFECNLNHNACSSSHLEKI
CKKARNAGLLGINTYALPLLKFYEYAQKLSLKSLSIDEVMLAEFLSIYTGGLSLATKKN
YRIALLGLFSYIDKQNQDENEKSYIYNITLKNISGVNQSAGNKLPHLNNEELEKFLESI
DKIEMSAKVRARNRLLIKIIIVFTGMRSNEALQLKIKDFTLENGCYTILIKGKGDKYRAVM
LKAFHIESLLKEWLIERELYPVKNDDLFCNQKGSALTQAYLYKQVERIINFAGLRREKNG
AHMLRHSFATLLYQKRHDLILVQEALGHASLNTSRIYTHFDKERLKEAASIWEEN

>PROKKA_01119 hypothetical protein

MAIRFGVIFIDNNIDNFFNDRRKTQQLIRSILERKSGLDIFIRFIEATASYTNLERKRKL
SFEECAVFYTDADFVRIEREITSFLVFQNDRSVTLNNDPSATFNLKNYLLVLAKKFYCSV
CYCENINTQKIEWISTNHSDFGEILQEFSDITSLSPNELPQFLADLKDKIIEFNHKKII
NNNLPFGIIFIVSDPENLIDINDKNKKLESCFHYYENSHLLYFPIITESYLILDNLKSYF
IFQNEKPFNLKQHLLLELKKELGFEPGIFYCENASTHKIELVSNEHKDFKGVLLLEFLQ
ITSLSPNELPQFLANFKFSKSSNGIDKAEFSLSSLSLENMTPIDNLLSTLKCRIKNTLKR
YSHWGYSKIPQEKRITQIKKQVSEECKVDPFFHSHWRVSSEQNRKVNKIADDLKSGKIVGE
KIIDNVFDINASYCFITPEDLKNLKERLFIQQDIDITHVINQKKGEAFDHILINDIKYNL
FDNTFHFIFDVDNSSLTIKVPRKDLENYELPNTKKDIFNKLINIRTIDENLILKNNKEI
TNDLNPLLGLQKTQEFPLSDAITKAINQKEKGIALDYALINDIKYNLLDNTFHFIFDVG
SLLKESSQLIIEVPREALDLENVDRLIENTLSPYETHSVYSESSLVYHLNQGSYIIDLHL
IDD

>PROKKA_01120 Aspartate aminotransferase

MLYSSKIQSLSESTTIAISTLAKELKSQGDILSFSAGEPDFDTPQAIKDAAIKALNDGF
TKYTPVAGIPELLKAIAFKLKKENNDYELSEILVSNQAKQSLFNAIQALIGEGDEVIIP
VPFWVTYPPELVKYSGGVSQFIQTDEKSHFKITPKQLKDALSPKTKMLILTTSPNPTGMLY
SKAELEVLGEVLKDTKVWVLSDEIYEKLVYKGEFVSCAAVSEEMKKRTITINGLSKSVAM
TGWRMGYAASKDKKLKLMNNLQSQCTSNINSITQMASIVALEGLVDKEIETMRQAFERR
CDLAHAKINAIGGLNALKPDGAFFYLFIHIGSLCGGDSMRFCHELLEKEGVALVPGKAFL
EGYVRLSFACSEEQIEKGIERIAARFVKS

>PROKKA_01121 Helicobacter outer membrane protein

MKKFVVFKTLCLSVVLGNSLLAAEGSTEVOQKQLEKPKEYKVVKGEQNAWYLGISYQVGQA
SQSVKNPPKSSEFNYPKFPVGKTDYLAVMQGLGLTVGYKQFFGEKRWFGARYYGFMGYH
AVFGANALTSNNGGVCLEHQPATKVGTMGNLSDMFTYGVGIDTLYNVINKEDASFGFFF
GAQIAGNSWGNTTGAFLETSPYKHTSYSLDPAIFQFLNLGIRTHIGHRQEFDFGVKIP
TINVYFNFHGNLSFTYRRQYSLYVGGRNF

>PROKKA_01122 hypothetical protein

MIFNFDYNAMKKRISEISLKSSSVLNELEKAFLLYHLGQGIQSFETLKINSKQAFRERNY
DVWYISLYNMYNIPLFYGYSDENKKLEKYHEERISIDLNESFYELPFYKREQLKYLKDI
GTILDNTLIKAYQLKEKALRDLEIWSSSDSSFSFNNNQIKAYGIFKKTLKYFHFLLLNG
NQEKKFFQOMTEIFFSMAIFQIQEKRRDNNTKIPITLKSEQIYCILKYFDNKILMQKLNQ
YFQKTNIIFKTERDIDLIGIFKNISSQFVNIDIFETEF SRLFKNFLVLSAWIELDQNAFD
AIIIEICQEKIDEDLLRNSYDSMGYFITQWKNFKMEIKTEIKFSILDRILFSFIRKLTEN
FSGYLIILESSPRCMQNLLFILQQYNIELDLIQQALINTLIKTIMELPNDTQIFISNYLI
CDLFPITKNNDGVNQNVKKFLLSIWEKNQNRKTIQEDEVYLLLTYNMHRIGILNSDQYQK

IFLELKNKYMNREIAKKFNNEQPIHEQELLEQAMQENALDRMLALLKDCENSFKKE

>PROKKA_01123 NAD-dependent deacetylase

MNSQQYKEFIETSIDNIRKAHNENYLSIFAGAGISAESKLPKWGDLINELQKCLYGDTHK
NEDYLVLAEFYFNQFGESFYQTLKSLIPDSAKKNDLHLEIVKLNKLNITTNWDNLFEQ
AINEEGRFFNIKSDKDIGSSTGFAGFIKMHGSLDENNIVFKEQDYLEYSNNFPLIENYI
KGVFSTDTVILLGYSLSQNVKQIISWVNSHSHSVKPIYFIKTAKEFDRIEFEFYKNKNI
HILYTQELFEKKDIMRNC

>PROKKA_01124 N-6 DNA Methylase

MPIALEESEIQNLDEAVNNTNFKNIWKVIKALRSHDPSLVDEATFKEKVKIFGSDDSNR
NDEKTLFDAILLQDLADAMYNVMPTKLGDRNYWENFTKKTGNIARTLNRLKMI FDKNPE
FFHGFLDSLRENIHQNIKEDEALDMITSHIITKPIFDALFGDNIQNPIAKALDKMVQKLS
TLGLQGETKDLKNLYESVKTEATHAKSQKSQQELIKNLYNTFFKEAFKKQSEKLGIVYTP
IEVVDIFILRATNGILKKHFNTDFNDQSITIFDPFTGTGSFIARLLSKENALISDEALKEK
FQKNLFAFDIVLLSYIALINITQAAQNRDSSLKNFKNIALTDSLDYLEEKTNGALPLY
EDLKENKDIKDTLANQNIRVIIIGNPPYSAGAKSENDNNQNL SHPKLERLVYEKYGKNSTS
RSVGQTTTRDTLIQSIRMASDVVKDKGVLGFVVNGGFIDSKSADGFRKCVAKEFSHLYVLN
LRGNQRTSGEVSKKEGGKIFDSGSRATIAIVFFVKDKSVSDNTIDYYDIGDYLKREEKLH
RLANFLNLDAIPFEKITPNDKGDWINQRNDDFEKLIPLKRDKKLKFDTIFDLNSNGVVS
GRDPWVYNSSPNALMQSVQKCIDTYNADLKRFNARFREAFKQRTKGVKSGDLYKHLNDKE
ITTDKTKIAWTDGLKNHLIKNKNLQESSEKRVRLSLYRPFNKQRLYWDKDWTRNQRFSK
IFPDKDAQNVVINTGVNGKDFSALVSDFISDCSLISPNQAYPLYYDDLGNRHYAISGY
ALNLFRRHYGDNSIAEEEEIFYIYAILHHKGYLEKYKNSLTKEAPRIALSDDFKELSALG
KELAKLHNLYESGEMHTSVEYNLLESAAIEGYDVEKMTKKGDCILYNNNIAITKIPKA
FEYVINGKSAIDWVIERY SITMDKDSLIENNPNDYAGGKYVFELLCRVIKLSEKSVDLIE
KISEKRFE

>PROKKA_01125 type I restriction enzyme EcoKI subunit R

MQEISAYELIKEKLHAIPNLRHKGSLFEKISKQFLQEHD SAN EYESIDLWSDWELRGKEC
DRGIDIVIQTTSKEYIAVQCKFHQNSISLNDISTFLTQLQSGVGEVRFKKGIIISTSNLT
SEALKAIEQIRSTGMGIDIDEITEEDFIYSRIDWEKFDPTKTEDEILLCDKKRPRPHQTE
AIEKTKEYFSSPKNTRGKLIMACGTGKTYTSLKIMEALDPKITLFLAPSIALLSQTFREY
AKEKSEPFYASVVCSSDKVGKSKDEDNDDINFSELPLKPSTRLEDILSVRKAQKENKRF
IIFSTYQSALRIKEAQEAGLNEIDLICDEAHRTVGAMYSSNERDDKNAFTLCHSDKNIK
AKKRLYMTATPKVYSESSKAKAKESDNVIYSMDDAEIFGEEIYTLNFSKAIALDLLTDYK
VII LAVRKENLSGVTNSVNKKISQLKAEGTKLDKKLINNEFVCKIVGTHKGLAQDLIVL
DDENKEDHNLQNQYDTAPSQRAINFCKSINTSKNIKDSFETIMCEYDEELKKKSFKNLKI
SIDHIDGTMNCKDRLEKLEKLEKLEFQNPNTCKVLSNARCLSEGVDVPALDSIVFFDGKSAMV
DIIQAVGRVMRKAKRKKRGYIIFAYRFRRE

>PROKKA_01126 Anaerobic glycerol-3-phosphate dehydrogenase subunit C

MNENVNENIFEEVGDA CVKCAKCVPGCTIYRIHKDEATSPRGFLDMLRNAQNKLQLDTN
LKHLLCTCFLCTACVETCPFHLPIDTLIEKAREKIAQKHGIAWYKKSYSLLKNRKKMDR
VFSTAHLAPCIFIQVGDLSLEPRVVFVKGLFKRFKKSALPPLNQKSFLQKHEMKPLENPI
QKVAIFIGCLSNIHYQQVGESLLYILEKLN IQAII PNQECCSAPAYFTGDKDTTLFLVKK
NIEWFESYLDGVDAIIVPEATCASMLINDYYKVFLGEKDKDLVYKRLEKITPKIYLASVF
LEKHTPLKHLLEKIPKGGKEVITYHNPCHAKKTLNAHKEVRNLLNSHYEIKEMLDNCCGF
GGITMQTEKAEFSLKVGLLRakeIMDTKARILSAECGACHMQLNNAKSLDDPNTPPFLH
PLELIAKALKSAE

>PROKKA_01127 Oxygen-independent coproporphyrinogen-III oxidase

MQTIDFEKFSQYSKPGPRYTSYPTAVEFKENFNEESLKTAFNHDHLKNPMPLSLYTHLP
FCRSACYFCACSVIYTSLEEKVRYISYLKELALLKNAMDTNREVAQFHYGGGTPTFFS
PPQLDEITQSIQEVFPNFSKDIEMSCEIDPRHFTKEHMQTLFDRGFNRLSFGVQDFDFEV
QKAIHRIQPFEMVQESVKLARDYGIKSINFDLIYGLPNQTKESFLKTLEWVLKLNPDRLA
VFNYAHVPWVKKTRKIDETLLPSPRDKLEILESLSIFLEKANYQMIGMDHFAKSDNELY
LALQKAELRRNFQGYTTKKFTQTIGIGVTSIGEGGDYTTQNYKDLHHYEKALDLGHLPE
RGVALSQEDVLRKEVIMQMSNLKLDYSKIEEKFSVDFKVHFKEKELEKLPYEEAGLLSF
NPKGFEMTRTGGMLVRNMAMEFDAYLRGGEKHFSKTL

>PROKKA_01128 hypothetical protein

MEKLPKKRVSKTKSQKLIHSLTTQKNRAFLKKISANEMLLELEKGAFKKNEAYFISDEED
KNYVLVPDNVISLLAENARKAFEARLRAELERDIIITQAPIDFEDVREVSLQLENLRQKD

GNLPNINTLNFVKQIKKEHPNLFFNFDNMFKQPPFNENNFFENFDNSDEENF

>PROKKA_01129 Chorismate synthase

MNTLGRFLRLTTFGESHGDMIGGVLDGMPSGIKIDYALLENEMKRRQGGRNVFITPRKED
DKVEITSGVFEDFSTGTPIGFLIHNQRARSKDYDNVKNLFRPSHADFTYFHKYGIRDFRG
GGRSSARESARVAAGAFAMLLREIGIVCESGIIKIGGIEAKNYDFNHALKSEIFALDK
EQEEAQKTAIQNAIKNHDSIGGVALIRARSVKRNQKLPIGLGQGLYAKLDAKIAEAMMGL
NGVKAIVEIGKGVESLLKGSEYNDLMDQKGFLSNRSGGVLGGMNGEEIIVKVHFKPTPS
IFQPQQTIDINGNECECLLKGRHDPICAIIRGSVVCESLLSLVLADMVLLNLTSKIEYLKT
IYNEN

>PROKKA_01130 Ribonuclease 3

MMKNKRSQNSPYITPNNPYLTLEKALGYSFKDKRLLEQALTHKSCKLALNNERLEFLGDA
VLGLVIGELLYHKFYQYDEGKLSKLRAISVSAHGFTKLAKAIALQDYLRVSSSEEISNGR
EKPSILSSAFEALMAGVYLEAGLAKVRKIIQNLLNRAYKRLDLEHLFMDYKTALQELTQA
QFCVIPTYQLLQEKGPDDHKEFEMALYIQDKMYATAKGKSKKEAEQQCAYYALQKLKEAK

>PROKKA_01131 Ribonuclease HI

MQEIEIFCDGSSSLGNPGPGGYAAILRYKDKEKIIISGGENFTTNNRMELRALNEALKILKR
PCHITLYSDSQYVCQAINVWLANWQKKNFSKVKNVDLWKEFLKVSKGHLIVAVWIKGHNG
HAENERCDSLAKLEAQKRVT

>PROKKA_01132 tetratricopeptide repeat protein

MQHFNFLYKDSLSIALFTFIIALVILLEQARAYFTRKKNKKFLQKFAQNQNAYASSEN
DELLKHAKISNLMFLARAYSKADVEMSIILKGLLNRLPKDEEKIAVLDLLAKNYFSVGY
LQKTKDVTVEILRFSPRNVEALLKLLHAYELEKDYSKALETLECLEEELEVPEIETIKNYL
YLMHLIENKEEAAKILHVSASLDLKKIALNYLKSYPDENLFWQEIDATERLENVLDLLWD
MNIPAFILEKHALLQDIARSQGLLLDNKFCQVFEEVLRLALLNSPMKARLTFEYRCKHCK
QIFPFESHRCPCVYQLAFMDMVLKISKESYGGGLNARN

>PROKKA_01133 peptidyl-prolyl cis-trans isomerase SurA

MRKIFSIVLKALLFFGIVYAEPESEKVEALEGRKQESSLDKKIRQELKNKELKNKELKNKE
EKKEIKAKRKPRAEVHHGDSKNPAQKITPPKIKKSAGTQNGQVQNNAPKLEEKETTPQI
LEKNKGASPSQFNSIFGNPNDAAGNTLEDKVVGGISLLVNGSPITLYQIQEEQEKSKVS
KAQARDRLIAERIKNQEIIEERLKIHVDDDKLDQEMAMMAQQQGMDDLHFKQMLMAEGHYKL
YRDQLKEHLEMQELLRNILLTNVDTSSETKMREYYNKHKEQFSIPTETVRYTSTNQED
LERAMADPNLEIPGVSKANEKIEMKTLNPQIAQVFISHEQGSFTPMNGGGGQFITFYIK
EKKGKNEVSFSQAKQFIAQKLVEESKDKILEEHFEKLRVKSRIVMIRE

>PROKKA_01134 Aspartyl/glutamyl-tRNA (Asn/Gln) amidotransferase subunit B

MPFEAVIGLEVHVQLNTKTKIFCSCSTSFGESPNSTCPVCLGLPGALPVLNKEVVKKA
QLGTAIEANINQYSIFARKNYFYPDLPKAYQISQFEVPIVSDGKLEIDTKEGAKIVRIER
AHMEEDAGKNIHEGSYSLVDLNRCTPLLEIVSKPDMRNSEEAIAYLKKLHAIVRFIGIS
DANMQEGNFRCDANVSIRPKGDEKLYTRVEIKNLNSFRFIAKAEIEYERQSAAWESGRY
HEEVVQETRLFDTHKGITLSMRNKEESADYRYFKDPDLYPVFIDEKLLKESQKINELPSA
KKTRYMKDFNLKEDDANLLVSDPLLAEYFESMLHLGVKAKTSVTWLCVELLGRKAEVTL
ENCGISAHMLGALAKRIDEKGISGKSAKDVLKLEEKGGDVDALIEQMGLSQVNDTEAI
VKVIEEMLKNNADKVLEYKSGDKDLGFFVVGQAMKNLKGANPSVNVNAILKEKLD

>PROKKA_01135 hypothetical protein

MEKIRVQEGLAYSVYIRSNFSKVAHFASGYLQTKLSTQAKSVALVKKIVKEFIEKGMTQQ
ELDDAKKFLLGSEPLRNETISSRLNTTYNYFYLGPLNFNQTLNLIQKMSLKEINDFIK
AHTINDLTFAIVSNKKKDK

>PROKKA_01136 Insulinase (Peptidase family M16)

MGLQASALTHQEIINQAKVPVIYEENHLLPMGFIHLAFRGGGSLSDKNQLGLAKLFAQVLN
EGTKELGAVGFAQLLEQKAISLNVDSTEDLQITLEFLKEYEDEAIMRLKELLKSPNFTQ
NALEKVKTQMLAALLQKESDFDYLAKLTLKQELFANTPLANAALGTKESLQIKILEDLQ
QFDKVFELNKLVVVLGGDLKIDQTLKRLDNALNFLPQGKAYEPEYFETSDKKSEKVLYKD
TEQAFVYFGAPFKIKDLKQDLAKSKVMMFVLGGGVWLSFNGKNQGSRLSL

>PROKKA_01137 Aminodeoxyfutalosine synthase

MRIDRKEILDLMKNAPLKELGQKALEIKQRLHPENLTTFIVDRNINYTNICFVDCKFCFA
KRTLKEKDAYVLSYEEIDQKIEELLAIGGTQILFQGGVHPQLKIDYYENLVSHIAQKFPT
ITIHFSAVEIDYISKISKLSLKEVLERLKNAGLSSIPGAGAEILSDRVRDVIAPKKLSS
DRWIEVHRMAHLGCIKSTATMMFGSVDNEEDVVEHLQVRDLQDETGGFRAFILWSFQPD
NTPLKEEIPSIKKASSNRYLRYLACSRIFLDNIQNIQSSWVTQGS MIGQLALLFGANDLG

SVMMEENVVKAAGTSFCMNEAEMIELIEDIGSVAAKRNTAYEILKRYPAKAKV

>PROKKA_01138 Outer membrane protein assembly factor BamA precursor

MTPPKAESQESQRNEAPKNEVQRNEAQKETPQSSQTPKEMKVKSISYIGLSYMSDMLANE
IVKIRVGDIVDSKKIDTAVLALFNQGYFKDVYATFEGGILEFHFDEKARIAGVEIKGYGT
EKEKDGLKSQMGIKKGDTFDEQKLEHAKTALKTALEGQGYGSSVVEVRTEKVSEGALLIV
FDVNRGDSIYIKQSIYEGSAKLKRRMIESLSANKQRDFMGWMMGLNDGKLRLDQLEYDSM
RIQDVYMRRGYLDAAHISSPFLKTDFTSTHDAKLHYKVKEGIQYRISDILIEIDNPVPLKT
LEKALKVVRKDVFNIEHLRADAQILKTEIADKGYAFVVKPDLDKDEKNGLVKVIYRIEV
GDMVYINDVIIISGNQRTSDRIIRRELLLGPKDKYNLTCLRNSENSLRRLGFFSKVKIEEK
RVNSSLMDLLVSVEEGRTGQLQFGLGYGSYGGLMLNGSVSERNLFGTGQSMSLYANIATG
GGRSYPGMPKGAGRMFAGNLSLTNPRIFDWSYSSINLYADYRISYQYIQQGGGFGVNIG
RMLGNRTHVSLGYNLNVTKLLGFSSPLYNRYSSVNEVVAPRQCSTPASVIINRLSGGKT
PLQPESCSSPGAITTSPEIKGIWDRDYHTPTSSFTLDVSYDNTDDYYFPRNGIIFNSYA
TMSGLPSSGTLNSWNGLGGNVRNTKVYKGFAAHYHHLQKYLIDLIARFKTQGGYIFRYNT
DDYLPLNSTFYMGVTTVRGFRNGSVTPKDEFGLWLGGDGIFTASTELSYGVLKAAKMRL
AWFFDFGFLTFKTPTRGSFFYNAPFTTANFKDYGVIAGGERATWRASTGLQIEWISPMG
PLVLIFPIAFFNQWGDGNGKKCKGLCFNPNMDDYTQHFEFSMGRF

>PROKKA_01139 Aminodeoxyfutalosine synthase

MLQCLSLKRGKGNLNMDFLEKVLNDQVTESKELVRLYDYDLYTLGEAADRMQRNMHQKIVY
FNVNRHLNPSNICADACKFCAFSARHNPNPYEMSLEEILEKVKNSYNKGIKEVHIVSAH
NPNYSYEWYLKVFEETIKQEMPNLHLKAMTAAEVHFLSVKFNKPFELVLEDMLKAGVDSMP
GGGAEIFDEEIRRICKKGKVGSSRWLEIHAYWHKLKMSNATMLFGHIENKIHRIDHMLR
IKKIQSPKNQVENKEGGFNAFIPLLYQKENNYLNVEKSPSAIEILKTIAISRILLNNIPH
IKAYWATLGLNLALVAQEFGANLDGTIEIESIQSAAGAKSRHGLEKEDLIFKIKDAGFV
AVERDSLYNFIQKF

>PROKKA_01140 Ferritin

MLSKDIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIF
LNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFL
QWYVAEQHEEEVLFDKILDKIELIGNENHGLYLADQYVKGIASRKS

>PROKKA_01141 Phosphoserine phosphatase

MQKLAVFDFDSTLVNAETIESLARAWGVFDEVKTITSQAMNGETDFHKSLILRVSKLKNM
PLKLAKEVCESLPLFEGAFDLISTLKEKNYKVVCFSGGFDLATNHYRDLHLDAAFSNTL
IVENNALNGLVTGHMMFHSKGEMLLALQRLNISKDRTLTVVGDGANDLSMFKHAHIKIA
FNAKEVLKQHATHCINEPDLALIKPLI

>PROKKA_01142 Glycosyltransferase family 10 (fucosyltransferase)

MFQPLLDAFIESASIEKMVSKSPPLKIAVANWWGDEEVKEFKKSFYLYFILSQRYTITLHQ
NPNKPSDLVFSNPLGSARKILSYQNTKRVFYTGNEVPNLNLFDAIGFDELDFRDYLR
MPLYYDRLHHKAESVNDTAPYKIKDNSLYTLKKPSHHFKENHPNLCAVVNNESEDPLKRG
FASFVASNPANPMRNFYDALNSIEPVTGGGSVKNTLGYKVKNKNEFLSQYKFNLCFENS
QGYGYVTEKILDSYFSHTIPIYWGPSVAKDFNPKSFVNVHDFKNFDEAIDYVRYLPRTO
TLI

>PROKKA_01143 hypothetical protein

MLYENPLNTLDGKAYFYQNLFSFKKILDFFKTILENDTIYHDNPFIFYRDLHEPLATIDDL
RVNY

>PROKKA_01144 hypothetical protein

MRVNYDDL RVNYDDL RVNYDDL RVNYDDL RVNYDXLRVNYERLLQNASPLLELSQNTTFK
IYRKIIYHKSPLLLHAIRRWVKK

>PROKKA_01145 Uracil DNA glycosylase superfamily protein

MGETYTNTNLTTTPQNKPLNKQVHEGIENCNLCKRHQHSKPVIGIFNPTSKLAFITLTPML
DSQLHFLNNLKAAMLESIIQKVFNYPKDCSILSLLKCDNSNLNEEEINACLFHFLTQWL
DNSASKVIVVFGEILPKRLNLSKEESFGRIVSLKTKHFLSTHALEDMLKNPTLKKEALA
HFKIALQFLNQS

>PROKKA_01146 Aspartate ammonia-lyase

MRIEHDFIGQMEISDEVYGIQTLRASENFFITNDKLCSYPVFIKSFAQVKKAAALANAQ
LGLIDEKCLKIAICHACDLLVDGKYHDQFIVDMIQGGAGTSTNMNMNEVIANLALEYMGHQ
KGEYQFCHPNHDVNRSQSTNDAYPSALKIAIYERLSNLVAPMKALRDAFAQKAKEFAHVI
KMGRITQLQDAVPMTLGQEFETYALMVDRDIEQVLDARNWVRELNLGGTAIGTGINSHPDY
RSLIEKKIQEVTGRPFVMANNLIEATQSTGAYVQVSGVLKRIAVKLSKVCNDLRLSSGP

RAGLNEINLPKMQPGSSIMPGKVNVPVIEVVNQVCFVIGNDLSVALAAEGGQLQLNVFE
PVIAYKLFHSFVILGRAIETLTTCVEGITANEKICHDYVFNSIGIVTALNPHIGYEKSA
MIAKEALKSDRSIYDIALEKKILTKEQLDDIFKPENMLSPHAFKKHKD
>PROKKA_01147 UDP-N-acetylglucosamine 1-carboxyvinyltransferase
MDFLEIVGQVPLKGGVEISGAKNSALPILAATLLSQQEVKIKSLPQVVDIKAMALLLQNL
GALEWLDPNTLQISAKSLHTEATYDLVRKMRASILVLGPLLARFKECLVSLPGGCAIG
ARPVDLHLKAMQQQLGAEIKIEQGYIHAKAPKGLKGNLFDKISVTGTENALMAASLAKG
ITRIINAAKEPEIAQLCAFLQSGGVEIEGVGSSELKIRGVESDALNLKDIQIIPDRIEAG
TYLCVGAITNSQLKINRIIPNHLQAITDKLIEIGFSLDIQENSIEIHPAKKRQAFEITTK
EYPGFPTDMQAQFMALATQCLGTSVIEETLFENRFMHASELQRLGANISLKTNVATISGS
TELTGSDVMATDLRASSALILAALVAKGVS RVHRIYHLDRGYERLEDKINALGAKVARLK
EK
>PROKKA_01148 hypothetical protein
MGNLTYAYMYLILFVCLLPVLLMGLAWRFFRPPLKQNI PNKSLSLEDLSEQIKNLKSAP
ALEKLKNSFNERFKICPKDKETLWLETIQNLVASEFFELEDAINFGQELNANPSHAQKI
ANATGLALKNKKEKG
>PROKKA_01149 UTP--glucose-1-phosphate uridylyltransferase
MIKKCLFPAAGYGTRFLPITKTIPKEMPLIVDKPLIQYAVEEAMEAGCEVMAIVTGRNKR
SLEDYFDTSYEIEHQIQGTSKENALKSIRNIIEKCCFSYVRQKQMKGLGHAILTGEALIG
SEPFVILADDLCISRDPYSLVKQMISLYQKYQCSIVAIEEVALEEVSKYGVIKGEWLEE
GVYEIKDMVEKPSQEDAPS NLAVIGRYILTPDIFEILSETKPGKNNEIQITDALLTQAKR
KRIIAYQFKGKRYDCGSVEGYIEASNAYYKKRL
>PROKKA_01150 Soluble lytic murein transglycosylase precursor
MRFLILFFIGALGVGFSQTEFNLDLEKKPAGIVRDYLLWRYISDKKTSLENAKKAYELT
QNKNNALQKAMQEKGS DNAEKNPDVKLPEDIYCKQITLESMLTTDTFQASCIAIALKSK
IRDFDKIPIETLKLPLQIKIKEAYPVLYEELEILQNKHVSASL FKANAQVFSALFNHLSYE
KKLQIFEKRIPIKELNRLLEDYPAFNRLIYQVILDPKLDHFKDALTKSNATHSNAQTFF
ILGINEILRKPKSKALKYFERSEAVIKDDDFSKDRAIFWQYLVSKKKKTLEHLSQSPALN
LYSLYASRKLKTTPSYRIISHIQNLSQEDPPFNTYDPFSWQIFKEKTL SLKDESAFNAML
KSLYYEKSAPELTYLLSQRNKDKIYYYYLSPYEGII EWQNVDEKAMAYAIARQESFLLPAV
ISRSFALGFMQIMPFNVGPFAKRLGMDNIDLNDMFNPNI ALKLGNYLNLHKKEFNHPLF
VAYAYNAGPGFLRRWLESSKRFEKNHFEPWLSMELMPYSETRLYGFRVMLNYLIYQEIF
GNFIPIDGFLEQTLNSKDKP
>PROKKA_01151 YGGT family protein
MIFSTLLNAIAVILSSLLTIYMWVVIYSLLG FVQPNPNPIMQILARLCEPVFYFLRSK
FKLVFNGLDFAPLVVVIVLKFLDLTLIQWL FALAKSL
>PROKKA_01152 Glutamate--tRNA ligase 1
MLRFAPSPTGDMHIGNLRAAIFNYIVAKQQHKPFLIRIEDTDKERNIEGKDQEILEILKL
MGISWDKLVYQSHNIDYHREMAEKLLKENKAFYCYASAEFLEREKEKAKNEKRPFYRLDE
WATLEKDKNHAPVVR LKAPNHAVSFNDTIKKEVKFEPDELDSFVLLRQDKSPTYNFACAC
DLLLYEISLIIRGEDHVSNTPKQILIQQALGSNNPIVY AHLPIILDEVSGKKMSKRDEAS
SVKWLLNQGLFPVAIANYLITIGNKVPKEVFSLDEAIEWFDLENLSSSPA HFNLYLKLHL
NHEHLKLLDDEKLLELTSIKDKNLLGLLRLFI EECGTLELLEKIKISLFLEPKDIVKTYEN
EDFKERCLALFNALKSMDFQAYKDFESFKKEAMRLSQLKGKDFFKPLRILLTGNSHGVEL
PLIFPYIQSHYQEV LRLKA
>PROKKA_01153 Putative NAD(P)H nitroreductase YfkO
MDREQVIALQHQRFAAKKYDPNRRISQKDWEALVEVGR LAPSSIGLEPWKMLLLKNERMK
EDLKPMAGALSSLEGASHFVIYLARKGV TYDSDYVKKVMHEVKKRDYDTNSRFAQI IKN
FQESDMKLN SERSLFDWASKQTYIQMANMMMAAAMLGIDSCPIEGYDQEKVEAYLEEKGY
LNMAEFGVSVMASFGYRNQEITPKTRWKTEVIYEVIE
>PROKKA_01154 2-(hydroxymethyl)glutarate dehydrogenase
MKIGWIGLGAMGTPMAARLCDANLEVS VYNRTESKAASLKEKGVAVYTNPIDLA AKVDLI
FIMLSDKTAIDAVLV PKFWEQMSSEKIVVN MSTIAPLESLSLEEIAQKHQVTYLEAPVSGS
VGAAGALLILAAGDEEVISKLKPVLVHLGSQTFYLGKVGQGAGAKLSINSLLAQMGVA
YSEALLLAKHLGVDAELFLQIIIGSGMNSPLFQAKKGMWLQDSYPAAFS LKMLKDIRLA
NNEAGEAIIQLPFLSKAQELY SQAESGLGALDMVAVYHYLEKGEH
>PROKKA_01155 hypothetical protein
MNAIQWSKKWILGETDSFVSNEVIVKGLDFNKVVQH LRDASCWEKYKNSGNVHMHNQDN

TILKDRTRFCFETFGFLVEAQVEEFELKDTILRLAWRGWNEAKGDEYLEVYHAWIVEKLD
NDRVRILTQESQSGVPAKALAKSVPNAMLNGHQAWLDGLVAYS

>PROKKA_01156 CCA-adding enzyme

MFELEKGLKELFDIYEKSPHELYLVGGCVRDCLMGITPKDYDLTSNALVNESKELLLKRH
FRVLETGIKHGTITALKNHQSYEITTFRIEKGHIKHKPKELVFSTRLTDDLKRRDFSMN
AIAYSPTKGLIDPFKGRSAIENQMIECVGGARLRFEDALRILRALRFSATLGFKIAPST
KRAVFACKDLLKHLKERLQSELNKLIMGKNAYEVAKEYQEILELVTQEKIENLGFLKNA
PFNLELRLLGFFKHQKSLENLRYPKKTCVLFSKAKECHKAFLENIHNKTELKFLKKNYDLE
PFNLALDFYALENPKHALKIKRLLKEIFDSNEPFFKKEHLALKGNALQSLDYQHQQIGEIL
NACLNLVIAHPKNNALEWLIKWVKDHYLPNDAINHSPIGRKN

>PROKKA_01157 7-cyano-7-deazaguanine synthase

MGQEICVIGFSGGQDSTTLAVWAKKRFFKVCVGFVDYVQKHSVELECAQKIASLLQLPYE
IISLDFLENITRSALFKNSNDLIGHSHAQNKDLPNSFVFNNAIFITLLHSAQKLGASN
IALGVSQADFSGYPDCKEDFIKSIEHALNLGSNTAIKILTPLMFLNKAQEFQMAKDLGVL
DLVIKETHTCYQGERKILHAYGYGCDECPACQLRKKGYEEFQTKALLQ

>PROKKA_01158 Helicobacter outer membrane protein

MKKALLLTLSLSLSFWLHAERNGFYLGINFAGESYIQGQGSIGEKASAENALNEAINNAK
NSLFPEQNTKAIRDAQNALNEVKDSTKIANRFAGNGSGGLFNELSFYKYFLGKKRIIG
FRHSLFFSYQLGGVGSVPGSGLIVFLPYGFNTDLLINWTNDKRASQKYVERRVKGLSIFY
KDMTGRITLDANTLKKASRHVFRKSSGLVIGMELGGSTWFASNNLTFFNQAKSHTIFQLQG
KFGVRWNNDEYDIDRYGDENYLGSSVELGVKVPFAKVNYYSDNYGDKLDYKRVVSVYLN
YTYNFKKNKH

>PROKKA_01159 hypothetical protein

MHQNNKTFLPNQSAHLSKIVLFLNTGLLAYLLSACGANVPPIEEVIVKDPKETKAQEVARE
ENAIQQENATIDARTTPLINRFTNYAYGSLNGFYNSVDNLNSPMQNGMYGGYMPYYM
PYGFMPYGSGLMPYGPYGYGAPGYFPYAFY

>PROKKA_01160 hypothetical protein

MRIVRNLFVSVFVAYSSAFATDLETETKSEKSGKKFYKFHKNHGSETETKNDKKLYDFT
KNSGLEVDLEKSPNLKSHKSDKKFYQLAKNNIAEGVSMPIVNFNKALSFGPYFERTK
SKKTQYMDGGLMMHIRF

>PROKKA_01161 hypothetical protein

MACNLQARFYSVYKDN TASFYLQASAETTLEFAQKLSEILPFSLDFSFLSLKEITEPLDE
NLFQTTSLSKPLFMNAKEHQDFLDKNSSLYANALGFVNAAFKGAI IHSPKELIDCLTQL
KGMLKTQDFIPIFTSRGALSLSLKKPSPSVIFSDDLSSVLTCTKLPLEDAQYLASLEKPCI
KASLKS VFKDIFKNDEIIAQLPYDPILNLLCHILQDEGIEFVFTHE SRCEALLHYEALF
KTPKRLITPTKKFVLENHLSTFPFKDELEFLSATPNSIVLYFSFKRPTRLLLHANGSLKT
LLSVSFDNFQIFNLLKQDEKASRMLKNYAAKFPDFYARIAELSKHNLGGANVLDFFRILG
FVLGYSED FHSHSVISLAKECLRPKGPRIDYKILKDDSLKMAFNFSKIMHSAMSFRLAGV
ENEILSLGILDSLAEFLGNFIWDNAQNFSVQEV TIAGDFFGEKVFLDLFVQYFPKTLALK
THAFLDYE

>PROKKA_01162 Hydrogenase 2 maturation protease

MSQKILILGIGNILFGDEGIGVHLAHLKRNFSFFPSVDIIDGGTMAQQLIPLITSYEKV
LILDCVSAKGVEIGSVYAFDFKDAPKEITWAGSAHEVEMLHTLRLTEFLGDLPKTFVVG
VPFVIGSETTFKLSSGILNALETALKAIETQLNAWGVQMQRTHIALECIAELSYKGF

>PROKKA_01163 Quinone-reactive Ni/Fe-hydrogenase B-type cytochrome subunit

MDKMNKVVLHKEYSGFVRFFHWVRALSIFAL IATGFYIAYPFLQPNSSFYKGVYLLQAYV
RSFHV MFGLLISALIFRTYLFFTKESLMERRSFSQLLSPKAWIDQMKAYFLISGKPHTK
GAYNPIQLVAYSTLIVLIVLMSLSGVVLYYNVYHAGLGAFLGSAFKWFEALCGGLANVRF
IHHLATWGFI LFVPVHVYMVFFHSIRYESSGADSMINGYGYTKE

>PROKKA_01164 Quinone-reactive Ni/Fe-hydrogenase large chain

MSKKIVVDPITRIEGLHRIEVIVDDDNVITDAFSSSTLFRGLETIIKGRDPRDAGFIAQR
ICGVCTYSHYKAGITAVENALGITPPLNAQLVRSMLNMALLFHDHVVFHYTLHGLDWCDI
MSALKADPIQAAKLSFKYSPYPINTGAPELKAVQKRLSDFAKSGSLGPFPSNGYYGHKTYR
LSPEQN LIVLSHYLKLLEIQREAAKMTAIFGAKQPHPQSLTVGGVTSVMDILDPTRLAEW
KSKFEVVANFINHAYYPDLVMAGEMFANEPSVIKGCGLRNFIAYEEVLLGRDKYLLSSGV
VLDGDISKLHPIDESLIKEEVTHSWYQYEDTKEVQLHPYDGQTNPHYTGLKDGESVG IEN
KIIPAKVLDTKDKYSWIKSPRYDSKPM EVGPLSSVVVGLAAKNPYVTEVATKFLKDTKLP
LEALFSTLGRTAARCI EAKTIADNGLLAFDALVENLKS DQSTCAPYHIDKNQEYKGRYIG

QVPRGMLSHWVRIKNGVVENYQAVVPSTWNAGPRDSQNQRGAYEMSLIGTKIADLTQPLE
IIRTIHSFDPICIACSVHVMDFKGQSLNEFKVEPNFAKF

>PROKKA_01165 Quinone-reactive Ni/Fe-hydrogenase small chain precursor

MFYDEKKTYYQKIEERLEIVSSFNAHNEHKNLQDEFKGAGISRRDLLKWAGMMSTTLALPA
SFAPLTLKAVEVANRLPVIWLHMAECTGCSESLRSADPTIDSIIIFDYINLEYHETIMVA
SGFQAEKSLHDAIEKHKNYYILMVEGGIPQGTETYFLTQGPNAETGAEECRKAAQYAAAI
AIGTCSSFGGVQAAYPNPSNAQPLHKIIDKPVINVPGCCPSEKNIVGNVLYYLMFGALPK
LDAYNRPSWAYGNRIHDLCERRGHFDAGEFVEHFGDENAKRGFCLYKMGCCKGPYTFNNCS
KLRFNSTSWPIGAGHGCIGCSEPNFWDTMSPFEEPLANRSIKTAFDGLGADKVADKVG
TLLSATAIGIVAHVLLSKAIKNKE

>PROKKA_01166 Modulator of drug activity B

MKKVLIINGAKAFGSSGGKLNELTHHAKKTLESGLLEVDTTIVDKGYNHSQEVEKIVSA
DATIWQMPGWWMGEPWIVKKYIDEVFTLGHGKLYASDGRSSQNPTKNYKGGMLMQGKKY
LSLTWNAPIEAFNDPSEFFEGVGVDVYVHLHLKAFQFLGLSALPTFICNDVVKNPQVEQY
LNSLTTHLRQAFGK

>PROKKA_01167 hypothetical protein

MKATQSTINDFFALTGTIFSIPIVYQRNYTWEEENCEKLLQDIINISQNKTHFMGSITYI
LHLIDDEKSLRKLEFVIIDGQQRITTLMLLLKAIETKIPNEEIKGEIDNLLNLSGQKLR
LKPIKSDEKAFDLVMQNRSEIQQGVSHIRNNYKFFTKELDNYISKGYRIEEIYGAFLRLK
IVAIGLELGEDDPQVVFESINATGVQLKGLDLIRNYLMMGENSDNQNHLYNTYWVPLEDW
LGEKDLNDFIKTYLRIYFEDRLKEGEREVYALKAHHRDNFSDNIRGLMSDMREYGRIFY
IFLDRDHYCLHRGDPQQLANLRLRVKDLVKIKFGVAKPFVLRCARDFEEGKLDYENFCEI
LQILISYFVRRSVCGBPDAVLNKKVLYSLYKRLNVSADALKRYLGKSVGQAAFPNDDKIK
EAFAVRNAYAAANQVCKFILLEIEKLSNAEPPKEENLEVEHFYPKTPTQEWDRMVGDIYFT
EQDYLNNGFNLTLSGQNQRLGNKPYEAKIALMEEYSSLHLNDYFINNTHSWGIEEVKARS
GYLADQFCQVGLFKDLLKEYRTREINKTLDLDDLTNHNLSQSVKLPNHQRKIARNAKELVST
VIDYLLENAREAFESYTDDESQKYIYWSKAKAEARDRDGTLVVPFEKYGFYFVSNASYQT
TGSNLRDLILGCELNPRDFIV

>PROKKA_01168 Putative beta-lactamase HcpD precursor

MLKKSLLLLLVFLILQLSGAEENNQAPKNTPPELNPANAKGAPNPNTQITPKNDNSNLLDK
LGSPENAOQTELSAGIDLAKKGDYQGAFLFSQSCDNGNAAGCFAVGAMYANGVGIQTNRL
KAARYYEMGCSGGDATAACANLAQMYENKKNADSNDKENALQLYAVACQGGDMLACNNLGW
MFANGSGVQKDYYKAISYYKFSCENGNDMGCYNLGLMSNVNNIYGIDKAKLSQVDLNYLA
CNAGDMMGCANLGWIYANGDLGAPLNNHYAAKYFQMACDGGILGSCNNLGVLVYQKGLGVP
QDDQRALDLFSYACDNGFESSCRNYGNFKEHLLRVNPNYGRFLMPYNSYEIP

>PROKKA_01169 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase

MMNKFKNFVSNYQQSSCYREPLGFGIARVDIAPISKILCATYPVLNWKEENLGSYAVFC
NSLSKEKILKESASERVVEIDESFVLKALDFYTPFLNEAYSNKMAHKNIQVVLELLKALE
ENRLKNSDGESLYRLVILYEDKPCESVESAYMKLLALSGLKAPLRSNLLEGIFNQLSNAA
WSGNKPYELEWLRMNEVALKMRGHFPSIDFIDKFPRYLMQLIPEFDNIRLLDSSKTRFGA
YLTGGYTQMPGASYVNFNAGAMGVCMEGRISSSVVVGAGTDIGGGASVLGVLSGGNNN
PISIGKNCLLGANSVTGISLGDGCIVDAGVAILAGSVIEIEENEFFKKLLEVNSALEKHAN
NLYKGKELSGKNGVHFRSNSQNGKLIAFRSVKKIELNQNLH

>PROKKA_01170 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase

MLENRVKTQIFIGGVAIGGDAPISTQSMFTFSKTADIESTKNQIDRLKLAGADLVRVAVS
NEKDALALKELKKVSPPLIADIHFHYKFALIAAQSVDAIRINPGNIGSKDKIKAVVDAC
KEKNIPIRIGVNAGSLEKQFDQKYGPTPKGMVESALYNKLLLEDLDFTNFKISLKASDVM
RTIEAYRMLRPLVIYPFHLGVTEAGNLFSSSIKSAMALGGLLMEGIGDTMRVSITGELEN
EIKVARAILRHSGRLEKGINWISCPTCGRIEANLVDMAKVEKRLSHIKTPLDISVMGCV
VNALGEAKHADMAIAFGNHSGLIIEKGKVIHKLAEKDLFETFVIEVENLAKEREKSLKD

>PROKKA_01171 LL-diaminopimelate aminotransferase

MTFEPYPFERLRALFKEITPKKKGLDLGIGEPREFETPKFIQDALKNHAHSLNIYPKSAFE
EGLRAAQRGFFKRRFKIELKENELISTLGSREVLFNFPSFVLFDYQNPTIAYPNPFYQIY
EGAAKFIKAKSLMPLTKENDFTPSLNEKELQEVLDLILNSPNNPTGRALSLEELIAWVK
LALKHDFILINDECYSEIYENTPPPSLLEACMLAGNEAFKNVLVIHSLSKRSSAPGLRS
FIAGDSRILEKYKAFRTYLYGTSANAIQKASEAAWLDDRHAEFFRNIYANNLKLARKIFK
NMLIYPYSFYVYLPVQNGENFAKTLYQNEGIITLPALYLGRNRIGADCVRALVYDTPLL
EKPLEIIETIYRENHA

>PROKKA_01172 UDP-N-acetylmuramate--L-alanine ligase
MLETPKVLLKLNQDCKIHFIGIGGIGISGLAKYLKAQGATISGSDIAISPSVKYLKALGV
EINIPHDPKAINNQDVIIHSAIIKEDNTEIQRAKELEIPILSRKDALYSILKDKRVFSVC
GAHGKSSITAMLSAIFPFFGAIIGAHSKEFGSNVRESANDSLVFEADESDSSFLFSNPYA
AIVPNTPEHLEHYDHDLERFFFAYEYFLDHAQKRVIYKEDPFLKNYSKDAIVLEKKDIY
NIQYILKDGEPTYTSFELKDLGAFLVWGLGEHNATNASLAILSLDELDELGEIRNNLLNFK
GIKKRFDILQKNALILIDDYAHHPTAISATLKSARIYADLLDTQEKIIVIWQAHKYSRLM
DNLEEFKKCFLEHCDRLIILPVYSASEVKRDIDLKAHFKHYNPTFIDRVKKGDFLELLV
NDNVVETIEKGFVIGFGAGDITYQLRGEM

>PROKKA_01173 hypothetical protein
MGAAVVLFLTLVLLFLVLRDFGLASPKQKLVAFLVIGIIGASISVYTYQQNQQNQQEIAL
QRAFLRGETLLCKGIKVNNTFNLVSGTSLFLGKKQTPMKDVLVDLSDCQTLQKDPLIQP

>PROKKA_01174 Endonuclease MutS2
MMXNNNTTPPKPLEESLDLKEFIALFKTFFAKERDTIALENDLKQTFITYLNEVDAIGLPAP
KSVKESDLIVIKLTKLGTLLHLDEIFEIVKRLHYIVVLQNAFKTFTHLKFHERLNAIVLPP
FFNDLIALLDDEGKIKQGANATLDALNESLNRLKKESAKIIHHYARSKELAPYLVDTQSH
LKHGYECLLLKSGFSSAIKGVVLERSANGYFYLLPESAQKIAQKIAQIDNEIDCCIVEMC
QTLSHSLQKHLLFLKFLFKEFDLDSLQARLNFAKAYNLEFVMPSTQKKNQKGRKLFTPH
FKRAKALKFEI

>PROKKA_01175 Endonuclease MutS2
MLAVTGVNAGGKTMLLKSLLSAAFLSKHLIPMKINAHHSIIPYFKEIHAIINDPQNSANN
ISTFAGRMKQFSALLSKENMLLGVDIEIELGTDADASSLYKTLLEKLLKQNNQIIITTHH
KRLSVLMAENKEVELLAALYDEEKERPTYTFLKGVIGKSYAFETALRYGVPHFLIEKAKA
FYGEDKEKLNVLINSSALERELKQKNEHLENALKEQEDLKNAWLLMEKQKEIFHNKKL
ELEKSYQQALNLIKSEVASKDTSSMHKEIHKASEILNKHKTNQEIIPQTITSFQANEKARY
KNESVLIAQILDKGYWIE TELGMRLKAHGSLLKKIQKPPKNKFKPPKTTIPKPKKASLH
LDRGQRSEEAALDLDLDAFLNDALLGGFEEVLICHGKSGILEKFVKEFLKNHPKVVSFSD
APINLGGSGVKIVKL

>PROKKA_01176 Inorganic pyrophosphatase
MNLDQLEVSHDADSLCVVIEISKHSNIKYELDKESGALMVDRLVLYGAQNYPANYGFVPNT
LGSDGDPVDALVLSDAFQAGSVVKARLVGVNMEDESGVDEKLLALPIDKIDPTHSYVK
DIDDLKSHTLDKIKHFFETYKDLEPNKWVKVKGFFENKESAIKVLEKAIKAYQG

>PROKKA_01177 Glycosyltransferase family 25 (LPS biosynthesis protein)
MTQVYIIISLKESQRRDLTEKLVSNEKFKGRCVFQIFDAISPKHEDFEKFIQELYDSSS
LLKSDWFHSDYCYQELLPQEFGCYLSHYLLWKECVKTNQPVVILEDDVALESNFMQALEDC
CLKSPFDFVRLYGHYWGGHKTNLCTLPITYTETEEAEAPIENHAEAEASMEKTPIENYEVT
PPPNPTQDAQQDCIETQQDPKELSEPCKIAPQKTSFNQVVFVKIKRKLNRFIGNILAR
TEVYKNLVSKYDDLTKKYDDLTDLDLNKNIAEKYDDLTDLDLNKNIAEKYDELTGKYESLLA
KETNIKETFWESRTDSEKEALFLEHFYLTSVYVATTAGYYLTPKGAKTFIEATERFKIIE
PVDMMFNNPTYHDVANFTYLPCPISLNKHAFNSTIQNAKKPDISLKPPKKSYPFDNLFYHK
FNARKCLKAFNKYSKQYAPLKTPEV

>PROKKA_01178 Glycosyltransferase family 25 (LPS biosynthesis protein)
MIGVYIIISLKESQRRDLTEKLVSNEKFKGRCVFQIFDAISPKHQDFEKFVQELYDAQS
LLQSDWYHSYVGAGLTLPELGCYLSHYLLWKECVKSNQPVVILEDDVALESNFMQALEDC
LKSPFDFVRLYGCYWYHETKFHVLPKEFVFPFDYPFKNNPILEKFKKFFDVSRLNLS
THKVIHYILKKIQKSYATHEKEAFFLEHFYLTSVYVASTAGYYLTPKGAKTFIEATERF
KIIEPVDMMFINNPTYHDVANFTYLPCPVSLNEHSLDSTIQKPQKKSLSYPLPPKKSTFK
NLFYYSFNAKKRLNAFHKYSKQYAPLKTPEV

>PROKKA_01179 Adenylate kinase
MKQLFLIIGAPGSGKTTDAELIAKNNSETIAHFSTGDLLRAESAKKTERGLLIEKFTSQG
ELVPLEIVVETILAAIKSSEKGIILIDGYPRSVEQMQUALDKELNAQNEVILKSVIEVKVS
ENTAKERVLRGSRGADDNEKVFNRMRVFLDPLVAIQNFYKAKHLHKVINGERSIEEIVN
EMQKYILSFAN

>PROKKA_01180 Aspartate--tRNA ligase
MRSHFCTEISEKDVGKTIKVGWCNTYRDHGGVVIFIDLRDKSGLVQLVCDPSSKAYEKAL
EVRSEFVLVAKGKARLRGPGLNPKLKTGKIEIVLEELVIENKSATPPIEIGNKSVNEDL
RLKYRYLDRSLNAYEIFKLRSEVALIARNALAQKGFLEIETPILSKTTPEGARDYLVPS
RVHEGEFFALPQSPQLFKQLLMVGGMDRYFQIARCFRDEDLRADRQPEFTQIDAEMSFC

ENDVMGVVEDLLQAIKFAIGHDISKPFKRMPIKEAMENYGSDKPDRLRFELPLIEVGDCFR
DSSNAIFSNAIQDPKNKRIKALNVKGADAIFSRVLKELEEFVRQFGAQGLAYLQIKEDG
IKGPLVKFLSEKGLKNILEKTGAKTGDIVFFGAGDKKIVLDYMGRLRLKVAETLDLIDKD
ALNFLWVVNPFMFTEKTENGYHAAHPFTMPKNIECEDIEEIEAHAYDVVLNGVELGGGSI
RIHKEEMQKKVF EKINIHEEEVQKKFGFLLEALKFGAPPHGGFAIGFDRILMLTKSHSI
RDVIAFPKTQKASCLLTDAPSPINEEQRLRELHRLRK

>PROKKA_01181 Chemotaxis protein CheV

MVRDIDKTTSLHLNNEAQFLCFRLDGEKDAQLYGLNIFKIREIIHYDGEVTEILGGSDGM
MLGFLSVRGESIPLDVVKRWLHYNASDKSRNLTECSVKDDHNLVIVCHFSNHSIALKVLK
IERIIHKNWTEISAGDKQGINEEGKLSAITRFDEERVVQILDVEKMISDVFP SLKDLDDL
TLRCIEAIQSQKLILIAEDSL SALKTLEKIVQTLELRYLAFPNGRELLDYLYEKEHYQQV
GVVITDLEMPNISGFVLTIKADHRTEHLPVIINSSMSSDSNRQLAQSL EADGFVVKSN
ILEIHEMLKKTLS

>PROKKA_01182 DNA ligase

MIKSQKEYLERIAYLNTLSHHYYNLDDPIVSDAVYDELYQELKAYEEKNPNSIQANSPTQ
KVGAITTNSFNKNSHLMRMWSLDDVFNQSELQAWLQRILKAYPSASFVCS PKLDGVS LNL
LYQHGLVKVATTTRNGLEGE LVTNNAKHIANIPHFIA YDGEIEIRGEV IISKDFDALNQ
ERLNANEPLFANPRNAASGSLRQLDSEITKKRKLQFIPWGVGKHS LNFLSFKECLDFIVS
LGFSAIQYLSLNKNHQEIEDNYHTLIREREGFFALLDGMVIVVDELDIQKELGYTQKSPK
FACAYKFPAL EKHTKIIGVINQVGRSGAITPVALLEPIE IAGAMINRATLHNYSEIEKKN
IMLNDRVVVIRSGDVIPKIIKPLESYRDSQHKIERPKVCPIC SHELLCEEIFTYCQNLN
CPARLKESLIHFASKDALNIQGLGDKVIEQLFEEKLIVNALDLYALKLEDLMRLDKFKIK
KAQNLLDAIQKSKNPPLWRLINALGIEHIGKGASKTLARYGLNVLEKSEAEFLEMEGFGV
EMARSLVNFYASNQEFIRSLFELLNPRNSDMAEEKQKSSSVFN NKTIVLTGTLSKPRQEY
AQMLENLGAKISSSVSAKTDFLIAGENPGSKLALAKKHGVSVLNEEEELLKRLKELD

>PROKKA_01183 hypothetical protein

MEHHKVHTTVQALQAKRKRLLTELAEEAEIKVSSERRSSFNISLSPSLLAEIEEIEYEE
KTSKERRINHSVLLSPSFMKVDEYMKEKGFPNRSLLFEKALEFYMLKHP

>PROKKA_01184 ABC transporter ATP-binding protein YxdL

MISNISIHPKTMFKNALNIQDFSFKSHTSTAIIGTNGAGKSTLINTILGIRSDYHFKAQN
NNIPYYDNVIPQRKQLGVVSNLFNYPPLNANDLFKFYQFFHK NCTPNLF EKNLLNKTYE
HLSDGQKQRLKIDLALSHHPQLIIMDEPETSLEQNALIRLSNLISLRNTQQLT SIIATHD
PIVLDSCWVLLKNGNIAQYKPLNSILKSVAKTFNFKEKPTTKDLLALLKDI

>PROKKA_01185 ABC-2 type transporter

MGAILSILKLEIKSYLTNTSALFWFTIYPILMLLLLIFVFSKNTTEIFYFN NIIGLMGLL
IISSAIFGLTQAITSSRSNIFLFYMLSPATFKQITLALIASRLIVVILYAFIFIVLSFY
ALNIITILNFKALILGFVS VFSSALFCFCLAIFVARIFQNEQSILGFCNIINLYALMSCN
VFVPLEYLPNIGQLFIKTSIFYYLNQLLIKAFQGIDTILVLATSTFFIIGGIILFLLSAN
RMLLTPKERMR

>PROKKA_01186 Vacuolating cytotoxin autotransporter precursor

MWGGNFKSEGLIFNLNNPTHQTIINVTGASTIMSYNNQALIHFN TQKQAYTLIDAKRM
VGYDNQTILGGSLS DYLKLYTLIDFNGKRMQLKGDSL SYDNKPVS IKDGGVTSFKDNQ
GRMVYSSILYDKVQVTVSEKPMDIHAPSLEYIYKIYIQGSAGLNAIKSAGNNSIMWLNELF
VAKGGNPLFAPYYLQDNPTKHIVTLMKDITSALGMLS KPNLKNNSTDALQLNTYTQMSR
LAKLSNFASF DSTDFSERLSSLKNQRFADAIPNAMDVILKYSQ RDKLKNLWATGVGGVS
FVENG TGTLYG VNVGYDRFIKGV IAGGYAAYGYSGFYERITNSKSDNVDVGLYARAFIKK
SELTFSVNETWGANKTQISSNDTLLSMINQSYKYSTWTTNAKVNYGYDFMFKNKSII LKP
QIGLRYYYIGMTGLDGVM DNALYNQFKANADPSKKS VLTIDLAFENRHYFNTNSYFYAIG
GIGRDL LVRSMGDKLVRFIGNNTLSYRKGELYNTFASIT TGGEXRLFKSFYANAGVGARF
GLDYKMINITGNIGMRLAF

>PROKKA_01187 Putative vacuolating cytotoxin

MTYRNSKTDLKNEHFSKNRSFKGVKKKIAKKYKIKNSPLTIYSLKTHSNSSLSINKKIFL
GLGFVSALS AEDYNSSVYWLNSVNNENSHKSYVVSPLRTWAGGNRNFTQNYNNSQLYIGT
KNASATPNNSSVWFGEKGYIGFITGVFKARDIFITGAVGSGNEWKTGGGAILVFESSNEL
TTNGAYFQNNRAGTQNSWINLISNHSVHLKNTDFGNQTPNGGFNAMGRKIT YNGGIVNGG
NFGFDNVD SNGTTTISGVTFNNNGVLT YKGGNGIGGSITFTNSNINHYKLNLTNSVTFN
NSTLGSMPNGNTNTIGNAYILNANNITFNNLTFNGGW FVFNRS DAHVNFQGITTINNPTS
PFVNMTGKV NINANAIFNIQNYTPSIGSAYTLFSMKNGS IAYNDVDNLWNIIRLKNTQAT

KDNSKNATSNNNHTYHVTYNLGGTLYNFRQIFSPNSIVLQSIYYGANNLYYTNSVNIHD
NVFNLKDIKDNRDDTIFYLNLNTWNYTNARFAQTYGGKNSALVFNAMTPWANGSI PKSN
STXRFGGYEGVNWGKTGYITGTFTADRVYITGNMMTGNGAQ TGGGATLNFVGATEINIAG
ADFKNLKTTSQNSYMTFIALGSGSGSGKINVSQSDFYDWTGGGYDFTGNSVFDSDVFNKA
YYKFQGPENSYTFKNTNFLAGNFKFQ GKTTIEKSVLDDASYTFDGTNNTFNEDKFNGGSF
NFNHAEQTDAFNNNSFNNGSFSFNAKQVDFNHNSFNNGVFNFNNTPKVSFTDDTFNVNNQ
FKINGAQTTFTFNKGVVFNMQGLLSLSVGTTYQLLNAKSLDYKDNNALYQMLRWISGEN
PSGTLVNKDQSAPSNAKIYNVHFTDNGLTYIYIKENFNNGITLTRLCTLGYTHCVNIHNDA
FNLKNVNNSASDTV FYLNGMTTWKIAGTG VFTQDYSGANSVLVFNQ TTPFLDGANPTSNS
VVGFGKTS GAEWGLVGYIQGTFKASQIDITGTIRSGNGAKTGGGATLVFNAQKRLNIANA
HLNSDKAGLQNSWMNFIVNNGNLNATNANFSNQTPHGGFNLKANNITFNNGSVNGGNFG
VDNASANGNAVIKNVNFSDNGTLIYKGGENSAGNSLTLENNTFNSYNINAKVQNLI FNNN
SFNGGSYSFNDTKNTTFKGTNTLINSDPFSRLQGSIAIENNSIFNIERDLTDNTTYTLLS
GNNIKYNNEIVADNAFSKNLWNLIHYGGEQGTLLRADNNTFFVQFTQSNQKQFVFEETFN
PGSITYKYFTIHSSPFHTDADSKDIWNQVRKQFDFIPGKTPVCVGV CYIAPYKNQDLIGS
SAFAWSLNF GATVVG TLLLSAQEKANNNGGSIWFGKNNLLYLHGNFNATNIFLTNNFNV
GNPNAGGGATINFNADETLNADGLNYTNFQTVAMGLQTSASQHSWANFNSRLSMEIKNSN
FRDFTWGGFNFN SGRITFENTTFSGWTNINGATESGSSYVMVANTDLMFSNSILGGGIR
YDLKANNIVFNNSQMVIDVSKNVNQSSLNGNVTFNNSRLSVKPNAAINIGGDQTQTALEN
ASSLSFYNNSVANFN GTTAFNGVSYLNLNPNAQVSFNQTNFNANVTFYGIPLFGKTPDF
GNSARLINFKGNTNFNQATLNLRAKNIHINFQGASTFENNSTMNLAESSQASFN TLSVEG
ETNFNLNGSSLLNFNGDSVFNAPVNFYANHSQISFTHLATFNADASF DLSNNSTLNFQSV
LLNGALNLLGNGTNALSINASGNFSFGTQGI LNL SNVNLFDAENKPLVYNILQAQNIQGL
MGNNGYEKIRFYGIQIDKADYSFNNGVHSWSFTNPLNTTETITETLHNNRLKVQISQNGV
SNNKMFNLAPSLYDYQKNPYNESSNSYNYTSDKAGTY YLTSNIKGFNQNEIPGTYNAQN
QPLQALHIYNQAITKQDLSVIASLGKEFLPKMANLIASGALDNLNLNSPNSFETLFGIFE
KYGITLTQANWKSLLGIINNFSNTANYHFSQGNLVVGAIKEGQTNTNSV VWFGGDGYKNP
CTVGDN TCQMFRQTNLGQLLHSSAPYLG YINANFRAKNIYITGTIGSGNAWGS GGSANVS
FESATNLVLNQAKIDAQGTDKIFSYLGKEGIEKLFGEKGLGNALSNI IYEESLKNNAI PK
DLASMI PKDFGSKTLLSLLSPTEVNLLGVSAFKNAIMEILNSKTVGDVFGENGLLNALD
PIKRKEIDQMLLEQIQAHSSGFEKFIVKTLGIENVENFINTWYQKQSLSSFANNFVPGGL
NQALDKIGSSTD AKDLQNFLDKTTFGDILNQMINQAPLINKLISWLGPQDLSVLVNIALN
SITNPSKELLGAISGMGQKVLNDLLGEGVVNKIMSNQVLGQMINKI IADKGLGGVYHQGL
GSILPKSLQDELKQLGFGSLLGSRRLHNLWQKGNFNFVAKNHVFNVNSSFSNATGGELNF
VAGKSII FNGKNTINF TQYQKLSFISKDFSNISLDTLNATNGLTLNAPKNDISVQKGQI
CVNVLCMGEKKAHSSSTTAPTNETLEVNANNFAFLGAIKANGLVDFSKVLQNTTIGTLD
LGP NATFKANHLIVNNAFNNSNRYRADISGNLNVVKGAAALSTNENGLNVGGEFQERRVIN
L

>PROKKA_01188 Helicobacter outer membrane protein

MKKIVFILALWVGLLGAFEPKKSHIYFGAMVGLAPIKITPKPASDSSSYTAFLWGAKGGYQ
FAFFKALALRGEFSYLMAIKPTALHTINTSLSLNIDVLSDFYTYKKYSFGVYGGLGIGY
FYQSNHLMGNSSFMGYNGLINVGLGSTIAHCHRIELGAKIPFSKTRNSFKNSYFLESVF
IHAAYSYAF

>PROKKA_01189 Multidrug resistance protein MdtC

MYKTAINRPITTLMFALAIVFFGVMGFKKLSVALFPKIDLPTVVVTTTYPGASAEIIESK
VTDKIEEAVMGIDGIKKVTSTSSKNVSIVVIEFELEKPNEEALNDVMNKISSVRFDDSN
KKPSINKFDTDSQAIISLFVSSSSVPATTLNDYAKNTIKPMLQKINGVGGVQLNGFRERQ
IRIYADPTLMNKYNLTADLFSTLKAENVEIDGGRIVNSQRELSILINANSYSVADVEKI
QVGNHVRGLDIAKIEIGLEEDNTFASF KDKPGVILEIQKIAGANEIEIVDRVYEALKRIQ
AISP NYEIRPFLDTTSYIRTSIEDVKFDLILGAILAVLVVFAFLRNGTITLVS AISIPIS
IMGT FALI QWMGFSLNMLTMVALTLAIGI I IDDAIVVIENIHKKLEMGM SKRKASYEGVR
EIGFALVAISAMLLSVFVPIGNMKGII GRFFQSFGITVALAIALSYVVVVTIIPMVSSV
VNPRHSRFYVWSEPF FKALESRYTKLLQWVLN HKLIIFIAVVLV FVGSLFVASKLGM EFM
LKEDRGRFLVWLKAKPGVSIDYMTQKSKIFQKAIEKHAEVEFTTLQVG YGTTQNPFRAKI
FVQLKPLKERKKEHQLGQFELMRVLRKELRSLPEAKGLDTINLSEVTLIGGGGDSSPFQT
FVFSHSQEAVDKSVENLKKFLLESPELKGKVESYHTSTSESQPQLQLKILRQNANKYGV
AQTIGSVVSSAFSGTSQASVFKEGKEYDMIIRVPDDKRVSVEDIKRLQVHNKYDKLMFL
DALVEITETKSPSSISRYNRQRSVTVLAEPNRNAGVSLGEILTQVSKNTKEWLVEGAN YR

FTGEADNAKESNGEFLVALATAFVLIYMILAALYESILEPFIIMVTMPLSFSGAFFALGL
VHQPLSMFSMIGLILLIGMVGKNATLLIDVANEERKKGLNIQEAILFAGKTRLRPILMTT
IAMVCGMLPLALASGDGAAMKSPIGIAMSGGLMISMVLSLLIVPVFYRLLAPIDDKIKRF
YQNQKALE

>PROKKA_01190 Cobalt-zinc-cadmium resistance protein CzcB
MSFLSMEAGERVYAI FNVKAMQDSKLTLDSTGIVDSIKVTEGSVVKKGDVLLLLLYNQDKQ
AQSDSTEQQILIFAKKQYQRYSKIGGAVDKNTLEGYEFTYRRLES DYAYSIAVLNKTILRA
PFDGVIASKNIQVGEGVSANNTVLLRLVSHARKLVIEFDSKYINAVKVGDTYTYSIDGDS
NQHEIKITKIYPTVDENTRKVSAEALLSKPMAVGLFGDGFIQTK

>PROKKA_01191 Outer membrane efflux protein
MNAIIRYASLWGLCAALTLAQTPSKTPDEIKQILNNYSHKNLKLIDPPTSSLEATPSFLS
SPKETATTINQEIAKYHEKSDKAALGLYELLKGATTNLSLQAQELSVKQAMKNHTIAKAM
FLPTLNASYNFKNEARDTPEYKHYNTQQQLQAQVTLNVFNGFSDVNNVKEKSATYRSNVAN
LEYSRQSVYLQVVQQYYEYFNNLARMIALQKKLEQIKTDIKRVTKLYDKGLTTIDDLQSL
KAQGNLSEYDILDIQFALEQNRLTLEYLTNLSVKNLKKTTIDAPNLQLRERQDLVSLREQ
ISALRYQNQQLNYPKIDVFDWSLFWIQKPAYATGRFGNFYPGQQNTAGVTATLNIFDDI
GLSLQKQSIMLGQLANEKNLAYKKLEQEKDEQLYRKSLDIARAKIESSKASLDAANLSFA
NIKRKYDANLVDFTTYLRGLTTRFDAEVAYNLALNNYEVQKANYIFNSGHKIDDYVH

>PROKKA_01192 Uroporphyrinogen decarboxylase
MIFIDACFRKETPYTPIWMMRQAGRYLSEYQESRKKAGSFLELCKNSDLATEVTLQPVEI
LGVDAAILFSDIILVVPLEMGLNLEFIPKKGPHFLETITDLKSVESLKV GAYKQLNYVYDT
ISQTRQKLSKEKALIGFCGSPWTLATYMIIEGEGSKSYAKSKKMLYSEPEVLKALLEKLSL
ELIEYLSLQIQAGVNAVMI FDSWASALEKEAYLKFSWDYLLKKISKELKKRYAHIPVILFP
KGIGAYLDSIDGEFDFVFGVDWGTPLEAAKKILGDKYVLQGNLEPTRLYDKNALEEGVEKI
LKVMGNQGHIFNLGHGMLPDLPRENAKYLVQLVHAKTRR

>PROKKA_01193 hypothetical protein
MIFRFFLILSLLKGVLLAKKDLNFFKPLEPTKKYFGSFKIGYLYQHAETTKRSPIRPKNR
PPILMDKIYHDASLGFQAGFVLKKKALLGGYLDAGMGDSYFMSAGLVAGVRLFRGWVIPK
IALGYQLQILGAKIDKYQFNIQSAVGSVGLFFNAAKNFGLSIEARGGIPFYFIQSKFSKA
FGTPRLNIYSVGITFTFYDFTRFLG

>PROKKA_01194 N-glycosylase/DNA lyase
MDSFEILKALKGLDLLKNAPAWWWPNALKFEALLGAVLTQNTKFEAVVKSLENLKNAFIL
ENDDETNLKKIAYIEFSKLAECVRPSGFYNQKAKRLIDL SKNILKDFQSFENFKQEATRE
WLLDQKGIGKESADAILCYVCAKEVMVVDKYSYFLKKL GIEIEDYDELQHFFFEKGVQEN
LNSALALYKNTISLAQLYARFHGKIVEFSKQKLELKL

>PROKKA_01195 Flagellin A
MAFQVNTNINAMNAHVQSALTQNALKTSLERLSSGLRINKAADDASGMTVADSLRSQASS
LGQAIANTNDGMGIIQVADKAMDEQLKILD TVKV KATQAAQDGQTTESRKAIQSDIVRLI
QGLDNIGNTTTTYNGQALLSGQFTNKEFQVGAYSNQSIKASIGSTTS DKIGQVRIATGALI
TASGDISLTFKQVDGVNDVTLESVKVSSSAGTGIGVLAEVINKNSNRTGVKAYASVITTS
DVAVQSGSLSNLTLNGIHLGNIADIKKNDS DGRLVAAINAVTSETGVEAYTDQKGRNLNR
SIDGRGIEIKTDSVGNGPSALTMVNGGQDLTKGSTNYGRLSLTRLDAKSINVVSASDSQH
LGFTAIGFGESQVAETTVNLRDVTGNFNANVKSASGANYNAVIASGNQSLGSGVTTLRGA
MVVIDIAESAMKMLDKVRS DLG SVQNQMISTVNNISITQVNVKAAESQIRDVDFAEESAN
FNKNNILAQSGSYAMSQANTVQQNILRLLT

>PROKKA_01196 hypothetical protein
MGYNKENLYQILLHANTNGYHKKKHSSKDFHFTL

>PROKKA_01197 hypothetical protein
MQTPMDTINKNIPLRTFILLYKSSPKCVVLASITVLFIGIIPSINILVMIRLIDIVVDLLQ
NHTHFEYSLLLPTLLLWGALLFLTHVFSGISSSLQTIIEAQFSINIITQLANKLTKVKNL
NFFENKDHTIKLNAIHNGLYIRPLNYVSNLFFNLQRIIGLVSLFGILFSISIYLPFIMIF
ATVPCIFISNHIAKKHSASIDKLQDKKESI QNYLYSGLDNQKNKDNLLFNFMNLFHHKFI
ENKELYINHVFVKIAQKNLTLTIIYADILTTILSIALFFLMVFIILSKSCGSNCWVYPSI

>PROKKA_01198 putative ABC transporter ATP-binding protein
MKTILISLYTLIKIYALVGRNASGKSTLINLLGFYTPNSGQIIINNKYSLQDLELNSYHQ
QMSAIFQDFSLYAGYSIDNLFMQNNPTREQ LKQKREMLKSF DENFQNC LNDYSNALFGT
QYNGIDFSLGQKQRIATIRAF LKPSHCVVLD EPSSAIDPIMEKEFLDFIFKQSQSKMTLI
ITHRMNSVKQADEIIVLDQGKLIEQGNFETLMKKQGLFYELYLKQQY

>PROKKA_01199 Methyl-accepting chemotaxis protein McpS
MFGNKQLQLQISQKDSEIAELKKEVNLYQSLNLCLHEGFVGIKNNKVFKSGNLASLNN
LEEQSVHFKEAESVNLQGVSYSLKSQNIQDGVQYFSLAKKTGCVGEYHKNDLFKTFCASL
KEGLENAQESMQYFHQETGLLLNAAKNGEAHSTEGGLTVNKTGQDIESLYEKMQNATSLA
DSLNQRSNEITQVISLIDDAEQTNLLALNAAIEAARAGEHGRGFAVVADEVKLAECTQ
KATKEIAVVVKSMQQEANDIQTNTHDINSIVSSIKGDVEELKSTVKNNMIVAQAQKYTIY
NVNNRVFCGLAKLDHVVFKNLYGMVFGINSFDITSHKNCRLGKWWYEGAGKENFSNTSG
YRALESHHASVHAEANDLVKAVQEDHITDLKYLEHKVHLMEDSAKHVRENIDKMFYEKQD
ELNKIIEKIQKGE

>PROKKA_01200 8-amino-7-oxononanoate synthase
MFSKSLEALHHAARYRKRELFDSLKDYASNDYLGSLVKKDLLQNAFNKLQSFGAHSKPA
SMLVNGYHPLHAELENQLADLLGFESALLVSGSGLGNLALIDTLVKNALLFMDEHYHAS
GIFSTKTKPSQVVFVFSHNNAKDLQQLFNAPKNKLKFIIEGVYSMDASVAPYDFYEIIQ
ETPNAFLIVDEAHSFGTIGENLLGFLEYRYRIKEKDIIKLSTFSKALASYGACILAPLQV
IEFLTNRKSVIYTTALSLLDTALTALHLEYFIAQKQELKNELSRHQIIFETLGIRTPT
GFFTLFESNLALLNAHFLKEKGLVGAIRPPTVSKPLLRVSLSLKNSLEDTKELANTL
LDYFKIQSSFKSD

>PROKKA_01201 Penicillin-binding protein 1A
MLKKIFYGFIVFLIVMGLLAILIAQVWVTTDKDIAKIKDYRPGVASQILDRKGRLIANI
YDKEFRFYARFEEIIPRFIESLLAVEDTLFFEHGGINLDAIMRAMIRNAKSGRYTEGGST
LTQQLVKNMVLTREKTLTRKLKEAIIISIRIEKVLKSEEILERYLNQTFFGHGYGVKTAS
LGYFKKPLDKLTLKEITMLVALPRAPSFYDPTKNLEFSLSRANDILRRLYSLGWISSNEL
KGALNEVPPIVYNQTSTQNIAPYVVDVLKQLDQLDGLKTQGYTIKLTIDLDYQRLALES
RFGHQKILEKIAKEPKTNASNEDEDNLNASMIVTDTSTGKILALVGGIDYKKSANFRAT
QAKRQFGSAIKPFVYQIAFDNGYSTTSKIPDTARNFENGNYSKNSEQNHAWHPSNYSRKF
LGLVTLQEALSHSLNLTINLSDQLGFEKIYQSLSDMGFKNLPKDLSIVLGSFAISPIDA
AEKYSLFSNYGTMLKPLIESITNQNDVKTFTPMETKKITSKEQAFLTSLVLMNAVENG
TGNLARIKGLEIAGKTGTSNNNIDAWFIGFTPTLQSVIWFGRDDNTPISKGATGGVVSAP
VYSYFMRNILAIEPSLKRKFDVPKGLRKEIVDRIPYYSTPNSITPTPKRTGDSEERLLF

>PROKKA_01202 tumor necrosis factor alpha-inducing protein
MFKILKICVTLCIEIKIGDSVLEKSFLKSKQLFLCGLGVLMLQACTCPNTSQRNSFLQDV
PYWMLQNRSEYITQGVDSHIVDGKKTEEIEKIATKRATIRVAQNIVHKLKEAYLSKSNR
IKQKITNEMFVQMTQPIYDSLNVDRGLGIYINPNNEEVFALVRARGFDKDALSEGLHKMS
LDNQAVSILVAKVEEIFKDSVNYGDVKVPIAM

>PROKKA_01203 Disulfide bond formation protein B
MDKETRFYNLFSLAVLGILIFPVGLANFYFGYVLKDSPCIFCWAQRINMILIGAVALLV
RFGFKPKYIALLLLMASSGLYESFYHTGSHALEDVGQGFALAILGLHTQFWALFVFFSVV
ALLAVLLFFAPNTQPFKVRLLNTLQKSAFYVFFIVVGSNAVQAFFSTGPFYIGQSDPVR
FSWNLKESVWSMENWDHLKFPRSVLGRRDVGEPLKLSALPKDNDYERSPLEITKALEIGK
KEELFLKLNGAITDLSFNEDEAILITENQDLYLVGNLKTIIHSHMVLDSEYYSATVGSFVG
ADFNEDENIVIMGNKTSVEITPNKNANALKNFPYFLEGANSFDEVERSRLKTSRAKNYY
VSAARRGAKFTYLISTPNKRYKDLMIISMRNSDKQVHGEFLLELGNALKEKRKLDELVI
SALALKDNKLYAFSKEFNTLLVIDPIKEEILEVYGLPKEIKNISAGGFRDNELILVSYEN
DKNILYTLNF

>PROKKA_01204 hypothetical protein
MEFLGLILSLAAILIAFKKPEKENWAFGILMVVWLVELIIFIAHSSSVLPNMNL

>PROKKA_01205 hypothetical protein
MQAGIAAVAYCQFIYKEIHKRNIDTLVFVARDGYLLQKIFNILYPNSYKTTYVYAPRILK
KAVFLEVVEGESLEILRILEGEEVKKKQITTNQQAYVYLYSNFEHCRHLALKCLDNRYR
YLSSSNLEGNIAIVDTITFGYSSQELIQKVLKSEVFGCYVDLLRILNYDCVSFLPFSHPK
SVYFHNWDFMEFLITSPEYPILNVENGVPYQKDVSSCEEHRSKAYEKIVEGAVGYASYF
KESQISLGIYDVIEWVNFIDHPSIQDQEQFKQIYFLPDATHKNALPLFCNDVSLLSCL
KPSQSYGVLKRSLRTNKQERLFLKILSLIKKIYGKLKKK

>PROKKA_01206 haloacid dehalogenase-like hydrolase
MHKFLDGAKEILKYDVISDFIDFTLLRPFIKPTDLFLYIETKYNIKGFHQARILAEMQ
SREISKRQDITLDEIYHQIPKEFHSYKGEVIAATEKEVLIPNLEMLELYRFAKENNKRVI
VSDMYLPLEVLEDILISKGFDGYTNFYLSNHIMLTKHSKDLFKHVLKQENIHTQTILHIG
DNSWADDAMPKSLGIATLFRKSVLKQFEEIFPKYKTFSPSTSVTQSFILGSLCVFHKNLYS

KT

>PROKKA_01207 2-oxoglutarate-acceptor oxidoreductase subunit OorC
MEAQLRFTGVGGQGVLLAGEILAEAKIVSGGYGTKTSTYTSQVRGGPTKVDILLDKDEII
FPYAKEGEIDFMLSVAQISYNQFKSDIKKGGIVVIDPNLVTPTKEDEEKYQLYKIPIISI
AKDEVGNIITQSVVALAITVELTKCVEENVVLDTMLKKVPAKVAETNKKAFEIGKKHALE
ALKK

>PROKKA_01208 2-oxoglutarate oxidoreductase subunit KorB
MAFNIDEYLVRVDKIPTLWCWCGDGVILKSIIRTIDALGWKMDDVCLVSGIGCSGRMSSY
VNCNTVHTTHGRAVAYATGIKMANPSKHVIVVSGDGDGFAIGNHTMHACRRNIDLNFIL
VNNFIYGLTNSQTSPTTPNGMWTVTQAQWGNIDNQFDPCALTAAAGASFVARESVLDPQKL
EKVLKEGFSHKGFSSFFDVHSNCHINLGRKNKMGEASQMLKWME SRLVSKRQFEAMSPER
VDKFPTGVLKHDTRKEYCEAYQEIIIEKAQKQ

>PROKKA_01209 2-oxoglutarate oxidoreductase subunit KorA
MREIISDGNELVAKAAIEVGCRFFGGYPITPSSDIMHAMSVALPKCGGRFIQMEDEISGI
SVSLGASMSGTKSMTASSGPGISLKVEQIGYSFMAEIPLVIADVMSRGPSTGMPTRVAQG
DVNFLKHPIHGDFKAVALAPANLEEAYTETVRAFNLAEMLMTPVFLLMDETGVGHMYGKVQ
IPDLEEVQKMTINRKEFEVGDKKDYKPYGVAQDEPAVLNPFFKG YRYHISGLHHGPIGFPT
EDAKMGGDLIDRLFHKIESKQDIINENEEMDLEGAEIVIIAYGSVSLAVKEALKDYNKES
KQKVGFFRPKTLWPSPAKRLKEIGDKYEKILVIELNKGQYLEEIERAMQRKVHFLGQANG
RTISPKQIIAKLKEI

>PROKKA_01210 Periplasmic [Fe] hydrogenase large subunit
MAKMSAPDGVAVVWVNE DRCKGCDICVSVCPAGVLGMGIEKERV LGKVAKVAYPESCIGCV
QCELHCPDFAIYVADRDKDFKFAKVSKEAQERSEKVKANKYMLLEETILEGRGK

>PROKKA_01211 putative aminodeoxychorismate lyase
MVVVPQGS LKKVFFSLKDQGVDMNALDLLLRLMGMPKKGYIDMGD GALRKGD FLVRLIK
AKAAYKSATLIPGETRYFFTQILSETYQLETS DLNQAYESIAPRLNGAVIEDGVIWPD TY
HLPLGEDAFKIMQTLIGQSMKKHEALSQWLGYHHKEEF EKIIILASIVQKEAANVEEMP
LIASVIFNRLKKGMPLQMDGALNYQEF SHAKVTKERIKTDNTPYNTYKFKGLPKNPVGSV
SLEAIRAVVFPKKTDFLYFVKMPDKKHAFSATYKEHLKNINLSNNHF

>PROKKA_01212 hypothetical protein
MNKRKHVSKKVFNVII LFVAVFTLLVVIHKTLSNGIHIQNLKIGKLGISELYLKLNNKLS
LEVERVDLSSFFHQKPTKKRLEVSDLIKNI RYGIWAVSYFEKLKVKEI ILDDKNKANIFF
DGNKYELEFPGVKGEFSLEDDKNIKLKI INLLFKDIKVQVDGNAHYS PKARKMAFN LIVK
PLIEPSAAIYLQGLTDLKTIELKINTSPMKSLAFLKPLFQRQS QKNLKTWIFDKIQFASF
KIDNALIKANFTPSEFVPSLLENSVVKATLIKPSVVFNDGLSPIKMDKTELI FKNKQLLI
QPQKITETMELTGSYATFSNLLEAPKLEIFLKTTPNYYGDS IKDLLSAYKVVLP LDKIS
TPSSADLKLTLQFLKNTAPLFSVQGNVNLQEGTSLSLYNIPLYTQ NANISLDITQEYQYIY
IETHTTRYENMLDLDAKIALDLNKKTLSDLSLVHKIRFNTNNNINMRSYGLNNDPDNPQP
TKFSLDLKSLHSVIQEGENSEVFRRKI IDTIKAQSEDKFTKDV FYATGDTLKNLSLNFDF
SNPNHMQWVSPQLLLEGEFKDNAYVFRIKDLKKIKPYSPI MKYFALEDGSLEVSTSDFIN
IDFFAKDLKITLPIYHSGKQFN SISLFGSINKDEISVYTPSKSISIKVKGDQKDITLNN
IDLSIDDFLDSKMPAIAGLFSKERKEKPSSKEIQDEYIFISAKQRYEKAHKI IPISTRIH
AKDVVLIYKKMPFFLENLDIVAQDDRVIDGNYKNAMIMADLVHGALYLKAHNFSGDYIN
TILQKDFVEGGLFTLIGALEDQVFNGELKFQNTSLKNFALMQNM INLINTIPSLIVFRNP
HLGANGYQIKKGSVVFGITKEYLGLEKIDLVGKTLDIAGNGI IELDKNKLDLNLEVSTIK
ALS NVLNKIPIVGYLVLGKGGKITTNVNVKGTLDKPKTQVTLASDIIQAPFKILRRIFT P
IDIIVDEVKKNIESKRK

>PROKKA_01213 FeoA domain protein
MTLNEAIKDKVYEIVEIANCDEALKKRFLSFGIHEGVQCTLLHYS MKKATLSVKINHIOV
ALRSHEAQYLVIKESV

>PROKKA_01214 Ultraviolet N-glycosylase/AP lyase
MGLKRAKTYQKAQQIKELLLKHYPNQTTTELHHKNPYELLVATILSAQCTDARVNQITPKL
FEKYPVNDLALASLEEVKEIIQSVSYFNNKSKHLISMAQKVVKDFKGVIPSTQKELMSL
DGVGQKTANVVL SVCFDANCIAVDTHVFRATHRLGLSNANTPIKTEEELS DLFKDNLSKL
HHALILFGRYTCKAKNPLCGACFLKEFCVSKASFKA

>PROKKA_01215 Flagellar motor switch protein FliN
MPETEANKLKIAEKEKEKANKERELELSTYLEELICDYKNLLDMEIVFSAELGSTQIPLL
QILRFEKGSVIDLQKPAGESVDTFVNGRVIGKGEVMVFERNLAIRLNEILDSNAIVYYLA

KNS

>PROKKA_01216 hypothetical protein

MRLFLLLSATLMLLAEEKIPLNDDAPIKLVHWQNALKEVQPD SNAPATPPIKAVQTTLT
FETPFNKTPKIMEVEGQKVIVLKN AQLD SKKTMDFKEASLNALEMFSYQNDIYLLSKKAT
TELEIQASNSKDKKRLRFLFLPKGFHLAPPPSLKEKSQHANLASQDEKERAQSPLNALEL
KPPLDLSHTYKGLAVIAALLLILYIVKKKIVPTQRSFSTKDFKLEISVLGRVDANHKIIS
IETNKERYLVLLSDKYGLLLDKISP KTSKEELIKEAENNIKNSKLG NLYAGKF

>PROKKA_01217 preprotein translocase subunit SecG

MPENSKLQPAKLGKNFDPVDHSNRNFFFSLILSVLLHWLIYFLFEHREDFFP SKPKLVKL
NPENLLVLKRGSQDPSKNNPGAPKPTLAGPQKPPTPPTPPTPPTPPTPKPIEKPKPEPKPK
PKPKPEPKPNHKKH KALKKVEKVEEKVVVEEKKEEKVVVEQKVEQKKIEEKKPVKKEFDP
NQLSFLPKEVAPPRQENNKGLDNQTRRIDELYGEEFGDLGTAEKDFIRNNLRDIGRITQ
KYLEYPQVAAYLGQDGTNAVEFYLHPNGDITDLKIIIGSEYKMLDDNTLKT IQIAYKDYP
RPKTKTLIRIRVRYLGGN

>PROKKA_01218 Dihydroorotase

MEITLFDPIDAHLHVRENALLKAVLEYSSEPFSAAVIMPNLSKPLIDTPTTLEYEEEEILK
NSSNFKPLMSLYFNDGLTLEELQRAKNKGIFLKLYPKGMTTNAQNGTSDLLGEKTLEIL
ENAQKLGFI LCVHAEQAGFCLDKEFLCHSVLETFALSFPKLKIIIEHLSDWRSIALIEKH
DNLYATLTLHHISMTLDDLLGGSLDPHCFCKPLIKTKKDQERLLSLAKAHPKISFGSDS
APHFVSKKHSANIPAGIFSAPILLPALCELFEKHNALENLQAFISDNAKKIYALDNLPSK
KARLSKKPFIVPHTLCLNEKIAILRGGETLSWNLQEIA

>PROKKA_01219 hypothetical protein

MEPSRNRLKHAAFFVGLFIVFLIIMKRQTPPYAFVHNQTLSTESPPYFTQLTIPKPDDA
LSVHASALINLPNDNLLSAYFSGTKEGARDVKISANLFNKN TNRWSEAFILLTKEELSKN
AHEYIKKLGNPLFLHDDKILLFVVGVS MGGWATSKIYQFESALEPIRFKFARKLSLSPF
LNLSHLIK NKPLNTTDGGFMLPLYHELATQYPLLLKFDQQNNPRELLRPNALNHQLQPSL
TPFKDCAIMAFRNHSFKDSLMELETCKTPTAWQKPMLTNLKNLNDALNLINLNKELYLIHN
PSDLSLRRKELL LSKLENSNSFKTLKILDKADEVSYP SYSLNPHFIDIVYTYNRSHIKHI
RFNMAYLKSLLK

>PROKKA_01220 hypothetical protein

MMLISSSYTLSFIWLF LIFFFFKNKPLSLRFSLSFIIVILSNIAFKDSL SLNEFLSSFTA
PLSPFSCLLILAYALF SCHLLQKPPLETLQSYSVMLFFNLLLLTDILGFLPFSIYHHFMA
SLIFSALFCSSFLSSPLLGVIALMALSSSLMRSNFQILD SLDFPLLLFVFFKTL SLV
KKRLY

>PROKKA_01221 phosphoglycerol transferase I

MKSLSNALFSLFLKGFYFTFFMSLLFVFN RIGFILYTGYKHALKNPVFDEI IKTLFNGA
RYDNRVSSLA LILFIIIGLLGLFAPKHQTKMLNIMACFSITIIILFLNIANIVYYGIYGNV
FDENLLEFLHEDTLTILKMSGEYPIFSSFS LFLILSVLTSFIYFKLQNALFKPKNAYQAA
HTKPLKTFILFALFSLTQMFYINAQLSFVGASLDLSIEPAKDPFLMKITPGA FRNLYLLA
RNYRQSHNLKFSDFAKETPLEVAKNYFHLKENPSNNLYELLSQTS HNDSLQTIQHVFIYIV
SESLSSWHFDKKFDSIGLTSALQDLVKKEHAYMLSAFIESAPRTVKSLDVQITGLPYIND
NNLVNSGVILPSFPMAIGNITKT LGYKNNFY YGGSGIWNKLT SFTKKQGFHALYFNNHLL
EFAKNKPYPKPIESNWGVHDN ILFDYILENTNPNEKTF SMVMTLSNHA IKNVNLKAFGVP
LEKIQKFVEKTPKSENLPDANSLGHIYWDKVI VSFIKKASQKF PNSLFIITGDHFD RSY
EYAKNDLYIIKS VPLILFAPTLNPKKISQVGS HLDIAPTIIELIAPKGFQFVSFGKPLFS
NNTTNPPSHPNYALGYEAIATKDYFYNPSLGLRYL NESPKPKDKQNDKKQNDKIEASKF
YQQLKSLKALSFYLLYHGANKLD

>PROKKA_01222 Bifunctional protein Fold protein

MPNRGVVLLDGQALAYDIEKDLKHKIQIIHTQTHKRPKLAVILVGKDPASITYVNMKIKA
CERVGMDFDLKT LQENITEAELLSLIKDYNTDQNISGVLVQLPLPRHIDTKMILEAIDPN
KDVDGFHPLNIGKLCTQKESFLPATPMGVMRLLEHYHIEIKGKDVAIIGASNIIGKPLSM
LMLNAGASVSVCHILTKDISFYTQNADIVCVGVGKPDLIKASMLKKGAVVVDIGINHLND
GRIVGDVDFNNAQKVAGFITPVPKGVGPM TIVSLENTLIAFEKQQRKGF

>PROKKA_01223 Signal peptidase I

MKFLRSVYAFCSWVG TIIIVLLVIFFFIAQAFIIPSRSMVGTLYEGDMLFVKKFSYGVPI
PKIPWIELPVMPDFKNNGHLIEGDRPKRGEVVVFIPPEKKSYVVKRNFAIGGDEVLF TS
EGFYLHPFESGNDKNYIAKHYPNAMTKEFMGKIFVLNPKYKSKHPGIHYQKDNETFHLMEQ
LATQGVEANISMQLIQMEGEKV FYKKINNDEFFMIGDNRDNSSDSRFWGSVAYKNIVGSP

WFVYFSLSLKNSLEMDAENNPCKRYLVRWERMFKSVEGLEKIIKKENATH
>PROKKA_01224 Peptidase family M50
MQFFDFSLESFITTLMKILALLIAIIGHEIMHGLSAFLFGDRSAKDANRLSLNPIRHLD
MGSVLLPALLLIFQAPFLFGWAKVPVDMRYIVSQKGSACVVVSLAGVAYNFTLAVLLA
FITHLSFQELGINALNINELNLYQLALVTFLIQGILYNLVLGVFNSLPIPPLDGSKALGF
LALHFKSAFLLEWFSKMERYGMLVVVFVFLFIPPLSEFFIHAPTRFLFSLLS
>PROKKA_01225 Putative sugar phosphate isomerase YwlF
MNKPLNIAQVFIGSDHAGHLAEFVKHFLEDKRFKVQAFLLPTMRVDYPDYAKLVCQKVL
NPQSYGILVCATGIGMSMGANRFKGIARAALCLDAYMAKMTRLHNNANVLCLEKISGIGV
AESILEAFFSTEFQGRHVLRIQKLDESLKS
>PROKKA_01226 hypothetical protein
MFTQWFIITIAVVFIYMGVTRTFFFKTVAKRQERTNASMKLTLQEAELIQKHQLQLQRA
LGNIDILTQEMSSSLKTELKALKQRNSEYKGESDKYKNRIKELEQKIEALL
>PROKKA_01227 Adenine phosphoribosyltransferase
MNETLKEELLQSIREVKDYPKKGILFKDITLLNYPKLFNKLIDALKKRYLALNIDFIVG
IEARGFILGSALAYALGVGFVPVRKKGKLPAPHTLSQSYSLEYGSDSIEIHSDAFRGIKGV
RVVLIDDLATGGTALASLELIKALQAECIEACFLIGLKELPGIQLLEERVKTFCLEEC
>PROKKA_01228 Inner membrane protein YohD
MEEYIIDLWNQHAATWGYLILFGWSILEGEIGLILAGIASYTGMMHLGLAILVAGIGGFV
GDQIYFYIGRTNKAYIQKKLEKQRRKLALAHLLQKHGWFIIIFIQRYMYGMRTIIPISIG
LTRYSALKFAIINLISAMVWASITIIILAWCLGEELLHALEWLKKHPYALILLVVSFLALV
LWYFQYYSSKKNR
>PROKKA_01229 Cytosol aminopeptidase
MLKIKLEKTTTFENAKAECSLVFIINKDFDHAWVKNKELLETFKYEGEGVFLDQENKILYV
GVKEDDVHLLRESACLAVRTLKKLAFKSVKVGVTTCATHAKDNALLENLKAFLGLKLGL
YEYDTFKSNKKESVLKEAIVALELHKPCEKTCANSLEKSAKEALKYAEIMTESLNIVRDL
VNTPPMIATPVYMAEVAQKVAKENHLEIHVHDEKFLEEKMMNAFLAVNKASLAINPPRLI
HLVYKPKKAKKKIALVGKGLTYDCGGLSLKPADYMTMKADKGGGSAVIGLLNALAKLGV
EAEVHGIIIGATENMIGPAAYKPDILISKEGKSIEVRNTDAEGRLLVADCLSYAQDLNPD
VIVDFATLTGACVVGLGEFTSAIMGHNEELKNLFTSGLESGELLAKLPFNRHLKKLIES
KIADVCNISSSRYGGAITAGLFLNEFIRDEFKDKWLHIDIAGPAYVEKEWDVNSFGASGA
GVRACAFVEELLKKA
>PROKKA_01230 Ribosome-binding ATPase YchF
MGLSVGIIVGLPNVKGSSFTNALTKTQNAQSANYPFCTIEPNKAIVNVPDRRLDALAQIVK
PERILHSVVEFVDIAGLIKGASKGEGGLGNQFLANIKECEVILQVVRFCEDDNITHVNDKI
DPLNDIETIELELILADIATLDRIDRLQKALKSSKDAKNLLECALSLKTHLEELKPAKT
FPLNTSEAFLELDKELRFLSHKKMIYAANVGEDLNTLNEHAKKVQNYAKDQKSEFVALC
AKLEEMVMSSEDEVKEFLQSLGIEESGLEKTIRLSFKELGLINYFTAGVKEVRSWTIKK
GSSAPVAAGVIHKDFEKGFIIRAETISYDDFIAYKGEAGAKEKGALRIEGKDYIVQDGDVL
HFRFNV
>PROKKA_01231 pyrroloquinoline quinone biosynthesis protein PqqE
MTPNKKLFKKIYIELSDICGLQCSFCPNPKNIRGVMPLELFEKVCKEAAPLTPITLHVL
GDPCKLKNLNHYLSTAKRFSKVDLVTSGVYLHDFETLLQDAIYQISISLDAGLDNHNKL
NQHRYIQKILEFCRYKFEKNSEVFLNLRIQDSTLDKHQNLIKPFLESFECVSLGGLKSQG
RARLFKKSFLNIQKTFKWPNLNAQNPLNQESKIPYCYGLIKQIATLSNGVVVPCCMDTQA
HINLGDNLNHTPLKDVLSQKAMAIAKTHFLKGAELELLCKNCSEYPLIRYKK
>PROKKA_01232 hypothetical protein
MKAQYFFWILFLIGFYWMIYLYQDFLMDALIAGLLCVGFFQVKVFLDKRFFNVVSSFLCV
LILASVLIVPLYFIVYKSSNIIIFEINFEKFSALIKWLKGTITENLSHFPPIHDGVSKFLE
NFSASITGYLLKISSYVGRYSLKLITDALFILGLLFFFFYYGERFYRYFLGVLPGLINQ
SKKIFEEVAGILRIVLLTSLITVILEGVAFGVMIVWFGHDGWSLGILYGLASLVPVAVGGA
LIWIPIAIYELYHGNVNEAIFIALYSILLISVLIDSVIKPILIVFIKKRIFKTTLKINEM
LIFFSMIAGISQFGFWGIIVGPTITAFFIALRLRYENYFIQNDQKACE
>PROKKA_01233 Diaminopimelate epimerase
MVFYKYSGSGNDFLITQSFKKRDFSNAKQVCHRHEGFGADGLVVVLP SKDYDYEWDFYN
SDGSKAGMCGNASRCVGLFAYQHVIAPEKHVFLAGKREISICIEEPNIIESNLGNYKILD
TIPALRCKNFFSSDSVLEHISTFYLIDTGVPHLVGFVNKEWLNLSNLTLELRALRHEFNA
NINIAFIENKETIFLQTYERGVEDFTLACGTGMAAVFIAARIFYNTPKKAALIPKSNESL

ELSLKNDGIFYKGAVRYIGMSVLGMRVFEENGCF

>PROKKA_01234 Inner membrane protein YkgB

MQALKSLLEVITKLQNLGGYLMHIAIFIIFIWIGGLKFVPYEAEGIAFPVANSPPFFSFMY
KFEKPAYKQHKMSESQSMQEEMQDDPKIVENKEWHKENRTYLVAEALGITIMILGILVLL
GLWMPLMGVIGLLVAGMTITTLSFLFTTPEVFNQHFVWLSGAGRLVVKDLALFAGGLF
VAGFDAKRYLEGKGFCLMDRSSVGKTKCSSGCCS

>PROKKA_01235 hypothetical protein

MELGNKNIKPGRKRVAVDELKRNFSVTIFYLSKDEHDVLRRLADEEVEESVNSFVKRHILKT
IIHKKGANQDLSVNDSSSRL

>PROKKA_01236 hypothetical protein

MAINTFLKHSFLICLLAVNSYAFDWNIFKYNLGFNMFIMDHEGSTPYWVNTNTNLKTRLT
PNFGIQFYTRGVEQSLTVGAYFFQNFHNYSTNFPYRWGPTMYKARGKHFTFYGGIFPRK
NLLGRYGLNIFAPYYWFIDPNARGFLLQFQNHYSKPYGHAEFMLDWFGGNCYNTCKF
GRNPYGNAMDRFQMNGSVAYNFFKDLLGIGGYFVLFHNEDKYLLNGADGMQFNEKKAIDN
NNIYLMDRLYFNAYIGTSLLDIAPFMEKLNASFGMVSESSRLRQIHKNVPFMNSVGGQLD
VEIQYKGFGIHNLFFFAKTPEMPFYNQYQYVEMYCTPSYCPTPIYRGVPFFQANLYNRFD
FYYNWKNDFASVRINFVLNAMRGGFDRSLPWSESYQVYMTVAFDPYNLINKIARKK

>PROKKA_01237 30S ribosomal protein S21

MPGIKVRREGDAFDEAYRRFKKQTDNRNLVVTECRARRFFESKTEKRKKQKISAKKKVLKRL
YMLRRYESRL

>PROKKA_01238 3-oxoacyl-[acyl-carrier-protein] reductase FabG

MQFTGKNVLITGASKGIGAEIAKTLASMGKLVWINYRSNAEVADALKNELEEKGYKAAVI
KFDAASESGFIEAIQTIVQSDGGLSYLVNNAGVVRDKLAIKMKTEDFHHVIENNLTSAFI
GCREALKVMSKSRFGSVNVASIIIGERGNMQTNYASASKGMIAMSKSFAYEGALRNIRF
NCVTPGFIETDMNANLKDDELKADYVKNIPLNRLGSAKEVAEAVAFLLSDHSSYITGETLK
VNGGLYM

>PROKKA_01239 Acyl carrier protein

MALFEDIQAVIAEQNLNVDTAQVTPEAEFVKDLGADSLDVVELIMALEEKFGIEIPDEQAE
KIVNVGDVVKYIEDNKLA

>PROKKA_01240 3-oxoacyl-[acyl-carrier-protein] synthase 2

MRRIVVTGMGMINSLGLNKEDSFLAIAKGECKIKHIESFDASAFPVRIAGEITDFDPTEV
MNPKDVKKAGRFIQLALKATREAMKDSGILDHNRCPPELANRMGVSSSGSGIGGLGNIEA
NSIFCFEKGPRKVNPFITISALVNMIGGFTSIEFGIKGNLSSVTACAAGTHAIIIEAVKT
IXLGGADKMLVVGAESTICPVGIGGFASIKALSTRNDEPKKASRPFDKDRNGFVMGEGAG
ALVLEEYESAKKRGAKIYAEFAGYGESGDANHITAPAPEGEGAFFRAMKMALEMAKVEVGY
VNAHGTSTHYNDWYESIALKNVFGSKEKVPVSSTKGQIGHCLGAAGALEAVISIMTMNQ
GILPPTINQETPDPECCLDIYPNAAREKQVDAVMSNSFGFGGTNGVVIFKKA

>PROKKA_01241 Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha

MAIYLDFFENHIKEIQNEIELALIRGDEDAKEILEKRLDKEVKSIYSNLTFQKLQLARHP
DRPYAMDYIDLILKDKYEVFGDRHYNDKAIVCFIGKIDNVPVVVIGEEKGRGTKNKLLR
NFGMPNPGCYRKALKMAKFAEKFNLPILMLVDTAGAYPGIGAEERGQSEAIKNLQEFAS
LKVPTISIIIGEGSGGALAIADKLAMMEYSIFSVISPEGCAAILWDDPSKTEVAIKA
MKITPRDLKEAGLIDDIILEPSKGAHRDKFSAANTIKEYFLDALRTIQQDPHFLDNRYQK
LMSLGSFVEGMN

>PROKKA_01242 hypothetical protein

MVIYEKIKSRFSRNLRLDRGRHFASASVYFFSLLVITAVNRSSAVAWLLMPKHLIWWFL
ISFSGEFVIDMAFGKSKIFKTRFGISIVSGVSLLLGALPAHLFFVWFGFINWWAVFFIE
AGADLLVGCMIQKIFFW

>PROKKA_01243 hypothetical protein

MKNLFLAFIVGGMLLSADALNDKVENLMGERAYHTNKLFLERLFKNRKAFYVMGRLDLKL
LLNTLKENGLLSFNFDKPSMLKITFKASSNPLAFKSSINNSLSMMGYSYVLPKMQSSSG
ENVFSYELKTEYVLDPNILTIETMKRHGFDFVDIRRVSLKEWEYDFSQEVKLPNARALVL
SSDPVEFKEASGKYWLSVNQAYLKISSNNPLWQPKIIFYDENLKI IQIIAKENRQQEIA
LNLLNGVRFIHITDAKNPIILKNGISVVFDAMP

>PROKKA_01244 hypothetical protein

MEQNKKSLNLDLSDVQNISKDISGAALEELSLKNLDKNLQILKEVGVAEICKATKIASK
NIHSILEKRYESLSRVHARGFIQILEREYKIDLSAWMKEFDRVCTLKESVSEEQNQETDP
EEKTKNPLKVEIDYSINQANTSLSKKSSKWKPFVLVLGVVVIILAVVIIQNSSSLKEERG

QESAISGTTKNSFNKVNPTTEENKPEPTPKLEEKPKQDKQEKEAIKEDPNTIYIIPKKD
IWVEVVDLDEKKNSFQKVFKKNYSLETKNHRLLLRFGHGHLSLKNNHQEQEYNDGKTKRF
LYEPNKGLTLINAEQYKELQQ
>PROKKA_01245 Putative TrmH family tRNA/rRNA methyltransferase
MQAVVYGKHVITHLLNSHQEKLQEIYLSKEIDKKFFFALKKACPNIKVDNKKQAQSLAKG
GNHQGVLAKEVELPLAASLKEIKKAQKLLVLCGITDVGNIGGIFRSAYCLGMDGVILDFAK
ELAYEGIVRSSGLMYDLPFVSMPNTLDLINELKTSGLCLGASMQGSSQVENLSLKKCA
LFLGSEHEGLSKKILAKMDAILSVMRRDFDSLNVSAAGILMDKIN
>PROKKA_01246 Ribosomal RNA small subunit methyltransferase I
MLYFLPTPIGNLADITLRALEVLERCEVFLCEDTRVSKRLLHLLAQNPISHSFPNIAAK
KREFIAFHSHNDQEFNLQIEPSFFDKEIAVMSDAGMPSLSDPCMSLVAAHALKHNIYDVL
PGANALTATFACASGFLEGRFFYAGFLPHKSKERRLKIAKILNALAYLEEKTPVVFYESPH
RLLETCLKDLNDLAKGMHLFAAKELTKLHQYYLGEVSQIIERLQKSNIQGEWVLVLLNEK
KIEPCMGLSALLELDLPPKIKAKIEAAMTQKKR
>PROKKA_01247 50S ribosomal protein L31
MKKGIHPEYIPCKVTCVTSKGIEVLSTKPEMRIDISSFCHPFYTGSDKIADTAGRVEKF
KQRYNLK
>PROKKA_01248 hypothetical protein
MNENAPTHKSSHVKVTHTPVSGYHIEDLRTPTEKLLLEIANKLKVENPQEFKRQDLMFEI
LKTQVTQGGYILFTGILEIMPDGYGFLRGFDGSFSDGHNDTYVSPSQIRRFALRNGDIVT
GQVRSPKDQEKYALLKIEAINYLPSDEIKNRPLFDNLTPLPDEQIKLEYEPTKVTGRM
LDLFSFVGKGQRALIVAPPRTGKTELMKELAQGITSNHPEVELIILLVDERPEEVTDMQR
SVKGQVFSSTFDLPANNHIRIAELVLERAKRRVEMGKDVVLLDSITRLARAYNAVTPSS
GKVLSSGVDANALHRPKRFFGAARNIEEGSLTIIATALIETGSRMDEVIFEEFKGTGNS
EIVLARNIADRRIYPAFDILKSGTRKDNILLGKDRLTKVWVLNRNVMQMQMDDIEALS FVYS
KMQQTKDNEEFNLNMNEK
>PROKKA_01249 Glutamate racemase
MKIGVFDŠGVGGFSVLKSLLKAQLFDEIIYYGDSARVPYGTKDPTTIKQFGLEALDFFKP
HQIKLLIVACNTASALALEEMQKHSKIPVVGVIIEPSILAIKQQVKDKNAPILVLGTKATI
QSNAYDNALKQQGYLNVSHLATSFLVPLIEESILEGELLETCMRYYFTPLEILPEVILG
CTHFPLIAQKIEGYFMDHFALSTPPLLIHSGDAIVEYLQQKYTLKKNAAHAFPKVEFHASG
DVIWLEKQAKEWLKL
>PROKKA_01250 hypothetical protein
MRSEVEKYGFKNFDELKIDTVDAFQGEKADIIIIYSTVKTYGNLSFLIDSKRLNVAISRAK
ENLIFVGKKSFFENLRSDEKNIFSAILQVCR
>PROKKA_01251 Cytotoxicity-associated immunodominant antigen
MTNETINQQPQTEAAFNPPQFINNLQVAFIKVDNAVASFDPDQKPIVDKNDRDNRQAFEK
ISQLREKFANKAIKNPTKKNQYFSNFISKSSDLINKDGLIDTGSSIKSFQKFGTQRYQIF
MNWVSHQKDPSQINTQKIRGFMEINIIQPPISDDKEKAEFLRSQAQALAGIIIRNQIRSDQ
KFMGVFDES LKERQEAENGE PNGDPTGGDWFDIFLSFVFNKKQSSDLKETLNQEPVPHV
QPDVATTTTDIQLPPE SRDLLDERGNFSKFTLGDMEMLDVEGVADIDPNYKFNQLLVHN
NALSSVLMSGHDGIEPEKVSLLYGNGGPEARHDWNATVGHNKQOGNNVATLINVHMKN
SGLVIAGGEKGVNPNPSFYLYKEDQLTGLKQALSQKEIQNKVDFMEFLAQNNAKLDNLSKK
EKEKFQNEIEDFQDKHAYLDALGNDAIAFVSKKDKKHLALVTEFGSGDLSYTLKDYGKK
ADKALDREIKTTLQGS LKHGVMFVDYSNFKYTNASKSPDKGVGATNGVSHLEANLSKVA
VFNLPLNLNLAITSYIRRDLEEKLGAKGLSPQEANKLIKDFLNSNKELVGKVLNLNKAVA
EAKNTGNYDEVKKAQKNLEKSLRKREHLEKEVVKKLENRNDNKNRMEAKAQANSQKDKIF
AIINEEASKEARVAACVQKFKGKIMELSDKFENINKNLKDFDKSFDDFKNGKNKDFSKAE
ETLKALKGSVKDLGINPEWISKVENLNTALNDFKNGKNKDFSKVTQAKSDLENSIKDVII
NQKITDKVDNLNQAVSVAKATGDFSGVEQALADLKNFSKGQLAQQAQKNEDFNTGKNSEL
YQSVKNGVNGTLVGNLSGIEATALTKNFS DIKKELNEKFKNFNNNNNGLKNSGEPIYAO
VNKKKTQGVASPEEPIYTQVAKKVAKIDQFNQVASGLGGVGQAGFSLKGHTKVDDLSKV
GRSVSPEPIYATIDDLGGPFPLKRHDKVDDLSKVGLSRNQELAQKIDNLNQAVSEAKACH
FDNLDQMIDKLKDKSTKKNVMNLYVESAKKVPTSLSAKLDNYAINSHTRINSNVKNGTINE
KVTGMLTQKNPEWLKLVNDKIVAHNVGSVPLSEYDKIGFNQKNMKDYSDSFKFSTRLSNA
VKDIKSGFVQFLTNTFSMGSYSLMKASVEHGVKNTNTKGGFQKS
>PROKKA_01252 hypothetical protein
MKTIRNSMFIGASLLGGCASXGYFDALRVACVKDAVIEKESHHTPKDFGSFYFSNQTDI

GF

>PROKKA_01253 hypothetical protein

MCKGLLYKHIIEESVFEVSPMYENAIWDKITKCLFHGIQLLLPIAYKEFKHKKYTPRDTL
KVWVTNYDSLKSAFIAEKEPEPKHMDVIMLFNDRLGSQLKYAYIFNDFNLNYSVLDYQ
KNGAIKNTKQFYDPENIIRWINVEREHKLEKTSSYNKNQVQKIIELLEQINRILNQKIR
KTIGIITPYNAQKRCLRSEVEKYGFKNFDELKIDTVDAFQGEKADIIISTVKTYGNLSF
LIDSKRLNVAISRKENLIFVGKKSFFENLRSDEKNIFSAILQVCR

>PROKKA_01254 hypothetical protein

MQEWMDERQEKIDRTQHFITKAQSKIKEIESNIDEILKHIQTMEPNHLLSQDFLYVGKLA
CVNKSEIYAHDGDEESGWNIIYQRLSLESGEHEKSYGILITLKRNFIF

>PROKKA_01255 hypothetical protein

MLPFIVGGAILAVGFGVLAFFDVETEKEKERQNTLKKDIENYTHKAQQAKKRHKHQOTL
KIADHCLNIIQRYLQEIERLQQEKKETPTQLEAVLEQIQQEMKSLPIQQKHSLQIYCNRF
EYAQRALAYLDYLSRLKTQLQEKEKRLKSDLGEIKKLQKVKENEKDIKKHQEALEQFE
E

>PROKKA_01256 hypothetical protein

MKSVNLPTPKGNSKGGFLDNLSDMLKLVTASKDLISSAINAQKEISKEQEKTKQVEIQ
ANQEMIASNNRLREKLAEFQRDITALNKIHEQEMLRLSNEHEKDMKLEIMQGIINKIDS
LNQFLICEHGEHQKQVDELHAQLIHYQQTLLALTSKVN

>PROKKA_01257 hypothetical protein

MFYRDRNTLQATMISIDLFKTDGRFVRGMCNTKIIYSVIRYGAKGIAKVSPVLVCVDVVL
SLADMIHSYDQYRAAKEQTKQLEIIRNTLRKQYENLLIELRLDKQKLRQLAKDLEKIDA
RIHKNA DKMHQLKLAYKCSFMVLKCIKGHLDEYEKKFPYDNTQRIVALCQQYHEALTVHC
QASLNFIGD

>PROKKA_01258 hypothetical protein

MERIEDIGFTFLFCLHAFGDGLNLNELSQVTEIDPITIQLHDLFLVKRGFVNEKHKISVC
GCNILKLHDEINKFNRTNRVVFLENNAVREKVWRERQELTDRSCGWLIAGEDLSLQEFE
ELYNARKDIDRVLCCLDNTKDFKDVSEIRVRLEPKKVERFLVFKINPKALTPTKIERKT
YARAIHGEQECFIEIETLCKQL

>PROKKA_01259 L,D-transpeptidase catalytic domain

MKKVLPALLMGFVGLNASDRLLIIMHLYQKQGLEVVGGKLD SYLADKSFWAEELQNKD TD
FGYYQNKQFLFVANKSKPSLEFYEIDNNMLKKINSSKALVGSKKGDKTLEGDLATPIGVY
RITQKLERLDQYYGVLAFTVNTYPNLYDTLKKRTGHGIWVHGMP LNGDRNELNTKGCIAIE
NPILSSYDKVLKGEKAFLITYEDKFFPSTKEELSMILSSLFQWKEAWTRGDFERYMRFYS
PNFTRYDGMKFNAFKEYKKRVFAKNEKKNIAFSSINVIPYPNSQNKRLFYVAFDQDYKAY
QQNKLSYSSNSQKELYIEIENNQVSIIMEK

>PROKKA_01260 GTPase Era

MKTKAGFVALIGKPNAGKSTLLNTLLNAHLALVSHKANATRKL MKCIVPFKDKEGYESQI
IFLDTPGLRHQEKLLNQCMLSQALKAMGDAELCVFLASVHDDLKGYEEFLNLCQKPHILA
LNKIDTATHKQVLQKLQEYQKYASQFVDLVLPSAKKSQNLNALLECISKHLSPSAWLF EK
DLMSDEKMRDIYKEI IRESLFDFLSDEIPYESDVMIDKFIEEERMDKVYARIIVEKESQK
KIVIGKNGVNIKRIGTSARLKMQEVGEKKVFLNLQVIAQKSWSKEEKSLQKLGYIHKRGM
D

>PROKKA_01261 ATP-dependent protease ATPase subunit ClpY

MSKLNMTPREIVAYLDEYIIGQKEAKKSIAIAFRNRYRRLQLEKSLQEEITPKNILMIGS
TGVGKTEIARRIAKIMELPFVKVEASKYTEVGFVGRDVESMVRDLVNNSVLLVENEHKEK
LKDKIEEAVIEKIAKKLLPPLPNGVSEEKKQ EYANSLLKMQQORIAQGELDSREIEIEVRK
KSIEIDSNVPPEILRVQENLIKVFHKEQDKVKKTL SVKEAKEALKA EISDTLLDSEAIKM
EGLKRAESSGVIFIDEIDKIAVSSKEGSRQDPSKEGVQRDLLPIVEGSVNTKYGSIKTE
HILFIAAGAFHLSKPSDLIPELQGRFPLRVELENLT EIMYMILTQTKNSIIKQYQALLK
VEGVGIAFEDSAIKELAKLSYNANQKSEDIGARRLHTTIEKVLEDISFEAEDYSGQSVTI
TKELVQSKLEDLVADENLVKYIL

>PROKKA_01262 ATP-dependent protease subunit HslV

MFEATTILGYRGEMGGKKFALIGGDQVTLGNCVVKANATKIRS LYHNQVLSGFAGSTAD
AFSLFDMFERILES KGDLFKSVVDFSKWEWRKDKYLRRLEAMMIVLNLDHIFILSGMGDV
LEAEDNKIAAIGSGGN YALSAARALDHFAHLEPRKLVEESLKIAGDLCIYTNTNIKILEL

>PROKKA_01263 50S ribosomal protein L9

MKVLLLEDVKNLGKAGEVCEVKDGYGNNFLIANQKAKLATNEVINKYKAEVKKKAEKEAL

EKAQKLQMVETLQTITLTIHKKVANGSLFGAITKEEITERLKEQHASLNLDKKDIELKH
PIKSTGIYEIEVKLGSGVVGAFKIDVVAE
>PROKKA_01264 hypothetical protein
MKANAIKLLDFIGKSQEQFVIPIYQRVYSWEKEQCKQLWDDIIKTGGNDQIEGHFIGSI
VFVHDGIYTTNYNELLIIDGQQRLTTITLLFIALRDHLNDEVKRKEIEDHYLINSNGNDGD
KKIRLILSESDRDTLLSLIDKDKRKPSSELSLKIMENFKLFEWIRKNTDKLETIFKGLEK
LMIVWIALKKEKDDPQLIFESMNSKGIELTQTDLIRNYIVMETEVGKQEDFYNQYWRAME
EDFKQNETLFRFVRHYLTIKTGDIPNISKVYEAFKRYQQERGIETEALLQDLQKYCGFF
CRIVFKKEADKDLNKALSFLVDLEMDVIYPLLELYSDYSDGVLISKQDFIPIIALIESYI
FRRAVCGLGTNSLNKIFASFTKKINKDQYLESIKAHFLSLETTKGKFPKDSEFKNLFITI
DFYNLKEKKYFFERLENFNTKEPVNTKECTIEHIMPCTLTEEWEKDLGENFQAIHDKYLH
TIGNLTLTGYNQEYSNRSFQEKRDMEKGFKQSPLRLNQGLRDLESFGEKEIEKRANLAD
WALKIWTPNLDAETLEKYKQKGKREKKAYDLNKFSGHRLRELFILSKEIKALDERITEK
FNQEYISYMFDFKNFVDIVPLKNGLLKLYLNMPFNELQDEKNLARDMTNKGHLGNGDIEVKL
ETKEDIPLYCLGLIKQVLEKQMGGKNRQ
>PROKKA_01265 3-hydroxylaminophenol mutase
MIVRTQNSESKIKEFFEFCKENEVEFVDFRFSIDIKGTWNHIAYSFGALTHGMLKEGIPFD
ASCFKGWQGIEHSDMILTPDLVRYFIDPFSADVSAVVFCDVDYVYKNQPYEKCPRSIACK
ALQHLKDSGLGDVAYFGAENEFFIFDSIKIKDASNSQYYEVDSEEGEWNRDKSFENGVN
GHRPGKQGGYLPVPPTDTMMDIRTEIVKVLNQVGLTFVHHEVAQAQGEVGVKFGDLVE
AADNVQKLKYVVKMVAHLNGKTATFMPKPLYGDNGSGMHTVSVWKNENLFSGETYKGL
SELALHFLGGVLRHARGLAFTNASTNSYKRLIPGYEAPSILTYSANNRASVRIPYGIS
KNSARFEFRFPDSSSNPYLAFGAILMAGMDGVKNKIDPGEAMDINLFKLTLDIREEKGIK
QMPHTLRRSLEEMLADKQYLKEGQVFSEEFIQAYQSLKFNAEVFPWESKPHPFEFITTYS
C
>PROKKA_01266 Beta-lactamase HcpA precursor
MFKDFYRTTSLFNPLLLLLVLLLPFSLCIADEYISISDDWDERARNQWDEIARNHKTYFF
ENGLDHFNQGGYQQAQFDFKLAQEYSIGLGSVYLAKMYLEGKGVKVDYKKAQFYAQNAIK
GYGSGLLGGALVLGRMQAQLGMKKDLKQVLKTYRHVVRMFSNKSTNYFANNFRLPNLAF
TSMILGSRFIDLSGLSANPIKFGKKFGILVMKSTQIKDKTLLWEDIAEISSNIILLKQOM
GEILYRIGIAYKEGLGTRKKKDRACKFLQKSAEFGYEKAMEAL
>PROKKA_01267 hypothetical protein
MADTINTTETTHETKPNFVNFNFKNSLADKRYDSLGLIGAGVLCCVLSGAMGIVGIIIFV
AIGIFLSFSNINLVKLVEKLSKKQPKAATTNNETQKSQATSVTNGPTEAKENKD
>PROKKA_01268 DNA topoisomerase III
MASRHTDDSNKFNKPKKEIDENVINALETSLKSKGVVSETLESIKEAKNNKEYENLSNR
GFLRAKLDQGVGLNFQESIGYNLSALQKKAISLLKISPNIQILELVQKLYMKGRISYPT
EYLFEEFTFKKNLYKIYKDSENDCFNKNVDDEEKVLLGFPNRKFTPNEFFTATSLTLNAREFC
LYINLKKKETNV
>PROKKA_01269 hypothetical protein
MFRKLATAVSLIGLLISNTLYAKEISEADKVIKATKETKETKKEAKRLKKEARQRQQIPD
HKKPQYASVDDTKTQALFDIYDTLNVNDKSFQDWFQNSALKDKTYLYAMDLLDYNLYLSI
ENPIIKTRAMGTYADLIIITGSLEQVNGYYNALKALNKRNAKFVLKINENMPYAQATFLR
VPKRSDPNAHTLTKGAIDENKLFQEQKKMYFNYANDVICRPDDEVCSPLRDEMAMPTS
DSVTQKPNIIAPYSLYRLKETNNANETQPSPYATQTAPESSEKELIEELIANSQLIANEE
EREKKLLAEKEKQEAELAKYKLKDLNQKKLKALEAELKKKNAKKPRVVEVPISPQTSNS
DETMRVVKEKENYNGLLVDKETTIKRSYEGTLISENSYSKKTPLNPNDLRSLEEEIKSY
IKSNGLCYANGISLYAKIKNDPYKEGMLCGYESVQNLLSPLKDKLKYDKQKLQKALLKDS
K
>PROKKA_01270 hypothetical protein
MFGKWIGLTLFLNSLAYPCQKVTITFKQYENLTQIRQKGCNEVVCRTLISIALLESSLG
LNNKREISLKDTSYSMFHITLNTAKKFYPTYSKTLLKFKLLNDVGFAIQLAQILKENFD
YYKQKYPNKSQYQLVEMAIGAYNGGMKHNPNGAYVKKFRCIYSQVRYNE
>PROKKA_01271 Conjugal transfer protein TraG
MEDFLYNTLYFIEDYKLVVIFSFIGLIALFFLYKFIKAQKKAQKDKANQPQKKKSFKEII
IDGLKERVKTFFGFWLQAILLSYSFITSGLFFLILLGNFYDDNRSPESSDDDLFDIWIYAI
QDFPAYYFKALCFSSLKIYGFNISLVVYSSILCSYIFITFFVWFLKYLTRTRDIGANKKV
DDLFGSASWETEEKMIKAKLITPNNKKRTFDKREVIVGRRGLGDFIAYAGQAFIGLIAPT

RSKGKGVGFIMPNNMINYPQNIIVVFDPKADTMETCGKIREKRFNQKVFIYEPFSLKTHRFNP
FAYVDFGNDVVLTEDILSQIDTRLKGHGMVASGGDFSTQIFGLAKLVFPERPNEKDPFFS
NQARNLFVINCNIYRDLMWTKKGLEFVKRKKIIMPETPTMFFIGSMASGINLIDEDTNME
KVVSLMEFFFGGEEDKSGDNLRVLNPNATRNMMWNNFKTMGGAKETYSSVQGVYTSAFAPYNN
AMIRNFTSANDFDFRRLRIDEVSIGVIANPKESTIVGPILELFFNVMIYSNLILPIHDPQ
CKRSCLMLMDEFTLCGYLETFFVKAVGIMAEYNMRPAFVFQSKAQLENDPPLGYGRNGAKT
ILDNLSLNMYYGINNDNYEHFEKLSKVLGKYTRQDVSRSIDDNTGKTNTSISNKERFLM
TPDELMTMGDELIILENTLKPIKCHKALYYDDPFFTDELIVSPSLSKKYKLGVNPNQAT
FYDDLQAAKTRGELS YDKSLVPVGSSEL

>PROKKA_01272 Type IV secretion system protein VirB11

MTEDRLSAEDKKFLEVERALKEAALNPLRHATEELFGDFLKMENITEICYNGNKVVWVLK
NNGEWQPFQDVRDRKAFSLSRMLMHFARCCASFKKKTIDNYENPILSSNLANGERVQIVLSP
VTVNDETISISIRIPSKTTYPHSFEEQGFYNLLDNKEQAISAIKDGAIGKNVIVCGGT
GSGKTTYIKSIMEFIPKEERIISIEDTEEIVFKHHKNYTQLFFGGNITSADCLKSCLMR
PDRIILGELRSSEAYDFYNVLCSGHKGTLTTLHAGSSEAFIRLANMSSSNSAARNIKFE
SLIEGFKDLIDMIVHINHHKQCDEFYIKHR

>PROKKA_01273 CagZ

MELGFNEAERQKILDSNSSLMGNANEVRDKFIQNYASSLKDSNDPQDFLRRVQELRINMQ
KNFISFDAYYNYLNNLVLASYNRCKQEKTFAESTIKNELTLGEFVAEISDNFNFTCDEV
ARISDLVASYLPREYLPFFIDGNMMGVAFQILGIDDFGRKLNEIVQDIGTKYIILSKNKT
YLTSLERAKLITQLKLNLE

>PROKKA_01274 Type IV secretion system protein virB10

MNEENDKLETSKKIQQHSPQDLSNEEATEANHFEDSSSKESKESSDHHLNNPTETKTNFDG
DKSEETQTQMDSGGDETSESSNGSLADKLFFKARKLVDNKRLLFTQQKNLDEETQELNEED
DQENNGYQEETQTGLIDDETSKKXQQHSPQDLSNEEATEINHFEDSSSKESKESSDPYLDN
PTETKTNFDKSEETQTQMDSGGDETSESSNGSLADKLFFKARKLVDNKRLLFTQQKNLD
EETQELNEEDDQENNGYQEETQTGLIDDETSKKXQQHSPQDLSNEEATXINHFEDSSSKES
KESSDPYLDNPTETKTNFDKSEETITNDSNDQEI IKGSKKYI IGGIVAVLIVIIILFS
RSIFHYFMPLEDKSSRFKDRNLVYNDEIQIRQEYNRLKERNEKGNMIDKNLFFNDDPN
RTLYNYLNIAEIEDKNPLRAFYEICISNGGNYEECLKLIKDKKLQDQMKKTLEAYNDCIKN
AKTEERIKCLDLIKDENLKKSLNQQKVQVALDCLKNAKTDEERKECLKLINDPEIREK
FRKELELQKELQEKDCIKNAKTEAEKNECLKGLSKEAIERLKQQALDCLKNAKTDEERN
ECLKNIPQDLQKELLADMSVKAYKDCVSKARNEKEKQECEKLLTPEARKKLEQQVLDCLK
NAKTDEERKKCLKDLPKDLQSDILAKESVKAYKDCVSAKNEAEKKECEKLLTPEAKKLL
EEEAKESVKAYLDCVSAKTEAEKKECEKLLTPEAKKLEEEAKESVKAYLDCVSAKTE
AEKKECEKLLTPEARKKLEAKKSVKAYLDCVSAKNEDEKKECEKLLTPEARKKLEQQA
LDCLKNAKTDEERKKCLKDLPKDLQKKVLAKESVKAYLDCVSAKNEAEKKECEKLLTPE
ARKLLEEAKESIKAYKDCVSKARNEKEKKECEKLLTPEAKKLEEEAKESVKAYLDCVSA
AKTEAEKKECEKLLTPEAKKLEEEAKESL KAYKDCVSRARNEKEKKECEKLLTPEAKKLL
EQQALDCLKNAKTEADKKRCVKDLPKDLQKKVLAKESVKAYKDCVSRARNEKEKKECEKL
LTPEAKKLEEEAKESL KAYKDCLSQARNEEERRACEKLLTPEARKKLEQEVKKS VKAYLD
CVSAKTEAEKKECEKLLTPEARKFLAQVLNCLKAGNEEERKACLKNLPKDLQENVLA
KESL KAYKDCLSQARNEEERRACEKLLTPEARKKLEQEVKKS VKAYLDCVSRARNEKEKK
EKEKLLTPEARKFLAKELQQKDKAIKDCLKNADPNDRRAIMKCLDGLSDEEKLKYLQEAR
EKAVLDCLKTARTDEEKRCQNLYSDLIQEI QNKRTQSKQNQLSKTERLHQASECLDNLD
DPTDQQAIEQCLEGLSDSERALILGIKRQADEVDRIYSDLNRNKTFDNMAAKGYPLLPMD
FKNGGDIATINATNVDADKIASDNPIYASIEPDITKQYETEKTIKDNLEAKLAKALGGN
KKDDDKEKSKKSTAEAKVESNKIDKDVAETAKNISEIALKNKKEKNGEFVDENGNPIDDK
KKAEQDETSVPKQAFIGKSDPTFVLAQYTPIEITLTSKVDATLTGIVSGVAKDVWNMN
GTMILLDKGTKVYGNYSVKGGTPIIMTRLMIVFTKAITPDGVII PLANAQAAGMLGEAGV
DGYVNNHFMKRIGFAVIASVNSFLQTAPIIALDKLIGLGKGRSERTPEFNALGQAING
SMQSSAQMSNQILGQLMNIPPSFYKNEGDSIKILTMDDIDFSGVYDVKITNKSVDDEIIK
QSTKTLREHEEITTS PKGGN

>PROKKA_01275 Conjugal transfer protein

MGQAFFKKIVGCFCLGYLFLSSAIEAAALDIKNFNRGRVKVVNKKIAYLGDEKPITIWTS
LDNVTVIQLEKDETISYITTFNKGWSIVPNSNHIFIQPKSVKSNLMFEKEAVNFALMTR
DYQEFLLTKKLIVDAPDPKELEEQKKALEKEKEAKEQAQKAQKDKREKRKEERAKNRANL
ENLTNAMSNPQNLNNKLNSEFIKQQORENELDQMERLEDMQEQAQANALKQIEELNKKQA

EEAVRQRAKD KINIKTDK SQSPEDNSIELSPSDSAWRTNLVVRTNKALYQFILRIAQKD
NFASAYLTVKLEYPQRHEVSSVIEEELKKREEAKRQKELIKQENLNTTAYINRVMMASNE
QIINKEKIREEKQKIILDQAKALETQYVHNALKRNPVPRNYYQAPEKRSKHIMPSEIF
DDGTFTYFGFKNITLQPAIFVVPDGLSMTDAAIDPNMTNLGLRWYRVNEIAEKFKLIK
DKALVTVINKGYGKNPLTKNYNIKNYGELERVIKKLPLVRDK

>PROKKA_01276 hypothetical protein

MKRTFLITIISSFFLIVPNWLK AIDLPIVSNLKIYQTVYCM LIPSYVL TNKSFADILTGYT
SIGASGSGKSSGQGVIEALSTPLATSLAASN LVKYLNTLGPLWGS AWASVATAIQGFALT
PSSGCNFGWNALINKNIDVSMDSVLDNLSNKIQNFTKGGVEDNVKGNILLQIIGSITAQA
STNITADGLIWLIGKEFTANKLQNNTTAMLAFVALESVVKGADAAVLPAVG VVNLPDIII
GQGSYLDVFSYLIYIVFGIFVFISFMKLRDISNGIQINIGFEYMR FVG GTLFKMAMVSFI
AYAGFGYLYKISYSIYFGLAGAFGLNQVLFWALDLVLNYTVNSILPAVRAVFSNVGNNA P
SLLQGLQVAGISLFAIFMQVTIIMRISTVVVKPLIAGAFSGIVFP IAVCLIVLDWFKDSM
KNILIWFINNLFILVLAIPILLFGVLALLAFNL TITPSVAIQNINQGGLGIDSTIASLIT
LFILKGFIETIIESVNAIVNTIFSSVSMDGSRMDRERDALMVGRVGGSMFKG

>PROKKA_01277 VirB8 protein

MLGKKNEEVLIDENLVGGVIALDRLAKLNKANRTFKRAFYL SMALNVAAVTSIVMM MPLK
KTDIFVYGIDRYTGEFKIVKRS DARQIVNSEAVVDSATSKFVSL LFGYSKNSLRDRKDQL
MQYCDVSFQTQAMRMFNENIRQFVDKVR AEAISSNIQREKVKN SPLTRLTFFITIKITP
DTMENYEYITKKQVTIYYDFARGNSSQENLIINPFGFKVFDIQITDLQNEQTVSEILRKI
KEVESKNKALKQ

>PROKKA_01278 hypothetical protein

MNDTTEHHGPNPLNAPPPSNSQSNDDLNL LDSLYPKGS LGEQRFHEALKNQEELKNILIE
IEKL PQEKRYELLMQIGQAKQRIMEAYAH SFLGYIGGLEHLLGLCMGGIFVLFAIYFVFL
RTSKNMELVESLKT KLQFYFYAFGVGAVLFFGLETIRSIYELYILGIGSTNDKVL FVL
KNICFIGMGYLIYKVIKVIGIKNFINGLFTSKKQE Q

>PROKKA_01279 CAG pathogenicity island protein 12 precursor

MKMRASVLIGATILCLILSACS NYAKKVVKQKNHVYTPVYNELIEKYSEIPLNDKLKDTP
FMVQVKLPNYKYDYL LDKQVVLTFKLVHHSK KITLIGDANKILQYKNYFQANGARSDIDF
YLQPTLNQKGVM IASNYNDNPNNKEKPQTFDVLQGSQ PMLGANTKNLHGYDVSGANNKQ
VINEVAREKAQLEKINQYYKTL LQDKEQEYTTTRKNNQREI LETLSNRAGYQMRQNVISSE
IFKNGNLNMQAKEEEVREKLQEERENEYLRNQIRSLLSGK

>PROKKA_01280 CAG pathogenicity island protein 13

MRKLFSMIADSKDKKEKLIESLQENELLNTDEKKKIMDQIKTMH DFFKQMHTNKGALDKV
LRNYMKDYRAV IKSIGVDKFKKVYRLLENETMELLHAI AKNPNFLFSKFDRSILGIFLPF
FSKPVMFRMSIREMDSQIELYGT KLPLLLKLFVMTDEEVNFYANLKTIEQYNDYVRD LLMK
FDLEKYMKEKGVQNA

>PROKKA_01281 hypothetical protein

MGKILASLLGGGTNLFTGLSSDLFSMILNFLFFLMLMMGLNEALGKKFNLPMDNIKNFMA
EVLKNGFDSIKNMG SALVSNNGFDSNKS DKTNTKMSVPQVRL

>PROKKA_01282 hypothetical protein

MLAKIVFSSLVAFGVLSANVEQFGSFFNEIKKEQEEVA AKEDALKARKKLLNNT HDFLED
LIFRKQKIKELMDHRAKVLSDLENKYKKEKEALEKETRGKILTAKSKAYGDLEQAL KDNF
LYRKLLPNPYAYVLNQETFTKEDRERLSYYY PQVKTSSIFKKT TATTKDKAQALLQMGVF
SLDEEQNKKASRLALS YKQAI EYSNNISNLLSRKELDNIDYYLQLERNKFDSKAKDIAQ
KATNTLIFNSERLAFSMAIDKINEKYL RGYEAFSNLLKSVKDDVELNTLTKNFTNQKLSF
AQKQKLCLLV LDSFNFDTSKKSILKKTNEYNI FVDS DPMMSDKTTMQKEHYKIFNFFKT
VVSAYRNNVAKNNPFE

>PROKKA_01283 hypothetical protein

MILKSIFKKLGSVALYSLVYGG LNAINTALLPSEYKKLVALGFKKI KTLYQRHDDEEIT
KEEKEFATNALREKLRNDRARAEQIQKNIEAFEKKNNSSI QKKA AKHKGLQELNEINANP
LNDNPNSNSSTETKSNKDDNFDEMINKVNEAFVKPAAPLPVNEWRTPEIEII INECI ISS
NDYDGLRKLKLIKIGIKDQKILAPLLEKIQEIETENNKFSRQHLSGLKLALNNSNNRTFLIA
SCAICEKRKKEME QENNYQDTTNTSEFGATDTKENEAKDATFSNNH SKSELPNSVINQIE
QSI AHGKK

>PROKKA_01284 hypothetical protein

MKTLVKNTISSFLLLSVLMAEDITSGLKQLDN TYQETNQQVLKNLDEIFSTTSPSANNKI
GQEDALNIKKAAIALRGDLALLKANFEANELFFISEDVIFKTYMSSPE LLLTYMKINPLD

QKTAEQQCGISDKVLVLVLYCEGKLKIEQEKQNI RERLETS LKTYQSNIGGTASLITASQTL
VESLKNKNFIKGIRKLM LAHNKVFLNLYLEKLDALERSLEQSKRQYLQERQSSKIIVK
>PROKKA_01285 hypothetical protein
MKCFLSIFSF LTF CGLSLNGTEVVIT LEPALKA IQADAQAKQKTAQAELKAIEAQSSAKE
KAIQAQIEGELRTQLATMSAM LKGANGVINGVNGMTGGFFAGSDILLGVMEGYSSALSAL
GGNVKMIVEKQKINTQTEIQNMQIALQKNNEI IKLMNQNALLEALKNSFEPSVTLKTQ
MEMLSQALGSSSDNAQYIAYNTIGIKAFEETLKGFETWLKVAMQKATLIDYNSLTGQALF
QSAIYAPALSFFSSMGAPFGI IETFTLVPTKCPYLDGLKISACLMEQVIQNYRMIVALIQ
NKLSDA DFQNIAYLNGINGEIKTLKGSVDLNALIEVAILNAENHLNYIENLEKKADLWEE
QLKLERETTARNIASSKVIVK
>PROKKA_01286 hypothetical protein
MAGTQAIYESSSAGFLSEISSIISSTSSVAGPFAGIVAGAMSAAIPIVVGFTNPQMTAI
MTQYNQSI AEAVSMPMKAANQQYNQLYQG FNDQSMAGVNNILNISKLTGEFNAQGNTQGA
QISAVNSQIASILASNTTPKNPSAIEAYATNQI AVPSVPTTVEMMSGILGNITSAAPKYA
LALQEQLRSQASNSSMNDTADSLD SCTALGALVGSSKVF FSCMQISMT PMSVSMPTVYAK
YQALATNALTSGVNPMTTPACPIGDKVLAVYCYAEKVAEILREYYIEFVKNNTNLLQNAS
QMILNQSGLATSTYDTQAISNISSLYNNIVANKSFLKSHLTYLDYIKDKLKGQKDSYLT
ERVQTKIIVK
>PROKKA_01287 hypothetical protein
MKTNFYKIKLLFAWCLII GMFNAPLNADQNTDIKDISPEDMALNSVGLVSRDQLKIEIPK
ETLEQKVAI LNDYNDKNVNIKFDDISLGSFQPN DNLGINAMWGIQNL LMSQMMGDYGPNN
PFMYGYAPTYSDSSFLPPLGGY
>PROKKA_01288 hypothetical protein
MKQNLREQKLLKILENDVLTILDSFSNYLFELKEELDFIEEEMEGEITEQNLTTLYDFS N
FLEDHVN VFYENVLNID DVKTEHLYSGLIDSLNANLHFVKSFLSNQDLD FRFFKEINEGQ
DPQKTL SRLISLQSGKNDANSFKANNSFVSLVYVYAYFMLETIMQSYRILRLLEKPINNN
ISED MQSDIENFFVQANFLEYVQNKIYPTNHAYDFMHLIMDSIVPNW IQVDMSVEAKKK
ELFEKYFQNI DEVTNKMLDQENQNKNSD
>PROKKA_01289 Type IV secretion system protein virB4
MASKQADEQKKLVIEQE VQKRQFQKIEELKADMQKGVNPFFKVLFDGGNRLFGFPETFIY
SSIFILFVTIVLSVILFQAYEPVLIVAIVIVLVALGFKKDYRLYQRMERAMFKKPF LFK
GVKNKAFMSIFSMKPSKEMANDIHLNPNREDRLVSAANSYLANNYECFLDDGVILTNNYS
LLGTIKLGGIDFLTTSKKDLIELHASIYSVFRNFVTPEFKFYFHTIKKKIVIDETNRDYG
LIFSND FMRAYNEKQKRESFYDISFYLTIEQD LLDTLNEPVMNKKHFADNNFEEFQRIIR
AKLENFKDRIELIEELLSKYHPIRLKEYTKDGV IYYSKQCEFYNFLVGMNEAPFICNRKDL
YLKEKM HGGVKEVYFANKH GKILNDDLSEKYFSAIEISEYAPKSQSDLFDKINALDSEFI
FMHAYS PKNSQVLKDKLAFTSRRIIISGGSKEQGM TLGCLSELVGN GDITLGSYGNLVL
FADSF EKMQSVKECVSSLNAKGFLANAATFSMENYFFAKHCSFITLPFIFDVT SNNFAD
FIAMRAMSFDGNQENNAWGN S VMTLKSEINS PFYLN FHMPTDFGSASAGHTLILGSTGSG
KTVFMSMTL NAMGQFVHNFPANVSKDKQKLT MVYMDKDYGAYGNIVAMGGEYVKIELGTD
TGLNPF AWAACVQKT NATMEQKQT AISVVKELVKNLATKSDEKDENGNSISFSLADSN TL
AAAVTNLITGDMNLDYPITQLINAFGKDHNDPNGLVARLAPFCKSTNGEFQWLF DNKATD
RLDFS KTIIGVDGSSFLDNNDVSPFICFYLFARIQEAMDGRRFVLDIDEAWKYLGD PKVA
YFVRDMLKTARKRNAIVRLATQ SITDLLACPIADTIREQCPTKIFLRNDGGNLSDYQRLA
NVTEKEFEIITKGLDRKILYKQDGSPSVIASFNLRGIPKEYLKILSTDTV FVKEIDKIIQ
NHSIIDKYQALRQMYQQIEEY
>PROKKA_01290 hypothetical protein
MINNNNSNKKLRGFFLKVLLSLVVFSSYGSAND DKEAKKEALEKEKNTPNGLVYTNLDFD
SFKATIKNLKDKKVTFKEVNPDI IKDEVDFVIVNRVLKKIKDLKHYDPIIEKIFDEK GK
EMGLNVELQINPEVKDFFTFKSISTTNKQRCFLSLRGETREILCDDKLYNVLLAVFNSYD
PNDLLKHISTVESLKKIFYTITCEAVYL
>PROKKA_01291 hypothetical protein
MKFFTRITDSYKKVVVTLGLVMTTNPLMAVTSPAAGVTETKSLVIQIISVLAIVGGCALG
VKGIADIWKISDDIKRGQATVFAYA QPIAMLAVAGGIIFLSTKFGFNIGESGGAS
>PROKKA_01292 4-hydroxy-tetrahydrodipicolinate reductase
MKIGVYGASGRIGKLLLEELKGGYKGLALSSVFVRQKCETDFSSFSHAPLVTNDLKAFVR
ACECVIDFSLPKGLDNLLEALLECPKILVSGTTGLEKETLEKMQQALAKAPLLHVHNMSI
GIMMLNQLAFLTSLKLDADIEIVETHHNLKKDAPSGTALSLEYETCAKARGYDEKNALAT

HREGLRSKESIGIAALRGGDVAGKHTIGFYLEGEYIELSHTATNRSIFAKGALEVALWLK
DKTAKKYEINEMFG

>PROKKA_01293 putative FAD-linked oxidoreductase
MLEQQHIKYFKNLVGGEDFFTDLAHLNAYCYDATKERHLP SGVIFPRNEQEISQILKYCN
EHRIIIVPRGAGSGFTGGALS VSGGLVLSVEKHLDKILEIDTKNLIARVEPGVINKHFQN
EVEKLNLFYPPDPASENQSTLGGNVAENAGGMRAAKYGITKDYVMALRVVLANGEIIRAG
KKTIKDVAGFNVAGLMIASEGCLGVISEITLKLAKPPLKQSAMGVFNHIEDAMNAVYKT
MSSGVTVPVAMEFLDNLSIRAVEERFSKGLPKNAGAILITQVDGVVKEQIAWQLNEIEKHF
KANGCVGFKIAQNEQEEQDLWFSRRNASQSISVYGKKKL NEDVTVPRASLPSLLQEVAKI
SQKYGFKIPCFGHTGDGNVHVNIMLEDPKKDLEKGHKAMEEIFQAAISLEGTLSGEHGIG
LSKAKFMPLAFNHSEMELFRNIKKALDPNNILNPFKMGL

>PROKKA_01294 Plasminogen-binding protein PgbA
MLRLLIGLLLSFISLQSASWQEPLRV SIEFVDLPKKIIRFPAHDLQVGEFGFVVTKLSD
YEIVNSEVVIIAVENG VATAKFRAFESMKQSHLPTPRMVARKGDLVYFRQFNNQAF LIAP
NDELYEQIRATNTDITFISDDLVTFLNGFNPKIANLRKACNVYSVGVYIYIVTTNTLNL
SCESFEILEKRELDTS GVTKTSTPFFSRVESIDAGTLGKLFSGSQSKNYFAYYDALVKKE
KRKEVRIEKKEVRIDARENKREIKQEAIKEPSDGNNATQQGEKQNAPKKNKAQKEEKKPN
SKEEKRRLLKEEKKKAKAEQRAREFEQRAREHQRERDEKELEERRKALEAGKK

>PROKKA_01295 GDP-mannose pyrophosphatase NudK
MSYFKNAFNQKSLIDDSSVYLEPCSSSNFIELKRMHYNEENTKKTWDIIKSLDSVAVLLY
EKESDCFVIVKQFRPAIYARRFHF KRDDQNDIDGYTYELCAGLVDKANKSLEEIACEEAL
EECGYQISPKNLETIGQFY SATGLSGSLQTLYYAEVHKNLRVSKGGGIDTEKIEVLFLE
SKALDFIMDFQYAKTTGLSLAILWHLKKFKNV

>PROKKA_01296 Murein DD-endopeptidase MepM
MVFFHKKIILNFIYSLMVAFLFHL SYGVLLKADGMAKKQTLLVGERLVWDKLTLLGFLEK
NHIPQKLYYNLSSQDKELSAEIQSNVTYYTLRDANNTLIQALIPISQDLQIHIYKKGEDY
FLDFIPIIIFTRKEKTL LLSLQTSPIYQDIVKATNDPLLANQLMNAYKKSVPFKRLAKNDKI
AIVYTRDYRVGQAFGQPTIKIAMISSRLHQYYLFSHSHNGRYYDSKAQEVAGF LLETVPKY
TRISSPFSYGRFHPVLKVRRPHYGV DYAACHGSLIHSASDGRVGFIVKVGYGNVVEIHL
NELRLVYAHMSTFAKGLKKG SFVKKGQIIGRVGSTGLSTGPHLHFGVYKNSRPINPLGYI
RTAKSKLHGKQREV FLEKAQHSKQKLEELLKTHSFEKNSFYLLLEGF

>PROKKA_01297 R.PabI restriction endonuclease
MDFCF SILELKT TTPLLNRTAALKEHALLTIHKTNALMFLEMLKIFGLLSQAHHNDVLKI
LEKILQN

>PROKKA_01298 hypothetical protein
MSLIKVSGDKKVIEFSIPLTSISGKARVKIRHAFSDYGISTATRKIPFSLKH YVEWQIGY
DAPIKDKEKFELTTLKDEKYHFLGANNKVKTLYELSEMIYYAKQLNLISLEI

>PROKKA_01299 hypothetical protein
MNLGAYYTPCYLVDCTYKLLKKHVGIEDY TILDTACGNKEFLKLHHPKKIGADIDPKCDA
LIINALVNP KRENYGISQDEPLIIVGNPPYNDRTSFAKQKLKDKNFHFEIDHHLKSRDLG
ISFLKSFAILKPAFICVLHPLSYLIKEANFKQLKLFKDNRYRLDAFV VSSKSFTKSNEFP
IVIALYERGRMDYAGIRRFVFTDCDTTLC LNDFDYIANYVDKYPNAKKVGACVGYFFPM
RDINALKRNKTFLNAPSENAVRI SQDKLIYYQYIHYFKEIAPKIPY YFGNLDIIIDHF AF
LEIKDAFLKDKRARLEYFKKLFQGHPCFD

>PROKKA_01300 DNA gyrase subunit B
MQNYQSHSIKVLKGLEGVKRPGMYIGDTNVGGLHMMVYEVVDNAVDESMAGFCDTINIT
LTDEGSCIVEDNGRGIPVDIHPTEKIPACTVVLTILHAGGKFDNDTYKVSGGLHGVGVSV
VNALSKRLIMTIKKEGQIYRQEF EKGIPISELEIIGKTKSAKESGTTIEFFPDESVM EVV
EFQAGILQKRFKEMAYLNDGLKISFKEEKTQLQETYFYEDGLKQFVKDSTKKELLTP IIS
FKSMDEETCTSIEVALAYADDYNENTLSFVNNIKTSEGGTHEAGFKMGLSKAILQYIDNN
IKTKESRPISEDIKEGLIAVSVLKMSEPLFEGQTKSKLGSSYARALVSKLVYDKIHQFLE
ENPNEAKIIANKALLAAKAREASKKARELTRKKNLSVGTLPGLKADCQSKDPLESEIFL
VEGDSAGGSAKQGRDRVFQAILPLKGKILNVEKSHLSKILKSEEIKNMITAFGCGIQESF
EIEKLRYHKIIIMTDADVDGSHIQTLTLM TFFYRYLRPLIEQGHVYIAQAPLYKYKKGKTE
IYLKDSVALDHFLIEHGINSVDIEGIGKNDLMNLLKVVRHYRYT LLELEKRYNLLEVLRF
LIETKDALSFD MKVLEKSILEKLEGLNYQILRSFATEESLHLHVQTPKGLVEFNLD DNLF
KEVLFEEAHYTYQKLMEYNLD FLENKDILAFLEEVENHAKKGANI QRYKGLGEMNPNDLW
ETTMHKENRSLIKLKIEDLEKTD AIFSLCMGDEVEPRRAFIQAHAKDVKQLDV

>PROKKA_01301 DNA polymerase III subunit beta
MKISVSKNDLENALRYLQAFLDKKDASSIASHIHLEVIKEKLFKASDSDIGLKSYIFTQ
SSDKEGVGTINGKKFLDIISCLKDSNIILETKDDSLVIKQNKSSFKLPMFDADEFPEFPV
IEPKVSIENVAPFLVDAFKKIAPVIEQTSHKRELAILMQFDQKHQTLSSVVGTDTKRLSY
TQLEKISIHSTEEDISCILPKRALLEILKLFYENFSFKSDGMLAVIENETHFTFTKLIDG
NYPDYQKILPKEYISSFTLGKEEFKESIKLCSSLSSTIKLTLEKNNALFESLDSEHSETA
KTSVEIEKGLDIEKAFHLGVNAKFFLEALNALGTTQFVLKCNPPSPFLIQESLDEKQSH
LNAKISTLMMPITL

>PROKKA_01302 Phospholipase A1 precursor
MIFIVCQLEGKKFSQDNFKVDYNYLRLKQDLHIKTQNDLSNAWYLPPQKAPKEHSWVDF
AKKYLNMMDYLGTYFLPFYHSFTPIFQWYHPNINPYQRNEFKFQISFRVPVFRHILWTKG
TLYLAYTQTDWFQIYNPQPSAPMRMMNFMPELIYVYPINFKPFGGKIGNFSEIWIGWQHI
SNGVGGAQCYQPFNKEGNPENQFPQPVIVKDYNGQKDVWGGRSVSAGQRPVFRVLWE
KGGLKIMVAYWPYPYDQSNPNLIDYMGYGNKVDYRRGRHHFELQLYDIFTQYWRYDRW
HGAFLRGYTYRINPFVGIYAQWFNGYGDGLYEYDVFSNRIGVGIRLNP

>PROKKA_01303 Sodium:neurotransmitter symporter family protein
MGKFSKLGFIATLGSSIGLGHWRFPYMTGVSGGAFVLLYLVLTLSLGIAMLLVEMLI
GNLGKKDVVSNYQILDPRKKYYPFTSFFILGGPLILSFYAVVLGWVLYLFFVVTFDLPK
DLEQAKMQFSMLQNGSLIWPVIGFSACLLPTIWFVSRGIEEGIEKLNVLMLPLLFVIFIG
LLIYAMTLESMPKALHFLNFIEIQIDFKVMDALGQMFFSLSLGVGTIITYSAFTPKKE
NLLKSSLFIVLPGILISLIAGVMIFTVFVEYHADVSQGPGLVFISLPLTFKMGMSGQIV
SLFFFMALVFAGITSTVSLIEPLALYLINRFNFSRTQASLWIGIVVYVLGVVLILSMNER
YAKFLSFANKSVFGWLDFITSSFLMPLGGLFSVLFGWILNKKRSFLATKHFFNINAFKA
WHFSVRFIAPVVILAIFILQFK

>PROKKA_01304 Sodium:neurotransmitter symporter family protein
MGNHFSKLGFLAALGSAIGLGHWRFPYMTGVSGGAFVLLFLFLSLSVGAAMFIAEML
LGQSTQKNVIEAFKELDLNPKKRKYAGILLISGPLILTFYGTILGWVLYLVSVSFNLP
NSIQESEQIFNDTLQSIRLQSIGLFSVLFITGWIVSRGIEGIEKLNVLMLPLLFATFFG
LLFYAMSMDSFSAFHMFDFKPKDLTSQVFTYSLGQVFFSLSIGLGINITYAAVTDKTQ
NLLKSTIWWVLSGILISLVAGLMIFTVFVEYGANVSQGTGLIFSLPVVFGQMGAIGVLV
SILFLLALAFAGITSTVALLEPSVMYLTEKYQYSRKFVTWGLVALIFIVGVVLIFSLHKD
YKDYLTFFEKSLFDWLDFASTIIMPLGGMATFIFMGWVLKKEKLRLLSVHFLGPKLFAT
WYFLLKYITPLIVFSIWLSKIY

>PROKKA_01305 Acyl-CoA thioester hydrolase YbgC
MRCRVYYEDTDEGVVYHANYLKYCERARSEFFKQNVLPENEEGVFVIRSIKADFFTPA
SLGQVLEIRTQIKELRKVFVFLFQEIYCIQNASLEPMKPFKVFASEIKFGFVNRSTYSPI
AIPKLFKELLNAV

>PROKKA_01306 hypothetical protein
MPSDSKKPTIIYPCVWDYRVIMTTKDTSTLKKLLETYQRPFKLEFKNTSKNAKFYSFNVS
MEVSSEEERNEIFQKISQLEVVVHAL

>PROKKA_01307 UDP-N-acetylmuramoylalanine--D-glutamate ligase
MKISLLGHGKTTLALARFFKKNHNEVKFFDDQFTSFHKDSEGFLCYPSKDFNPNDSQL
VSPGISFTNPLAIKAKHLVSEYDYIDSLFDSVFTPTIISISGTNGKTTTMITMLLEDF
KAVSGGNIGTPLIELFEKQSPLWVLETSSFSLSHYTNKAYPLIYLLINVEADHLTWHCNFE
NYLNAKLKVLTLMPKTSAILPLKFKEHPIIQNSQAQKIFFDKSEEVLERLKIPSNALFF
KGAFLDLAALALLVYEQFLKIKNLKWQDYRENALKRLNAFKIGSHKMEEFKDKEGRLWVD
DSKATNIDATLQALKTFKNQKIHLILGGDIKGVNLTPLFEEFKNHEVSLYAIGSSAFIIQ
ALALEFNVSCQVCLELEKAVQEIKSVLQNEIALLSPSAASLDQFSSYKERGEKFKAFVL
KD

>PROKKA_01308 Phospho-N-acetylmuramoyl-pentapeptide-transferase
MAILGDFMLYSLLYGYFNINLFQYLTFRAGLGFFIAFFLTFLMPKFLWAKAKKANQPI
SSFVPSHQKKDTPTMGGIVFVFATIVASVLCASLGNLYVLLGLIVLVGFSFVGFRDDYT
KINQQNNAGMSAKMKFGMLFILSLIVSVLLSLKGLDTFLYAPFLKNPLFEMPTMLAVGF
VLVFLSTSNVAVNLTDGLDGLASVPSIFTLLGLSIFVYVAGNAEFSKYLLYPKVIDVGELF
VVSALVGSLSFGFLWYCNPNASVFMGDSGSLALGGFIAYNAIVSHNEILLVLMGSIFVIE
TSLVILQVGSYKTRKKRFLMAPIHHHFEQKGWAENKVIVRFWIIISMLSNLVALLSLKVR

>PROKKA_01309 Neuraminylactose-binding hemagglutinin precursor (NLBH)
MERSLIFKKVRIYSKMLVALGLSSVLIGCAMNPSAETKTPNNAKNQVPOTHERIQTSSSEHV

TPLDFNYPVHIVQAPQNHVVGILAPRIQVSDNLKPYIDKFQDALVNQIQTIFEKRGYQV
 LRFQDEKALNAQDKRKIFSVLDLKGWVGILEDLKMNLDKDPNNPNLDTLVDQSSGVSWFNF
 YEPESNRVVHDFAVEVGTFQAMTYTYKHSNSGGLNSSNSIIHEDLEKNKEDAIHKILNRM
 YAVVMKKAVMELTEENIAKYRDAIDRMKGFKSSMPQKK
 >PROKKA_01310 50S ribosomal protein L28
 MAKRCALTFKGSMIGNHVSHANNKNKRLLPNLRSIKIQLDDGTTKRIKVAASTLRTMRK
 GA
 >PROKKA_01311 Voltage-gated potassium channel Kch
 MFEKLFFFKIKKDDIEIQPEVNLNSEIYEQFKVFRPLPLILIQLLVLLGTLGYFALENYSLM
 QAFFQTYYTMTATGFGALNESQFGPISIFLTSILMFFGAGIIAFSVAILVSVVNKGTLTR
 LIKEKGMIIYKIARLKDHYVICYHNEYTIELSKQFRSAQIPFVVVDNDPSFEEEAIAKHYP
 YIIIGDPHTNLAMLKTHLSSARGVVALSKILPVNVALMVSVRLFEKELKRKPYIIASAH
 SDEGLEKLKLLGADMVVSPTKLMAQRVSAMAVRPDMENILERFINKKDTLLDLEEVIVPK
 TSWLVLRKLKEAHFREIAKAFVIGITQKDGYIPMPDGETIIASESKLLMVGTSSEGVATC
 KQLIANHQKPKEVDYISL
 >PROKKA_01312 F0F1 ATP synthase subunit B
 MINYPNLPNSTLEISEQPEVKEITNELLKQLQSALKSNALFTDQIELSLKGIVRILEVLL
 SLDFFKNANEIDSSLRNSIEWINNAGESLAKMKKEYEGFFSDFNTSMKANEQEVNTNLNA
 NTENIKSEIKKENQLIETTTKLLTSYQIFLNQARESATTQISNNQTQAITNINEAKNNA
 NNEINTNKTQSLEAITQAKTNANNEINTNKTQSLEAITQAKTNANNEITANKTQAITNIN
 EAKENATTQINANKQEVLLNITQEKQTATSEIIEAKKTAFNELLETLPKPKFSGLFAGVYY
 IRNVIYIQGGWESKVKDLSDYKLEKSKKGE
 >PROKKA_01313 hypothetical protein
 MVCLSVSFTNKPFLWLWFLVVGCYLVYEWQKKQKDFQSAKSLKFDSVSELEKDLKMEVTN
 DEWENIKKLNENLKIFFNVENVSIIFFLLVSYLIVGMRVTLSLIDNYGVLK
 >PROKKA_01314 hypothetical protein
 MLNFTPKSRYIYEKAIKEYYNIDLKPEDELLVRVPHNRLNNTTEINNLAASSNQGRFNSESD
 HALAVLSHYEAKLKELDKKLDADSIYSLKNIVAKNLNFDQATHPNVTDNLALLMFNMMPR
 TKTQGIELLRWQKEFSNDIKSYEKVKMFVDNAGSFHNLIDHMFNPKVSLNAYLSDIMD
 RSFANLKNYQSTSESLKDLSEKFYKTSSLEMFEKSDQSVSDITEILGGAIARFARFDDPS
 KALFEALKSDNIKKGLKEFKIADITKDMFDPKSKEFKDIDIYDFTHYLLMVNREPNNNP
 TLNRLIQAVKGMQKESEKGIKKPKLETPSEWGHNYSEFKNDGLGAINKLETKKGFVAGA
 FHKEGLGDIDLWVGTPKTKDSNGYGLAHILEKREKQYKRLGLTNEQAKERTKELLRQIPE
 VIDKDIKDNDYIGHVTIRHNGIEVGLSSQKGNTPLNHYMITSFETDEKVLRELETIGTL
 PNDYRDGINYSASNLPNPTTKKLKDD
 >PROKKA_01315 hypothetical protein
 MIEKASNKELELLKDANFKHPENIRASLDHDAITHILKRHGVSNSVVRNGEIPITNEDI
 ANYRYIVNNADAILRTIDKYNQEAITAFKQVNGYAVVVEQAINKKNELALKTMYKNNGSY
 KNEVYKEFSSTSLNANAKVPHRLSSYSGATENTTQKDLIDQEDLLKRTLENNQTTPKPT
 NLSPLEQANAEKLAKLESEALESEQEFKAKEQEAQRKEALKKKLEHEHGNAGNIESQTK
 IEVGKDIPTQIQAQIPKSRVRLNEREYIDLDAIVKAKDLKPSFTTGGTQKRTDMNEE
 >PROKKA_01316 Holliday junction ATP-dependent DNA helicase RuvA
 MIVGLIGVVEKISALEAHIEVQGVVYGVQVSMRTSALLEVGQKARLKILQVIKEDAHLLY
 GFLEESEKILFERLLKINGVGGRIALAILSSFPNEFESIIATKEVKKLQQVPGIGKKLA
 DKIMVDLIGFFIQDENRPARNEVFLALES LGFKSTEINKVLKTLKPHLSIEAAIKEALQQ
 LHS
 >PROKKA_01317 hypothetical protein
 MSLERFAPIKVERCKDIQKELKKVAAENKLQTEDLWFEILKTSIFIKNSAKDDFSEAFGG
 ELQQLEDEYYEKKELTYQTHDIKIKSNAYKRFFEVKVDENLSKIEIILDEC FIVLDTE
 EHYQEMFAYIKECLAFQGVVFRHLSQMYENLKTRELRYQKEAQNKRFILYASSTFIPNAE
 EQSHFLLEEYLPHTIFLSDQEEFVKENYYIAKENQKVACVSYPKQGRDGRNLKGLYI
 ELPKVANSPTPIGHDKNAFEEREENNALVYYSKALQGVKMEKGRLVSKQNFIFKNGIKSI
 EVPNLLGGVESGLALEIQAKDELSDAIDSNLILEASVINVKGNVGNVILVAKEITIEGQ
 IHPEYVYANQARITNHKGVCYAKEFECKYLERAKVYANSVKVEASAGSVVYAKEIALEK
 LKSDNKLYFSKQCWIDEVDGNGNRFIIFYAFGGRENQEELKNAKQKLNALGLKSKKIIAQH
 QSLNHLVKNHQAIMEKLKNATEEIKRSLMQQESVKDAYSEFMFALKRLKILKAQMLELQK
 INNECYAKLISIENSFQHASITTKNPFQENIVIYHRNYPKVSNSTAMLSHNESVNVIYE
 DHKIKKIPKSAIKG

>PROKKA_01318 putative peptidoglycan biosynthesis protein MurJ
MLKKIFLTNSLGILCSRIFGFLRDLMMANILGAGVYSDIFFVAFKLPNLFRRIFAEGSFS
QSFLPSFIRSSIKGSFASLVGLIFCSVLLVWCLLVALNPLWLAKLLAYGFNEETLKLCA
IVAINFWYLLLVFITTTFLGALLQYKHSFFASAYSASLLNLCMISALLISKEKTHLEALYY
LSYGVLGGVAQILLHFYPLVKLGLNLLWKGFLSFKTKNAAKKKYRSKRAKKDLKAFFK
QFLPSVLGNSSTQIASFLDTTIFLASGSVSYYANRVFQLPLALFAIAISSALFPSI
AIALKNNQODLVLQRLQKAWFFLVGVLLLCSIGGIMLSKEITELLFERGQFSPKDTLITS
QVFSLYLLGLLPFGTLKFLSLWLYAKLEQKKAAKISLISLFLGLAASLSLMPLLGVLGLA
LANSLSGLFLFVLTIKAFGFQLFLGIIKNLKS WLVLFLACVEILLLLAFKSWVTHLYLF
YYFQGF

>PROKKA_01319 Cysteine--tRNA ligase
MFIYDTKSKQKVPFEPLIQNKANIYVCGPTVYDDAHLGHARSAIAFDLLRRTLELSGYEV
VLVRNFTDIDDKIINKALKENKSIQELSSIIYESYTRDLNALNVKKPSLEPKASEYLDAM
VGMIE TLLEKNIA YRVSNNGDIYLDTSKDKDYGSLSVHNSSIEFGRIGLVQEKRLEQDFVL
WKS YKGDNDVGFDSPLGKGRPGWHIECSSMVFETLALANTPYQIDIHAGGADLLFPHEN
EACQTRCAFGVEIAKYWMHNGFVNINNEKMSKSLGNSFFVKDALKNYDGEILRNYLLGVH
YRSVLNFNEEDLLMSKKRLDKIYRLKQRVLTGGINPNFKKEILECMQDDLNIKALS
VLESMLSSTNEKLDQNPKNKALKGEILANLKFIEELLGIGFKDPSAYFQLGVSESEKQEIE
NKIEERKRAKAQKDFLKADHIREELLKQKIALMDTPQGTIWEKFF

>PROKKA_01320 Vacuolating cytotoxin autotransporter precursor
MEIQQTTHRKINRPLVSLALVGALVSITPQQSHAAFFTTVIIPAIVGGIATGTAVGTVSGL
LSWGLKQAEAEANKTPDKPKDVWRIQAGRGFNNFPNKEYDLYKSLLSSKIDGGWDWGNAAR
HYWVKDGQWNKLEVDMQNAVGTYNLSGLINFTGGDLVDNMQKATLRLGQFNGNSFTSFKD
GANRTTRVDFNAKNILIDNFVEINNRVGSAGRKASSTVLTQLQASEKITSRENAEISLYD
GATLNLVSSSNQSVLDYGVVWMLQYVYGAYLAPSYSTINTSKVQGEMNFRHLAVGDHNA
AQAGIIANKKTNIGVLDLWQSAGLNI IAPPEGGYKDKPSNTTQNNANNNQQNSAQNNSNT
QVINPPNSAQKTEIQPTQVIDGPFAGGKDTVVNIFHLNTKADGTLRAGGFKASLSTNAAH
LHIGEGGVNLSNQASGR TLLVENLTGNITVEGTLRVNNQVGGA AVAGSSANFEFKAGEDT
NNATATFNNDIHLGKAVNLRVDAHTAYFNGNIYLGKSTNLRVNGHSAHFKNIDASKSDNG
LNTSTLDFSGVTDKVNINKLTTSATNVNIKNF DIKELVVTTTRVQSFGQYTIFGEIIGDKS
RIGVVS LQTGYNRAYSGGVTFKSGKKLVIDEIYHAPWNYFDARNVTDVEVNKRILFGAPG
NIAGKTGLMFNNLTLSNASMDY GKDLDTIQGHFTNNQGT MNLFVQDGRVATLNAGHQA
SMIFNNLVDSATGFYKPLIKINNAQNLTKNKEHVLVKARNIDYNLVGVQGASYDNISASN
TNLQEQFKERLALYNNNNRMDICVVRKDNLDIKACGMAIGNQSMVNNPNKYLEGKAW
KNTGINKTANNTTIAVNLGNNSTPTENGGNTTNLPTNTTNKARFASYALIKNAPFAHSAT
PNLVAINQHDFGTIESVFELADRSKIDITLYTHSGAQGRDLLQTLIDSHDAGYARTMID
ATNANEITQQLNAATTTLNNIASLEHKTSGLQTL SLSNAMILNSRLVNL SRRHTNHIDSF
AKRLQALKDQRFASLESAAEVLYQFAPKYEKPTNVWATLLGERA

>PROKKA_01321 Vacuolating cytotoxin autotransporter precursor
MGNAIGGTS LNNGSNASLYGTSAGVDAYLNGEVEAIVGGFGSYGYSSFSNQANSLNSGAN
NTNFGVYSRIFANQHEFDFAEQGALGSDQSSLNFKSALLQDLNQSYNYLAYSAA TRASYG
YDFAFFRNALVLKPSVGVSYNHLGSTNFKSNSNQVALKNGASSQH LFNASANVEARYYYG
DTSYFYMNAGVLQEFARLGSNNAASLNTFKVNATRNLNTHARVMMGGELQLAKEVFLNL
GFIYAHNLISNIGHFASNLMGRYSF

>PROKKA_01322 putative ABC transporter ATP-binding protein
MVLEVKNLSFKYSQKLILDKLSFSVPKNSITSILAPNGSGKTTLKCLLGLLKPLEETEI
KACNKDILPLKPYEKAKLIAYIPQVEYYAFNFSVLDFVLMGKATHLNLFAMPKAKHIKEA
TSVLERLDLES LKDQGINDLSGGQRQMVLLARSLLQRTPLLLLDEPTSA LDKNQALFFD
AIKDEMKKRELSVLVNIHDPNLVARHSTHVVMLKDKKFLQLASTPIAMTSHNLSALYDTP
LEAIWHDNKL VVYAL

>PROKKA_01323 Hemin transport system permease protein HmuU
MLKTYHIALACVILAVVLLFGGESLSFEEWQEVCLNVKNHFLHNEELSSLSIIILEIRL
PRVILALLVGASLSGSGVVMQTIFRNPLVDPFLLGISSGAMLGVAMAIAVVESNIAILAF
FGAILASLAVLAMNRVLGNSVLSLVLSGVVLSAFLSALAGAIKFFVIPQKAQAI VVWLLG
SLSLSSYKDCLIAFIGLSLGFIPFLLRWRINLLSLSDAQSLSLGINPVLLRSLCLVCVS
VASALAVSVSGTIGWIGLVIPHVARLFFGANLQKLLLSSLLMGAFFLLADVVAKTITPY
DLPVGIATSVLGAPFFLWLLFRTRGV

>PROKKA_01324 Cyclopentanol dehydrogenase

MGEKKESQKVAVITGASSGIGLECALMLLDQGYKVYALSRHATLCVALNHALCESIDIDV
 SDSNALKEAFLSISAKEDHCDVLINSAGYGVFGSVEDTPIEEVKKQFGVNFFALCEVVQF
 CLPLLKNKPYSKIFNLSSIAGRVSMFLFGHYSASKHALEYSDALRLELKPFPNVQVCLIE
 PGPVKSNWEKTAFSVENFESEDSLYALEVDAAKSFYSGVYQNALSAKEVAQKIVFLSMSh
 KIKARYLIGLKTQLLLALYQILPSSWYDSLFRLLIVLKRKRDA
 >PROKKA_01325 acyl-CoA esterase
 MPQIQSSHYSHFDFETIDTADRTKLLMSYLVVPTTANFNVMHGGELLNLLDKVAYVCSTR
 YCAKGTVTLSVDGVTFKYPIPVGNLLTFLASINYVGNTSCEVGIKVLSEDIKTREITHTN
 SCYFTMVAVENGKPTPMPKYEPKTEVEIRRYEGALKRKEMRTRGYLKSGKHGEV
 >PROKKA_01326 mRNA interferase YafQ
 MREKRLWVLKLNLLKKSFKQKDFDKLLNGFDDSVLKEVILRKKEPLDPQFQDHALKGKWKP
 YRECHIKPDILLVYLKDDDELILLRLGSHSELFCKSPTQFITLKPTAQPTTPLF
 >PROKKA_01327 hypothetical protein
 MPNTTNKDYTYQSEKQLFNLIHQLEQKIKKMQNDRVSFKEKMAKELEKRDQNFKDIDAL
 NELLQKISQAFDDKRDCCCLGRKIPNIQTQQAMRDVYNKETDLIVEDFSSYSDERKKALGV
 ETQS
 >PROKKA_01328 Helicobacter outer membrane protein
 MKKHILSLALGSLLVLSALSAEDDGFYMSAGYQIGEAQMVKNKTKGIQELSDNYEKLNNLL
 TRYSTLNTLIQLSADPSAINLARENLGASAKNLIGDTKTSPAYQAVLLAINAAVGFWNVL
 GYATQCGGNANGTKSISSTTIFNNEPGYRSTSITCSLNRYTPGYYGPMsidnmkklneay
 QILQTALKQGLPALKENNGKVNVTYTYTCSGEGNNNCXAPVSNGGTKTETQTIDGKSVTT
 TISSKVVDSKAKGNTRGVSYTEITNKLDGVPDSAQALLAQASTLINTINTACPYFSVTNX
 SSGPQMEPTRGKLCGFTDEISAIQKMITDAQDLVNQTSAINANSQSTSVGESNKPFNPFT
 DASFAQGMLANASAAQAKMLDLANQVGQAINPSNLSGTFKNFVTSFLATCNPSTSGANTT
 QGSAPGTVTTQTFASGCAYVQETIANLTNSIAHFGTEASQIQQAADIADTLVHFkskyse
 LGNTYNSITTALSKVPNAQSLQNVVSKKNNPYSPQGIETNYLQNSYNQIQ TINQELGR
 NPFRKVGIVSSQTNNGAMNGIGIQVGYKQFFGQKRKWGARYYGFFDYNHAFIKSSFFNSA
 SDVWTYGFGADALYNFINDKATNFLGKNNKLSVGLFGGIALAGTSWLNSEYVNLATVNNV
 YNAKMNVANFQFLFNMGVRMNLARPKKKDSHAAQHGIELGLKIPTINTNYYSFMGAELK
 YRRLYSVYLYNVFAY
 >PROKKA_01331 Hydrogenase isoenzymes formation protein HypD
 MSVDHLISPFDRKRTLLALSDAIKKLAFKLEKKLVIMEVCGGHTHSLMKYGLLDLMPKNL
 EFVHGPVCVCMVPRARLDEAYELATIKDSIILSLGDMMKVPGSYGSLIQAREKGLDARF
 LYSPMQALEIAKENPTKKVIYIAIGFETTTTPMSASVLLNAKKEKLKNLFFHINHILVPPS
 VSAILKDPACQINALLAPSHVSVISGAQIYSPLVDRFKLPPIIVSGFEPVDILESVLMLLK
 QALNKEAKLEIQYKRAVSFEGNVKAQALVNACMEVRENFEWRGLGNIKRSALKLKEAFAS
 YDAEKVFKEYLSHKIIKENKACKCGEILKGIAPLDCSLFATTCTPQNPIGSCMVSSEGA
 CAAYYRYKRV
 >PROKKA_01332 Hydrogenase isoenzymes formation protein HypC
 MCLAIPSKVIAINDNVAILETLGVQREASLDLMGESVKVGDYVLLHIGYVMSKIDEKEAL
 ESIELYQEMIakmnethe
 >PROKKA_01333 Hydrogenase/urease nickel incorporation protein HypB
 MSEQRKESLQNNPNLSKKDKIVEKILSKNDIKAAEMKERYLKEGLYVLNFMSSPGSGKT
 TMLLENLADFKDFKFCVVEGDLQTNRDADRLRKKGVSAHQITTGEACHLEASMIEGAFDLL
 KDEGALEKSDFLIENVGNLVCPSSYNLGAAMNIVLLSVPEGDDKVLKYPTMFMCADAVV
 ISKADMIEVFNFVRVSQVKEDMQKLKPEAPIFLMSSKDPKSLEDFKNFLEKKRENYQSTH
 SF
 >PROKKA_01334 hypothetical protein
 MMKAVFMNTWYLLLLTIILLIIIIWRIIRLVLKNKDDSNERD
 >PROKKA_01335 hypothetical protein
 MKMVHFLESVHFELHIETLNENSSNKEIRICMPKGAVMDKHKAPGAISVQVLEGKIIFE
 VENEKIEMPKGALISLEAQVLHRLDALENSVIRLSLSKK
 >PROKKA_01336 Acetate kinase
 MRNIEARKQKGDKEAKLAFEMCAYRIKKYIGAYMAVLKKVDAILFTGGLGENYSALRESV
 CEGLEDLGIALHKPTNDNPGNGLVDLSQPDAKVKVLRIPPTDEELEIALQAKEIVEKLK
 >PROKKA_01337 Acetate kinase
 MQEIGKLSILAPLHN PANLAGIEFVQKAHPHIPQIAVFDTAFHATMPSYAYMYALPYELY
 EKYQIRRYGFHGTSHHYVAKEAAKFLNIAYEEFN AISLHLGNGSSAAAIIQKGKSVDTSMG

LTPLEGLIMGTRCGDIDPTVVEYTAQCANKSLEEVMMKMLNHESGLKGICGDNEKHRSQKT
KRR

>PROKKA_01338 Acetate kinase

MEILVLNLGSSSIKFKFLFDMKENKPLASGLAEKIGEEIGQLKIKSHLHHNDQELKEKLVI
KDHASGLLMIRENLTKMGIIKDFNQIDAIGHRVVQGGINSMPQS

>PROKKA_01339 Phosphate acetyltransferase

MQGLWIYPEDTEVLGFACKSLLKALTTRYQKIALFSPIDGGRESFWERYGLNSLEFHSAI
DKQKALELVSTAQEELLFETILKRYDELQTHDFVINLGYAPKFFLNALLDLNTILAKHL
NAPIVAVAQTSLEYLKAMHFHILKKEAPFAVGLFAGEMPEKPDFLSASLCKQQCELEASA
IENLLQTKSEIITPLAFQMSLEKKAKKQIKKVVLPESEDERILKAAHRLNAMGAVGLILL
GDKEAINSQAKNLNLLENVEIIDPNTSHYKEEFANHLIELRKSGLSEQEAQQLVLDKT
YFATMLVHSGYAHAMVSGVNHTTADTIRPALQIIKTNPVSLVSSVFLMCLDTQVLVFGD
CAIIPNPSPKELAEIAITSAQSAKQFNIAPKVALLSYATGNSAQGEMIDKINEALTIAQH
LDPQLEIDGPLQFDASIDKSVAKKKMPNSQVAGQASVFIFPDLNAGNIAYKAVQRSKAV
AIGPILQGLNKPINDLSRGALVEDIVNTVFD

>PROKKA_01340 Flagellar hook-length control protein FliK

MPSPINPIHTNANANANALINSKAKNEDTKNAPKSASKDFSKILNQKISKDKTAPKEDSN
ASKVTPKDAKEDAKALEKTPTLPHQHAQNPAKDQQAPTLDWLNHKKTTVSHEAQHETHE
ANETNPKTPNETLNKNEKKNPVGTVSNAHQTNLASKNPITPNHANNAIKNPPTPTHSAKEP
KTLKDIQTLQKHDNLASNIQATTTPENKNPLNASDHLALKTTQTPTNHTLAKNDAKNTA
NLSSVLQSLEKKEPHNKEHANPQNEKKTPLKEALPMNAIKRDKTSLSKKSEKTQTKAQ
TTAPSIATENAPKIPLKTPPLMPLTGANPPPNNAPTLLKEEETTKEASDNKEKTKESSN
SAQSAQNAQSSDKTSENKSAAPKETIKHFTQQQLKQEIQEQYKPPMSRISMDLFPKELGKVE
VVIQKVGKNLKVSIVSHNNSLQTFLDNQDLKNSLNLALGFEGVDLSFSQDSSKEQPKEQL
RELKFEQESSPLKENALKSYQENTDHENQETSMQITLYA

>PROKKA_01341 Basal-body rod modification protein FlgD

MAIDLAEVGTAKAAQERKKEQPTIANGLDKNAFMKLFLEQLKNQDPTAPMETDKIITQTA
QLTQVEMQEENKKTMQEVASAMKSNKETNESLKDFQGALKDTMENLNKGMDDSLKANNAL
REVTALNSVSMIGKIAETDVSGANFDGNNKLSFSLFFDEKIDASKGVPAIQILNENNELV
KTIPLKDYNGQKGYINFWDGTNEKGEKVPKGNKYIKAEYNLDSSHKQYLQTRIGRGEVE
SVIFDKGKPMLRMGEMILPIDSAIEFYKPDQKPLEQKPIDQKLSKQKPIEQKISEQKPID
QKLSKQKPIEQKISEQKPIEQKISEQKPLDQKPKQISSKETA

>PROKKA_01342 Flagellar hook protein FlgE

MNDTLLNÄYSGIKTHQFGIDSLSNNIANVNTLGYRSNDPEFKTLFSSSHLDALNAKSVVAN
DRNYGVTGSGNVLSNKDGEYMPSEGEFHMAYQGKGWVFIGPNKNGEMTINKDGFSSKKQDN
FLTRAGNFARDADGYLVTPEGYVYVYGIDLKKIKDGTLNSTARDEDIEKLHGNTLSPLQIP
QDLTYQPVVLSTKVGISVNLNPKDHLKGVQDFFLNDKGEIIKERFLNQDINALANDDNKPI
DAITNRKLNVSIOKEDGKKEDFVFTYGDAAKGENQFKTLGDLQKLLKEKTGLDLNLIKSE
KDAKSPPLLEIANPSETPITFSLSGGIADKLGLKADGMELKKGISRDSVAIKIPYYSTE
VDIYDKAGDKYLLQSEYYMTNSNDPTSSPTS KRKNQTWEVKSIIADPKNKTPIINDPTWEI
VGFD SATHKMK SAPMTLDFKGNKLTYSLDKSENHDSSDLSYQDSKLLEASQDGKPRGIFR
DMRIEENGVISLAFSNGVVEPVARIGILAF TNDQGLRKIGGNLYEMQEGTINGENRPLSG
NPILGWDEEGKLKFGKIRHKYLETSNVNAGNALTNLILMQRGYSMNARAFGAGDDMIKEA
ISLKK

>PROKKA_01343 hypothetical protein

MIPTQLNEIAEFLKTNPHYLSHPLQDGRNLSSVNEEEILNTIKHSFPIQLPKAREWWDFS
FEENDIFYPVNIKTTTTKTADNLNCKLGIYYALCGLLPEFNNEIAWEKYFQKLHKDLGKN
TNRDYYFLIINKNDPKDIFINSLKGIQTLQPNLFPQCKWDNNREIVQRDFTESKNFELS
ALAKSVTLRANIYLLKFKEVFGEFFE

>PROKKA_01344 Modification methylase TaqI

MNNLDIKTLGQVFTPKKIVDFMLTLKQNGQSVLEPSAGDGSFLKRLKKA VGIEIDPKICS
KNALCMDFFDYPLENQFDTIIGNPPYVKHKDIAPSTKEKLHYSLFDEERSNLYLFFIEKAI
KHLKPKGELIFITPRDFLKSTSSVKLNEWIYKEGTITHFFELGDQKVFPNAMPNCVIFRF
CKSDFS RITNDGLQFVCKKGILYFLNQSYTQKLSEVFKVKVGAVSGCDKIFKNEKYGNVE
FVTSITKRTNVLEKMFVNKPNDYLLQHKDSLMQRKIKKFENNNWFEWGRMHHSPPKKRI
YVNTKTRQKNPFFIHQCPNYDGSILALFPYNQNLDLQNLCDKLNAINWQELGFVCGGRFL
FSQRSLENALLPKDFFNLG

>PROKKA_01345 DNA helicase II

MLETQLNPEQLKAAKALKGYNLVIASAGTGKTSTIVGRILYLLDNGIKPEEILLTFTN
KASNEMARVAKYFKSSSKIEAGTFHAVAYRYLKEHYPNLSLKQPKELKKLLESIVDTKN
AIDDDKKPYTSQHLIALYSLYTNALKQEDFSAWLSSKNPEHTPYAALYENILEEFENTKK
KHDIYDNDLLLLFKKAMLETSPYKEVLCDEFQDTNPLQESILDAINPPSLFCVGDYDQ
SIYAFNGADISIISNFTQKYKNARVFTLTKNYRSSKEILDLANQVIQRNERIYPKNLEV
KSGKFNKPTLLNYNNNIAQCQDIKRIVMRKNFKEVAVIFRNNASADQLEAALRSHNVPS
KRKGSASFFESKEVALALDICALIFNPKDIMAAIHILSCISDIGSNTAKDIHEALMLLGN
GDLKLALTHPSKEAKIYTKKEITSMGLFEEIFALENSSRFNSVIDKAFHSHPVLMHPKI
SLNGAKTSLDFFTLTKAPTHSPSALIKHILESFAFFQTFKTRLLKERSKNKDGSYNEFKK
LQAQKRFNEKMDLLSSLAKNYQNLGRFLNGTLIGSSEATQGEVNNLLSVHASKGLEFKDV
YIIDLMEGFRPNHKLMMNTGGGIEEERLFFYVAITRAKENLWLSYAKNELRENAKPKEHKP
SVFLYEAGLLKPSK

>PROKKA_01346 Helicobacter outer membrane protein

MIKKNRTLFLSLALCASISYAEDDGGFFTVGYQLGQVMQDVQNPGGAKSDELARELNADV
TNNILNNTTGGNVAGALSNAFSQYLYSLGAYPTKLNNDVSNALLSGAVSGTCAAAG
TAGGTSLNQTQACTAAGYYWLPSTDRILSTIGSQTNYGNTNFPNMQQQLTYLNAGNVF
FNAMNKALEKNETGSDGQTYHQAIQYLQGGQNILNNAANLLKQDELLLEAFNSAVAANI
GNKEFNAAFTGLVQGIIDQSQLVYNELTKNTISGSVDNAGINPNQANAVQGRASQLPN
ALYNAQVTLTDKINALNNQVRSMYPYPQFRAGNSRSTNINLNGFYTKVGYKQFFGKKRNIGL
RYYGFFSYNGASVGFRSTQNNVGLTYTYGVGTDVLYNIFRSYQNRSVDMGFFSGIQLAGE
TFQSTLRDDPNVKLHGKINNTHFQFLFDFGMRMNFGLDGLKSNRHNQHTVEFGVVPTIY
NTYYKSAGTTVKYFRPYSVYWSYGYSF

>PROKKA_01347 Helicobacter outer membrane protein

MTQSQKVRFLAPLSLALSLSFNVPVGAEEEDGGFMTFGYELGQVVQVKNPGKIKAEELAGL
LNSNTTNTNINIACTGGNVAGTLGNLFMNQLGNLIDLYPTLNTSSIQQCSATNSGGSGS
GATTAAATTSNSPCFQGNLALYDEMVSIIKTLNQNISKNIFQGNNTTSANLSNQLSELN
TASVYLTYMNSFLNANNQAGGIFQNNNTQAYENGVTAAQIAYILKQASITMGPSGDSGAA
AAFLDAALAHVFNNSANAGNDLSAKEFTSLVQNIWNNSQNALTLANNANISNSTGYQVSY
GGNIDQARSTQLLNNTTNTLAKVTALNNELKANPWLGNFAAGNSSQVNAFNGFITKIGYK
QFFGENKNVGLRYYGFFSYNGAGVGNPTYNQVNLTYGVGTDVLYNVFSRSFGRSLNA
GFFGGIQLAGDTYISTLRNSPQLASRPTATKFQFLFDLGLRMNFGILKKDLKSHNQHSIE
IGVQIPTIYNTYYKAGGAEVKYFRPYSVYWVYGYAF

>PROKKA_01348 Putative outer membrane protein

MKQEKYFLTSSLSLLSFLLCVPAEAFDYRFSGRVENFSKIGFNNSQINTKKGIYPTESFID
IVTLAQVKVNLPLPKGTENHRLSVSLGGAIAAIPYDKTKYYINQANGKVFGSIVENFIGGY
HGYFFNKYLGPAYAGTSQSASYHARPYVVDATAFLRYDYKDVFGFKAGRYEANIDFMSGSN
QGWVYVYQPYKTETQRLRFWWSSFGRLAFNSWIYEFFATVPYLLKGGNPNNSNDFIN
GWHGITTYSYKGLDAQFFYYFAPKTYNAPGFKLVYDTNRNFQNVGFRSQSMVMTTFPLY
YRGWYNPETNTYSLEDSTPHGSLLRNGVTLNIRQVFWWDNFWNSIGFYNTFGNSDAFLG
SHTMPRGNNTSYIGSEISITTRHAGMIGYDFWDNTAYDGLADAITNANTFTFYTSVGGIH
KRFAWHVFGRVSHANKNALGQVGRANEYSVQFNASYAFTESVLLNFRITYYGARINKGYQ
AGYFGAPKFNNPDGDFSANYQDRSYMNTNLTKF

>PROKKA_01349 Hemin receptor precursor

MAQGGIYDATKQTCSTIFKSLGGSVVANKNLFIINSGFNANVIHTIDHKNDNLLEYGLN
YQNLTTFDKAI PNSELVKPGDAPDACLRVTSPNDPNMNGRCQRNGASANVVGVAQANYT
LHPMVTLGAGTRYDVYTLVDKDWQLHITQGFSPSAALNVSPLENLNFRLSYAYVTRGMP
GGLVWMRQDNLRYNRLKPEIGQNAEFNTEYSSEYFDFRAAGFVQLISNYINQFSSTLFV
TNLPAQDIYVPGYEVSGTAKYKGFSLGLSVARSWPSLKGRLIADVYELAATTGNVFILT
ASYTIPRTGLSITWLSRFVTDLSYCSYSPYRNGPTDIDRRPSNCPKTPGIFHVHKPGYGV
SSFFITYKPTYKKLKLGLSLNAVFNNVFNQQYIDQASVPMSPDEPNQDKYARGMAEPGFNA
RFEISYKF

>PROKKA_01350 Heme transporter BhuA precursor

MNDKRFRKYCSFSIFLSLLGMFELEAKEEEEEKEKKTERKKEKKENAQHTLGKVTTQAAKI
FNNYNQTTISSKELEERRQANQISDMFRRNPNINVGGAIVIAQKIYVRGIEDRLARVTVDG
AAQMGASYGHQSNITIDPGLKSVVVTKGAAQASAGPMALIGAIAIKMETKSASDFIPKGD
YAMSGAATFLTNFGDRETVMGAYHNHFDALLYTHQNIFFYYRQDGNAMKNLFDPKADNK
VTASPSEQNNVMAKINGYLSERDILTLSYNMTRDNVTAL

>PROKKA_01351 hypothetical protein

MALVYLVQSDTTIGLLSKDSEKLNALKGRPKNQSVLIESADFSTLKSIVRAPNAFKNLIR
RSAKTTFIYPNSKAVRVIKGRHGDFLKRFKTLYSTSANLTQCAVDKEIASNLADVIVSDE
RGLFESASSKIFRIYKDKKVRIR

>PROKKA_01352 Carbamoyl-phosphate synthase large chain

MPKRTDIŠNILLIGSGPIVIGQACEFDYSGTQSCCTLKSLGYRVILINSNPATVMTDPEF
SHQTYIQPITGENIAAIIIEKEKIDAILPTMGGQTALNAVQMOMHQKGMLEGVELLGAKIEA
IKKGEDRQAFKEAMLKIGMDLPKGRYAYNELEALEAINEIGFPAIIRASFTLAGGGSGVA
YNIEEFQELAKNALDASPINEILIEESLLGWKEYEMEVIDRSDKNCIIVCCIENIDPMGV
HTGDSITIIAPSLTLTDKEYQRMRDASFAILREIGVDTGGSNVQFAIHPETLRMVVIEMNP
RVSRSALASKATGFPIAKVATMLAVGFSLEIKNDITNTPASFEPSLDYIVVKIPRFAF
EKFAGVSSTLGTSMKSIGEVMAIGGNFLEALQKALCSLENNWLGFESLSKDLEAIKKEIR
RPNPKRLLYIADAFRLGVSVDEVFELCQIDRWFLSQIQKLVKAEESINSSVLTDAKKLRG
LKNLGFSDARIAAKIKENENLEVSPFEVELARSNLQIAPNFEEVDTCAAEFLSLTPYLYS
TYAPNPLPPIENKQEKQEKKILIIIGSGPNRIGQGIEFDYCCVHASFALKDLNIKSVMLNC
NPETVSTDYDTSDTLYFEPHFECVKSIIQRERVDGIIIVHFGGQTPLKLAKDLAKMQAPI
IGTPFKVIDIAEDREKFSFLKELDIKQPKNGMAKSIDEAYSIANVIGFPIIVRPSYVLG
GQHMQILENIEELHHYLESVTHSLEISPKNPLIDKFLEKAVELDVDVICDKKEYIAGI
LQHIEEAGIHSGDSACFIPSTLSPEILDEIERVSAKIALHLGVVGLLNIQFAVYENSLYL
IEVNPRASRTVPFLSKALGVPLAKVATRMVLEDLKEALKFYDKKNIVEYSKGVYKPKMP
HFVALKEAVFPFNKLYGSDLILGPEMKSTGEVMGIARSLGLAFFKAQTACFNPIKNKGLI
FVSIKDKDKEEACVLMKRLVELGFELCATEGTHKALEKAGVKSLSKVLKISEGRPNIMDLN
MNGEISMAINTSDHKSQDDAKLIRASVLKNHVSFYTTLSAIEVLLLALAESSQQDELLAL
QDYLK

>PROKKA_01353 Modulator of FtsH protease YccA

MALYDRANSHNAYAENSLHSESLVGFVKTTYKFFAGSLLLATIGALLGLMNFQAVVQYK
WVFFIAEIAALFGLMFSKSKPGLNLFMLFAFTSLSGVTLVPLLGMVIAKAGLGAVWQALG
MTTIVFGLMSVYALKTKNDLANMGKMLFIALIVVVCSLINLFLGSPMFQVVIAGASAIL
FSLYIAYDTQNIIVKGYDSPIDAASVSLYLDLFLNVFISILQIIGIFSDDRK

>PROKKA_01354 Glyceraldehyde-3-phosphate dehydrogenase 1

MKIFINGFGRIGRCVLRILERSDTPNQLEVVGINDPANWEILAYLLEHDSVHGLLPKEA
HYSNYKLIIGSLEIPVFNISIKDLKGVVDVIECSGKFLKPKTLENYLLLGAKKVLLSAPFM
GEYDEKQYPTLVYGINHFCYQNAIVSNASCTTNAIAPICAILDKAFGIKEGMLTTIHSY
TSDQKLIDLAHPLDKRRSRAAASNIIPTTTAAALALHKVLPNLKNKMHGHSVRVPSLDVS
MIDLSLFLEKKAPKDPINDLLIEASKGVLEIDLERVSSDFISNPSSVIIAPDLTF
TLENMVKIMGWYDNEWGYSNRLVDMAQVLYHY

>PROKKA_01355 Vacuolating cytotoxin autotransporter precursor

MAFKKARLISKLSKSGSKLSKISKIIFTLNQILKREKPLKRHKALKPIKKLSNRNKSF
LKASVLLIGALGGLSHLRANECRYWSWSSWSYQDNIESGNSPHTNSYCLFSSTQSGSTY
YLNLTLTYSAGGASFTQKFNNGTNLNVGGNIRFGGTGINGGDVGYITGTYDAQTINFNSSH
LTTGNSYADGGGATLNFNATNHTINQASFDNSDAGTQKSYMNFKGSNIKVSFSGSSFKDDT
DGGFSFSGNSNNSTISFNQTSFNQGTYHFSNSATLSFDNSSFNQGTYHFSNAQSVFENS
FNQGTYNFNDNTSFNNDTFNQGTYSFNSTKVSFSGANTLNSSSPFASLKGVSFSGGAVF
NLNQTLSNSQTYDILTNGTIQYGVYQSYLWDLINYKGDKAISHVGVGNNTYDVTFDING
QDETLEQTFNNQSIITQFLGDDLQAKAQKTYQQDLNSQSALNNAASDSKIANSDTSYTQ
SKNTTVAKDAQNLEHTNQQIAQDEQALQGDLDKQLANSPTGFNEQAFNQAKQEQQDE
QTLQNDENAFNTEQEGLKQAIANAKPTSPTPSPPTPTKHTAQNTPPNKVPPTPTPTQN
LPTTNVWDGVYNLQNQYTSQKGIYYIDPNLSGQSGQSANTLSTYTANLFGRSFGVNIQNG
TLIIIGNNTESVNDGLIWIGHGGFGYITGTFAANIYLTNNFKTGEGVSNSDGGGANITF
KASDNITMDGLNYNDAETVTKMIQTGTQSQHSYAAF DALNNISVTNSSFSMTWGFKFSFA
KNISFSNASFSGFTNPGGSSITISANAANSLSFVNSRLNGGAIYNLQANSIIFNNTQAVFN
VLYSRGTSNFNATTQLLGNNTFTLSSQSLNFGDNTLQNNANITLGNKSQAFAKNSLTL
DNNSNLSDLNQSVLNASGTSAFNNQASLNIYNGSQAAFNSLFFNGGTLNASSKLNASN
ASFSSNTTINLDDSVLSASNTSSLNANINFQGASQADFGGNTTIDTASFNFDSASSLSFN
NLTANGALNFNGYAPSLTKALMSVSGQFVLGNNGDINLSDINIFDNITKSVTYNINLNAQK
GITGISGANGYEKILFYGMKIQNATYSNNNIQTWSFINPLNSSQIIQESIKNGDLTIEV
LNNPNSASNTIFNIAPELYNYQASKQNPTGYSYDYSDDQAGTYYLTSNIKGLFTPKGSQT
PQAPGTYSFPNQPLNSLNIYNKGFSSENLKTLLGILSQNSATLKEMIESNQLDNITNINE
VLQLLDRIKITPAQKQALLETINHLTDNINQTFNNGNLIIGATQDNVTNSTSSIWFGNG

YSSPCALDSATCSSFRNTYLGQLLGSTSPYLGYNADFKAKSIYITGTMGSGNAFESGGG
ADVTFQSANNNVLNKANIEAQATDNIFNLLGQKGIEKIFNQGNLANVLSQVAMEKIKQAG
GLGNFIENALSPLSKELPASLQNETLGQLIGPNNLDDLLNNSGVMNAIQNIISKKLSIFG
NFVTPSIIENYLAKQSLKSMDDKGLLNFIGGYMDASELSSILSVVLKIDITNPPTSLQKD
IGVVANDLLNEFLGQDVIKKLENQGLVSNINNIISQGGLSGVYNQGLGSVLPPSLQNAL
KENDLGALLSPRGLHDFWQKGYFNFLSNSYVFVNSSFSNATGGS LN FVANKS IIFNGDN
TIDFSKYQGALIFASNDVSNINITT LNATNGLSLNAGLNNVSVQKGEICVNLANCPTTKN
SSSTNSSVTPTNESLSVRANNFTFLGVASNGAIDLSQVKNNSVIGTLNLNENAALQANN
LTITNAFNNTSNSTANINGNFTLNQQATLSTNASGLNVMGNFNSYGD LVFNLSHVS HAI
INAQGSATIMANNNNPLIQFNTSSKETGTYTLIDSVKAIYYGYNDQITGGSSLDNYLKLY
TLIDINGKHMVMTDNGLTYNGQAVNIKDGG LVVGFKDSQNQYIYTSILYNKVKIAVSN DP
INN LQAPT LKQYIAQIQGTQSVDSIDQAGGSQAINWLNKIFETKGSPLFAPYYLESHSTK
DLTTIAGDIANTLEVIANPNFKNDATN ILQINTYTQQMSRLAKLSDTSTFARSDFLERLE
ALKNKRFADAIPNAMDVILKYSQRNRVKNNVWATGVGGASFINGGTGTLYGINVGYDRFI
KGVIVGGYAAAYGYSGFHANITQSGSSNVNMGVYSAFIKRSELTMSLNETWGYNKTFINS
YDPLLSIINQSYRYDTWTTDAKINYG YDFMFKDKSVIFKPQVGLAYYYYIGLSGLRGIMDD
PIYNQFRANADPNKKS VLTINFALERSHYFNKNSYFVIADVGRDLFINSMGDKMVRFIG
NNTLSYRDGGRYNTFASIITGGEIRLFKTFYVNAGIGARFGLDYKDINITGNIGMRYAF

>PROKKA_01356 Helicobacter outer membrane protein
MPKTL LHSSFFLPLFLSFCIAEENGAYASVGFEYSISHAVEHNNPFLNQERIQIISNAQX
XIYKLHQVKNEITNMPKTFAYINNALKNN SKLTPTEIQA EKYYLQSTFENIEKIVTLSSG
VASNPKL VQALEKIQEPITNPLEFEENLKNLELQFNQSQNR TLSSSLSSQIAAISNSLNAL
DPTSYSKNISSMYGVSLSVGYKHFFTKKNQGF RYYLFYDYG YTNFGFVGNGFDGLGKMN
NHLYGLGIDYLYNFIDNSQKHSSVGFYVGFALAGSSWVGSG LGMWXXQTDFINNYLINYQ
AKMHTSFFQIPLNFGVRVNVNRHNGFEMGLKIPLAVNSFYETHGKGLNTSLFFKRLVVFN
VSYVYSF

>PROKKA_01357 putative tautomerase
MPFINIKLVPENGGPTNEQKQQLIEGVSDLMVKVLNKNKASIVV IIDEVDSNNYGLGGES
VHHLRQKN

>PROKKA_01358 Recombination protein RecR
MNTYKNSL NHFLNLVDCLEKIPNVGKKS AFK MAYHLGLENPYLALKITHALENALENLKT
CASCNALSESEICEICSD ESRQNSQLCMVLHPRDVFILEDLKDFLGRYYVLNSIEEVDFN
ALEKRLIEENIKEIIFAFPPTLANDSLMLYIEDKLQHFHLTFTKIAQGVPTGVNFENIDS
VSLSRAFN SRIKA

>PROKKA_01359 tRNA pseudouridine synthase D
MNLNFMPL LHAYNHASIDFHFNSSARDFCVHEVPLYEFSNTGEHAVIQVRKSGLSTLEML
HIFSQILGVKIAELGYAGLKDKNALTTQFISLPKKYAPLLEKNTHNLQERNLKILSLNYH
HNKIKLGH LKGNRRFFMRFKMTPLNAQKTKQVLEQIAQFGMPNYFGSQRF GKFN DNHQEG
LKILQNQTKFAHQKLN AFLISSYQSYLFNSLLSKRLEISKIISAFSLKENLEFFKQKNLN
INPNTLKALKNQAH PFKILEGDVMCHYPY GKFFDALKLEKESERFLNKEAVPTGLLDGKK
ALYAKNLSLEIEKEFQHLLSSHAKTLGSR RFFXVFAENVTSQYMK EKAQFELGFYLPKG
SYASALLKEIKHEKGENNDEF

>PROKKA_01360 hypothetical protein
MTNFEKI IIAQNRLKTN AVLATYCVIFAFIGLLVD AIRINANDLGIALFKLMTFQIFPTIT
IVMFVVA FVII LVCIQNFSSIMLSGDGYKLIDTSKVLSSKENQIHRLLLELLEEANLHFE
PKLYIINAPYMN A FASGWNEANS LIALTSALIERLDKDELKAVIAHEL SHIRHNDIRLTM
CVGILSNIMLLVANFSVYFFMGNRKNKSGANLARMILLVLQIILPFLTILQMYLSRTREY
MADSGAAFLMH DNKPMIRALQKISNDYANNDYKGIDQNSTRSAAYLFNAEMFSTHPSIKN
RIQSLSR RVI

>PROKKA_01361 GTP cyclohydrolase 1
MENFFNQFFESIGEDKNREGLKETPKRVQELWKFLYKGYKEDPKVALKSAYFQGVCDEMI
VAQNIEFYSTCEHHLLPFLGNISLG YIPKEKIVGISAI AKLIEIYSRRLQIQERLTTQIA
ETFDEIIEPRGVIVVCEAKHLCMSMQGVQKQNAIKTSVLRGLFKKDPKTRAEFMQLLKS

>PROKKA_01362 Farnesyl diphosphate synthase
MSNPNL SFYYNECERFESFLKNHHLHLESFHPYLEKAFFEMVLN GGKRF RPKLFLAVLCA
LVGKKDYSNQQT EYFKIALSIECLHTYSLIHDDLPCMDNAALRRNHPTLHAKYDETTAVL
IGDALNTYSFELL SNALLESHIIVELVKILSANGGIKMILGQALDCYFENTPLNLEQLT
FLHEHKTAKLISASLMMGLVASGIKNEELLKWLQAFGLKTGLCFQVLDDIIDVTQDEEES

GKTTHLDSAKNSFVNLLGLEKANDYAQTLKTEVLNDLNLKPAYPLLQENLNALLNTLFK
GKT

>PROKKA_01363 5'-nucleotidase SurE

MKKILLTNDGDYHAKGIKALEQALENMAEIYVVPKHEKSACSQCITITAPLRAEKIKGK
EDRHYRIDDGTPSDCVYLAINELFKHVCFDLVISGINLGSNMGEDTIYSGTVAGAIEGTI
QGVPSIAISQILSNKNKNTPLSFDLAQKIIQDLVQNIFTKGYPLKGRKLLNVNVPNCSLQ
EYKGERITPKGYRLYKKEVHKRTDPKNESYFWLGLHPLEWQKRENEEDRLSDFDAIASNHA
SITPLNLDLTSYDDLKSLESWHKGMLK

>PROKKA_01364 hypothetical protein

MGYYSEVTGDYLFNYNSTIVVAYDRSDAMTSYYINVIVYELQKLGFYNVFTQAEFPLDKA
KNVIYARIVRNISAVPFYQYNYQLIDQVNKPCYFLGGQFYCSQTPTDYYAINGFSEQILM
SANSHFILDWYDVVLQKRVLVVDGSVSGRTCGYQMLYRDLIKSTIKRIDFNRPERYYYNL
RLPLYQPCYRQ

>PROKKA_01365 6-pyruvoyl tetrahydropterin synthase

MVIRRLYQFCASHVVRNCSSSLKCAQNIHGHNYEVEVFIETNRLDNANMALDFGLMQQEMQ
VFIESFDHAHHFWDKESLEFQRFIENHCVRYVKCSFNLSAESYALMFLYYLTILQKSVF
SNNEGELKVSSVRVHETKNGYAESFLKDLENPHFKSLVHNHCVSFSQGIQNLWHDKDFFN
KIISDEKQCFHAKPLHQIP

>PROKKA_01366 7-carboxy-7-deazaguanine synthase

MKLPPVESFFSLQEGGKNIGKPSLFLRLGGCNLSCKGFNCKTPFNDEILTGCDSLYAVHP
KFKETWDYNDPKPLIERLEGLTPNYKHDFILTGGEPSLYFNNPILISVLEHFYRKKIP
LFVESNGSIFFEFSPILKELHFTLSVKLSFSLEEEKRIHLKALQNILNNAKSVHFKFVL
ESQNATQSIIEIQNLLKRLSLKNEIFLMPLGTTNNELDKNLKTLAPLAIKHGFRSLDRL
HIRLWDNQKGF

>PROKKA_01367 ribosomal-protein-alanine N-acetyltransferase

MTIKVFSPKYPTELEEFYAERIAADNPLGFIQRLDLLPSISGFVQKLREHGGEFFEMREDK
KLIGICGLNHINQTEAELCKFHINTAYQSQGLGQKLYESVERYAFIKGYTKISLHVSKSQ
IKACNLYQKLGTVRIKEEDCVVELGEETLIFPTLFMEKILS

>PROKKA_01368 Helicobacter outer membrane protein

MRKLFIPLLLFSALEANEKNGFFIEAGFETGLLEGAQTQEKRHNTTKNTYATYNYLPTDT
ILKRAANLFTNAEAIKSKLFSSLSPPVRVLYMYNGQLTIENTFLPYNLSNVKLSFTDAQGNV
IDLGVIETIPKHSKIVLPGEAFDSLKEAFDKIGPYFFFPKIEATSTSISDTNTQRVFETL
NKIKTNLIMKYSNENPSNFNTCPYNNNGNTKNDWCQNFTPQTAEFTNLMLNMI AVLDSQ
SWGDAILNAPFEFTNKGGGGECDTSKENDCVNPGTNGLVNSQNKSYVLNKQDIVNKFRNK
ADLDVVVLKDSGVVGLGSDITPSNDDGKHYGQLGVVASALDPKKLFGDNLKTINLEDLR
TILHEFSHTKGYGHNGNMTYQRVPTGQENENGKPKDSDGLPYNVCSLYGGSNQPAFPSNYP
NSIYHNCADVPAFLGVTA AVWQQLINQNALPIDYANLSAQTNYNLNASLNTQDLANSML
GTIQKTFVTSSVTNHYSSASQSFRSPILGVNAKIGYQNYFNDFIGLAYYGIIKYNYSKA
LNQKFQQLSYGGGIDLLDFITTYSNKNSPTGVQTRNFSSSFQIFGGLRGLYNSYYVLN
KVKGSGNLDVATGLNRYKHSKYSVGISIPLIQKKASVSVSSGGDYTNSFVFNEGASHFKV
FFNYGWVF

>PROKKA_01369 Proline/betaine transporter

MAHFGDKFGRKNMFMLSILLMVIPTFALALMPTFDHLVSFGVDSMGLSLKNAHYLGYYIAP
VFLVFVRICQGI AVGGELPGAWVFVHEHAPQGQKNTYIGFLSASVVS GILLGSLVYMGIIY
MVFDKPVVEDWAWRVAFLGGIFGIISVYLRRFLEETPVFQQMKQDNALVKFPLKEVFKN
SRFGISISMLITWVLSACILIFLVPNFNTLTYPNFNFTPF EKTYFQILGLVGIVSSIIL
TGFLADKIKPHKVCMAFSVAFVAFGFLFFKEFYSNAPSLVNTIVLYFLACFCAGIMNFCP
IFMSDVFSKIRFSGISFAYNIAAITAGFTPQLSSWLNAKAIAAPESLQSYGLSFYIFV
VALIAFIVSLLMAPIYNKSNTQHEVSPTA

>PROKKA_01370 hypothetical protein

MYLNLKYQPNMDNPEDNYRFEFHQKPENDKKHWWFKVEDILELKS VWDYAQEHDLKGDK
LKLETLNKA FHDKQLISFFEET EKNLNKVLNIFIRVNSGGVKLSYSDLLMSILTASFSS
DIREKMNGLVDTLKDKGFPNVGKDQVLKTCLLIGKDTTFELKNFNKKNIKEIEDNWEKI
TESIYNATKLETFGYAGYLGSAIYLSSLAYFYFLNSKMNESDKEQALKFVRNAQITSYF
TPSTDTKLNNIANSMKDAQTFESFNHNLA KHQTWPLKITNDAIEEMMCSSSRARVFPILQ
ILYPNLKYKTTTFHIDHIYPKSKFKKENKKLNQDFYGWKDYL FNLQLEGAENIAKKDKD
PEVWLKEEYKNERAIEEYKKNYIDPTLKLEWENIKEFREKREEAIKTLKEVLLPKSS

>PROKKA_01371 hypothetical protein

MGVFLDKNIKEVV DGLNVRYFLPDIQREYVWLQNADEKKIEQLFDSILRGYPISGFLFWK
LQKEDIAKSDEQDSKLNLFQLYQFITNYDERKPHNEKIRIEQIKHDELYIVVNNA
>PROKKA_01372 L-cystine transport system permease protein TcyB
MSASPNLSLFFESLDLSKERLELLLEAFYPMLKAAFCISLPLAVTSFILGLFIAVFVALI
KIAPPKHFVHKALLTG VNFYVSIIRGTPLL VQIVVVFYGLPALGVYMDPI PAGIIAFS FN
VGAYASETLRASFLSVPKDQWDSSLSLGLNYLQTFWHVIF FQALKVATPSLSNTFISLFK
ETSLASVVTIAEVFRIAQQKANASYDFLPIYLEAALIYWLFCLVLEVIQKRVEKILN
>PROKKA_01373 putative amino-acid ABC transporter-binding protein precursor
MKKVLFLLVISFFGGFLNASSLYEKLINKETISVGTEGIYPPFTYHNKEGKLTGYDVEVA
RELAKELGVKIKFHETSWDIMLTGLKSGRFD MVANQVSLTTKKRQA AFDKSLPYSYSGTI
MLVRKDENRIKDIKDIKGLKAANTLSSTYGEIAFKYDAQIVSVDSMAQALLLVAQKRADL
TLNSSLAILNLYLNTHKNNPFKIAWESKEKDGGASFVINKHQEKALELINQAMQRLINKGV
LKRLGEQFFGKDV SQP
>PROKKA_01374 Alanine racemase
MLKRASFVEVD TASLRHNFS AVKSIVPKDAHIMAVVKANAYGAGAIKASEIFLQEGANYL
GVATLDEALELRSHFPKTPILILGYSPNANASMLIDNDLSAMVFSLEQAEVFSQMALKAK
KRLKIH LKIDTGMHRLGLEPNFKSIEI IKKIRALKGLEVEGIFTHLSNADAKIKTHAKNQ
MKAFNAFLEQLLNQKIEFQYRHAYNSAGILSLCNGNENRLLNLYRPGIMLYGFYPSNEMK
ESCPTILKNVISLKAQIVQIRSVKKEFIFGYGEHFYTNEETLVGV LALGYADGLARALGN
RIQVVINNQLAPLIGKVCMDQC FVKLN DIQAKEGDEVILFGDKSAKANDASEIATLLNTI
AYETISTLSKRLE RYI
>PROKKA_01375 Amino-acid carrier protein AlsT
METIDS VVRLLSNLVWGI PMQILLVGTGLFLTFYLRGLQFSKIFYA I KILFDKESQSKGD
ISQFSALMLS LGATVGIGSIVGVATAISIAGPGAVFMMWVTGLVGMATKYSEGILAVKYR
EKGAFGYNGGPMYYIKNGLNMPKLAMAF AIFTIIASIGTGNMTQSN AVSSILSEQANLPT
WVSGLLLTLLTAFIVIGGIKSIGKFTSYLAPVMVLLYLIAI IYIIVSHFDLAIQAIKLIF
EEAFNPKPVVGGASGALVATMIKTGVARGLYSNEAGLGSSAIIAASAQTRHPVRQALVSM
LQTFIVTLIVCSATASVILMAPEYNTLLPNGEKLSANLLTLKSTEYFLGSLGTVVIFLTM
IFFAYSTIIIGWAYYGEKCTEYAFGEKKVKYYRLIFLASVMVGAMAKIDFVWNLADLSNGL
MAIPNLIALILLHKVVYSETRWYFSKHSNK
>PROKKA_01376 D-amino acid dehydrogenase small subunit
MKKEVVVIGGGIVGLSCAYSMHKLGHKVCVIEKSDGTNGTSFGNAGLISAFKKAPLSCPG
VVLDTLKLMLKNQAPLKFHFGNLNKLYQWILKFMKSANAKSTHRTMALFERYGWLSIDMY
HQMLEDGMDFWYKEDGLLMIYTL EESFEKKIKTCDDSGAYKILSVKETKEYMPVVDNIC
GSVLLTENAHVDPGEVMRSLQEY LQNAGVEFLYNEEVIDFEFKNNLIDGVITHKEKIQAE
TIILATGANPTLIKKTKNDFLMMGAKGYSITFKMPEELKPKTSSLFADIFMAMTPRRDTV
RITSKLELNTNNA LIDKEQIANMKNLAAFTQPFEMKDATEWCGFRPLTPNDIPYLG YDK
RYKNLIHATGLGWLGITFGPAIGKIIANLSQDGAN EKNADIMLFSAFFRD
>PROKKA_01377 Enamine/imine deaminase
MKEVIHSTLAPKAIGPYSQAIATNDLVFVSGQLGIDASTGEFKGADIHSQTTQSMENIKA
ILKEAGLGMDSVVKT TILLKSLDDFAVVNEIYGSYFTEPYPARATFQVAKLPKDALVEIE
AIAIK
>PROKKA_01382 Malate-2H(+)/Na(+)-lactate antiporter
MGLSASSLIVPISVILMVFTKRVALSLFVGILVSAVLMHSLHLSYIVEYAYTKITSVFY
TYEPEKGLHFNL SNLYVFGFLIFLGVL SQVILKSGSVQNFVKKAKKYSKNAKTPEFIAFF
SGIIIFVDDYFNALT VQGISKSLNDAHNSTRERLAYIIDSTSAPVCLLVPISSWGAYIMG
IMNNDSSPLLKDSFSVLVQSLSSNYAIFALIAVFLTILWQINLPSMRKYQNIGVKDFYS
EQEESSPKLAPLSLLPLSILLIVSIS SLIFYTGVILKNTDASFSLFYGGFLSLIVTYLL
AYPFLEKGSFFKLMDGFKSVGPAILVLT LAWAIGPVIRDDAQ TGLYLANISKGFLNSGG
SVYMPLIFFLISGFIAFSTGTSWGAFAIMLP IAGMANESDIILIVSAILSGAVYGDHTS
PISDTTILSAAGAGCSVQSHFITQLPYATIAMLCSAVSLGVASFMSRSRLALLIGVALLV
GVFYLLKRLYGENLKA
>PROKKA_01383 hypothetical protein
MNGDDLRLYSKPLVYSAGSGIIGIDIDKRTFYKRAFAFTMKSLFGENLLL FVKLKHSALM
SKHMKGPLENRHHHSFTKNYEKAVNGCQKYFH IKLPEGAPSNFKSGSYMATMVVRF
>PROKKA_01384 hypothetical protein
MRFYII FTFLFIVGFGVFVYSIDPQAYAFNLG SYSFNFPIAVWLMGVLMFAFFSWVFLF
KHNLSHKIRLYHEKRDFDKLLKQILSQDTQKTF LKTKFKSDLAKNLSQILARYDLKADLN

TPSSGCEKVDNLFKHYHNIENNTLEPKDHAKHSLAYEHAYFSKRLKAFIHNDLKNAFEVL
TNAQIPLELRRYAFIEIAQKGDKEVLKALNTMQENLDKECVKSFLKAFFEKSLNTDTLK
ISELCCKVGYDKNDYLKLAQKAQKFLVPDQWFQFFEILSQEDDKAQKAFLEVLLELEMND
LAKEHLAVLSFEEYMLLNAYMDLKQEHKKAYKLEAFL

>PROKKA_01385 Ribosomal RNA large subunit methyltransferase H
MRCVVYSIAKSSPLELVKIYQKQCKQFDCELELVDFPKNTANAQKVSKELAQKSYSLAF
EPYLNPNNAKNIALHPKAQRGDSFAFSKMLENHLNINFFIAGAYGFEENFLKDCQAWSLSE
MTFSHEVAKIVLCEQIYRALSIIFFKHPYHK

>PROKKA_01386 Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta
MGFADFFKNFKINKLRTVPSKEEQPSHWVKCPKCYALMYHKEVFGKYSVCLKCHYHFRMK
AAERIEFLCDVGSFEEFDKHLRPNDPLNFVDKESYKQRIKKYEKRTNRPSSVISGEAKIN
RMPLQIVVDFDSFMGGSLSVEGEKIVRAINRAVAKKEALLIVSASGGARMQESTYSLMQ
MAKTSAAALNRLSEAKLPFISLLSDPTYGGVSASF AFLGLDIIAEPGAMIGFAGPRVIKQT
IGADLPEGFQTAEFLLHGLIDMIVHRKDLKKTLSDLIAMMMHKTSKIF

>PROKKA_01387 hypothetical protein
MQGFLLQTQSIRDEDLIVRVLTKNQLKTLYRFYGKRHSVLNVGRKIDFEEENDDKFLPKL
RNILHLGYIWEREMERLFFWQRFALLFRHLEGVHSLDSIYFDTLDDGANKLSKQHPLRV
ILEMYATLLNFEGRLQSYNSCFLCDAKLEHSVALAQGFILAHPSCLKAKSLDLEKIQAFF
RTQSTIDLETEEVEELWRTLNLGF

>PROKKA_01388 Nicotinamide-nucleotide amidohydrolase PncC
MKFKFLNMDNESGFILIEKELKRLNILAQVKEDCIELKGENVERARIYLQTLFNSNIVEL
DDHQKSANALIERLKSLLDKIAVAESCSGGLLSHAFTSISGASAVFMGGVVCYSEEVKRE
LLKVNATTLKVFGVYSEECVKEMLLGVFLNFKVDLALAMSGVAGPNGGNKANPVGTIYIG
VQKLGSQALIDRCFFEGNRESIQNKSVHEALNMLARML

>PROKKA_01389 hypothetical protein
MFKRMVLIALLGVFSSVSLSAKSLLRDDGILVSDLKGMKSELSDAPAWVFEDPKVPYEEM
GVAYIPVNNKYLGIEQATLNAKLSLIVVFHEIMMKYKKRFMEQFHESEQTATNISYAVYN
YLSTKIQVSNTYTNLKSEVAVVKIKLVGCQIEQIKRYLKASVENLNDNEIAYIAKVAQKE
FGSVCALR

>PROKKA_01390 Oxygen-insensitive NADPH nitroreductase
MKFLDQEKRRQLLNERHSCMFDSDHYEFSSEELIEIAEIALRLSPSSYNTQPWHFVMVTNK
DLKKQIAAHSYFNEEMIKSASALMVVCSLRPSELLPHGHYMQNLYPESYKVRVIPSFSQM
LGVRFNHSMQRLESYILEQCYIAVGQICMGVSLMGLDSCIIGGFDPKVGVEVLEERINKP
KIACLIALGKRVAEASQKSRKSKVDAITWL

>PROKKA_01391 Prolipoprotein diacylglycerol transferase
MNAWNTIYDQFNPIAFSLGSIEVHWYGLAYACAIVIAFYMALRMIQKDPKRFPIERKEFE
SYFLWAEELGIVLGARIGYILIYEPNSSYYLTHFWQIFNPFDSHGDFVGIRGMSYHGGLVG
FLIASYLYSRKDLKLLIYLDLIAISLPLGYVFGRVGNFLNQELVGRVVPKDSHLGQIIG
IMVDNQLRYPQLIEAFLEGVIVFLMVMWAKKHTKTHGLLIVVYGLGYSLMRFIAEFYRE
PDSQMGVYFLNLSMGQILSLFMVVVSLGILLYATKNSKKIKENQ

>PROKKA_01392 Ribosomal large subunit pseudouridine synthase D
MEKAYKILSVQENISHKKAKALIDLGLVSIIGGKKLMVARKELPKNTHFSVQKVEKPGVIF
EDENILALFKPPFIESYDLLSFFKGWVLLHRLDKETSGVILLVKENSEFHLKAKKAFKNR
AVKKEYLALTQGIIEEEREINAPILTIKTKAFSKISKKGQEAVTIITPLKIINKKTLK
VGIKTGRTHQIRVHLKHINHPITIGDTLYNNEPGSAKRLMLHAHKIALLLGYEFEAIPPKEF
EF

>PROKKA_01393 3-deoxy-D-manno-octulosonic acid transferase
MFKFFYLLLLTLGHLLGAPFIFWFSFKEKYRHSLSKARFFLKDNLLKSEPVFWFHACSYGE
VKSLEPIIQALKEPILISVTTNTGFELAAQTYRHSKHIEVRYLPFETLLFAWKKNLKLKCLK
TLVVTEAELWFNVFDTAQKLGAKTMLINARISVRSYPKYQRFSSFFYALLFKRIDFVLAQS
KEDQKRLNLGAKKVVDVFSNIKRFKSPVIASFYPKNPSALNIVLASTHEGEEELGLKAFL
EFKKTQHONARLIVVPRHPRFKSVRNLLQDALKTTSFSLECFSSKGFVECDILLVDRLGE
LNNFYKIADIVILGGSFVKMGHNPLEPAFFNARLITGEYIFNQVALFELVKPYKIVQKE
DLLDALLDYKNLGVARFLENGHDLNELLAFIKY

>PROKKA_01394 Putative zinc ribbon domain protein
MNTHLKQLIEISHLDKEIDSLEPLIREKRKDLKALNDKEAKNKAILNLEEEKLALKLQV
SKNEQTLQODTNAKIASIQKKMSEIKSERELRSLNIEEDIKERSNQANREIENLQNEIKH
KSEKQEDLKKEMLELEKLALALESLVENEVKNIKETQQIIFKKKEELVEKTEPKIYSFYE

RIRRWAKNTSIVTIKKQACGGCFIRLNDKIYTEVLTSGDMITCPYCGRILYAEGAYENSA
QPPKESQEESEQELV
>PROKKA_01395 GTP cyclohydrolase 1 type 2
MALVKEVLGVLNRLSPFELQESWDNSGLNVGSENSEFSEIVACLEITLKIALNVPKNALI
ITHHPLIFKPLKTLNDEAYPGNLIKILIQKNISVISMHTNFDKTHLNKHFHALLGFDNL
MEKGLILVKENANMEFDALVKKIKSSLGVGSLACVKSSQTIKDLAFVCGSGASMFSSLKA
QSCLITGDVKYHDAMIAQSLGISLIDATHYYSERGFALIVAEILHSFNVLVTIENFKNPL
QII
>PROKKA_01396 Glycine--tRNA ligase alpha subunit
MQDFSSLLKLQEYWKNGCLVIQPYDIPAGAGTFHPATLLRSLDKKPWNVAYVAPSRRP
TDGRYGENPNRLGSYYQFQVVIKPSPSNIQELYLKSLEVLGINLNEHDIRFVEDNWESPT
LGAWGLGWVWLDGMEVTQFTYFQVGGIACSPIPEITYGLERLAMYVQKVENILEIEW
AKKNHDSVRYAQVHLESEYEFSKYHFEIASVKRLLEMFKNAQAEALHCLENKLPLPAYDF
VMLCSHFFNILDARKAISVAERQNYILQIRDLAKGCAVLYKEQEEEREERLKNALTKAEN
GVS
>PROKKA_01397 Glycerol-3-phosphate dehydrogenase [NAD(P)+]
MEIAVFGGGAWGRALAFAGKEKNEVKIISRRDLNEPLKKLNDALISKGSAPIEQVDLQRG
LKAALYVIAISVQHLREWFQNASLPKNAKVLIAASKGIEVLNRAFVSEIAKDFIDSNFLCF
LAGPSFAAEIIQGLPCALVIHSNNQALALEFANKTPSFIRAYAQQDIIGGEIAGAYKNVI
AIAGGVCDGLKLGNASAKSLLSRGLVEMQRFGAFFGGKTETFLGLSGAGDLFLTANSILS
RNYRVGLGLAQNKPLEVVLEELGEVAEGVKTNAIVEIARKYGIYTPIASELALLLKGS
VLESMNDLIRRA
>PROKKA_01398 hypothetical protein
MADDDKEESKRENLAQLNEVISLFKESIDKVFDRVSAFTFEKYKEENDDERTMIEENYKE
ERYTERVNQGGVWGSFKRKFLSWADDDAGYDEEKDYENLLNNEKLQKKLREWRRSKRQ
>PROKKA_01399 GTPase Era
MSVNFFKGIFNDNSRAENHPNNHQVGLKERYDLIARILNARIENEGLEEYQSVLDNEFLE
FASGVDSLKEKEIALTLQEIKKELQLVASYPSTLQKTMVAVGGGSAGKSTFLNNLLGL
KLKLPEHINPTTAIPTYCLKGKKEVLMGFSQNGGMVELPHLAFDHQFLNSLGFNLKEIMP
FMLLSAPSVPFEFCLFIDTPGFNSAKQGYTDGDEQASKESLKHAQYILWFISCENGEIHE
DDLEYLQELYEYGKQVFIVLSRADSRQLEEIATIKIKETLEDGIEFLGIGAYSATRY
QEIKEFSEKSPVFDSEFEFLKKNQRSEKQNKILSVLYDVCLAYEKAIEQDANQFKRYQK
ALHSVKLDLMQKGFDDFNDDAFNRIHSLKKEFSEKESKRENLAQLNEVIYLFKESIDKV
FDRASAFTFENYKEERYTERVNQGGVWGSFKRKFLSWLMMIKKNLKERI
>PROKKA_01400 GTPase Der
MNAQELIKKSALIEKTLKEQGLQEKAKPFISDNAVITTEELEKTLKEMQDTRNLKVGII
GRVKAGKSSLLNALIFEGVEVLPAATPMTASLTVLKYAQTLSAEVEFYSPKDILELKNE
HERYVREFNRIVEEVKKQKEKQSLSNRTKEVVKNFGKMFNRKSDKEAPKERILSDEEIL
EKARRIAKSELEKDTKLVSSSHQYKMKKSGLTNTETKDSRIQANSLEELNQKLLQFVGA
NGKFMPCCTKAVEISLNNPNLKDLEVIDTPGVNDPIVSREERTKALLKDCDVVFIISPSNQ
FLTDSMDLDFDRVSNKEGLEIYFVASQTDASVAPSEVQNSRHHLPTALENAQKSLSS
LNATMSSLEKKYPHQKIFEKAIKNGVILTSGACFSMYKDFKNQASWESKKEEYYNAWRN
LTNAYPDAFNSADSKESLLLLSNMGAIKERLEKAAQRKEEIIISQLQDYAQSQANNLHN
LIVQLLDLEDEKKRVKNADMGAIKQIEAYEKLSGNIETGFREAYEEFILHFIKNTRDG
LDETTLTKAIQKASALAQNEEGEEYYTERVKQGGFLGFSFKRNFLFWADDDAGYDEVSRTRV
VVKAGAVVDYLIEMHKICEKALNDSVGSFKIVFRKELYAKVFPVLRQIIRDDDLIDENAF
KKSMAVTDKIKFEEFHYTLTPSEIGAKTGFLKGDGAVQFIQSVEMHLRDFETETNDVKE
YCTDLRKNLEKQDFANGVLSKLKDKDMQNLNQVQNKESIAQLDAQIKALKGFNER
>PROKKA_01401 Dynamin family protein
MILRAQTNFVEFLEQVLEVLKEVEIDKTECSTLLASIQKQQLVIPVVGNFSAKSTLLNR
FLGSSVLPTGITPETSALTELHYSANERIEAFSNNDKKAESFELNEQSFEAIKENAPKYS
YLVYLNNEALKNSAPLVFVDMPGFDSPISSHATHAILEYLERGVHVFILASVEEGLNLTNR
MVRELKNLLEFDKGLSFILSKTDLRTPSQVGEISRYIQDQIQDHLDLTTHLIRSDKDNN
LLEVADKIDAEKLFSALEYLRLKFLNSQLQNSLKSVIKSFYDYSKEKALEEIQALDLGVKD
IEKTYEKLRLANLEEEYSSVAVGSVKKVLEEVRQKSYLASLTNKPNEFNSEIESIMQQS
LIKNAKLEIEKINLSFSKDFCAEFASLNNTQLSSDFSVNLEHSLELGINALSUILAKNPV
TRPFALILQGLKPLLDLLTLLPNIIASFNRNEEKERTKLENLIEVKVMPEIQHKLKKVL
PDLFNEALQNSLSKSLKDRCELEITHKKQEIALAQTEKEKHLNDLEAQKQALENKINALSD

LEQQYLKG

>PROKKA_01402 Endoribonuclease VapD

MYAVTFDLDTNCLNENAVNLSKVYS DIAKFMEQHGFKWQQGSVYFGDETINAVTCVATVQ
ILAKQIPSF AICVKDVRMLKIEENNDLMPA IKIVL

>PROKKA_01403 hypothetical protein

MLALEIYIDICLKDALIDYLF EKGFDDFF YVECYKYAASSLLLSQKEQVSGRKDYAKFKL
FLSDEIALSLAKDLKNQFASKEMKLFYSQTHGLSKNML

>PROKKA_01404 Cobalt-zinc-cadmium resistance protein CzcA

MLASII EFSLRQRIIVIVGAILVLFFGTYSFINTPIDAFPDISPTQVKIILKLPGSSPEE
MENNIVRPLELELLGLKGQKSLRSVSKYSISDITIDFDDSVDIYLARNIVNERLSSVMKD
LPVGVEGGM APIVTPLS DIFMFTIDGNITEIEKRQLLDFVIRPQLRMISGVADVNSIGGF
SRAFVIVPDFNDMARLGVSISDLEAAVRVNLNRNSGAGRVD RDGETFLVKIQTASLSLEDI
GKITVSTNLGHLHIKDFAKVISQSRTRLGFVTKDGVGETTEGLVLSLKDANTKEIITQVY
QKLEELKPFLPSGVSINVFYDRSEFTQKAIATVSKTLIEAVVLIITLFLFLGNLRASVA
VGVLPLSLSVAFIFIKLSDLTLNLMSLGGLVIAIGMLIDS AVVVVENAFEKLSANTKTT
KLHAIYRSCKEIAVS SVSGVVIIVFFVPILTLQGLEGKMFRPLAQSI VYALLGTLVLISI
TIIPVVS SLVLKATPHSETFLTRFLNRIYAPLLEFFVHNPKKVILGAFVFLVASLSLFPF
VGKNFMPALDEGDVVL SVETTPSVSLDQSRDMLNIESAIKKHVKEVKSIVARTGSDELG
LDLGGNLQTDTFISFIPKKEWSVKTKDELLEKIMDSLKDFKGINFSFTQPIEMRISEMLT
GVRGDLAVKIFGDDISELNELSFQIAQVLKGIGKSSEVLTTLNEGVNYLYVTPNKESMAD
VGISSDEF SKFLKSALEGLVVDVIPTGISRTPVMIRQESDFASSITKIKSLALTSKYGVL
VPITSIAKIEEVDGPVSVVRENSRRMSVVR SNVGRDLNSFVEEAKKVIAQNIKLPPSY
ITYGGQFENQQRANKRLSTVIPLSILAIFFILFFTFKSIPLALLILLNIPFAVTGGLIAL
FAVG EYISVPASVGFIALFGIAVLNGVVMIGYFKELL LQGKSVEECVLLGAKRRLRPVLM
TACIAGLGLIPLLF SHSVGSEVQKPLAIVVLGGLVTSSALTLLLLPPMFMLIAKKIKIN

>PROKKA_01405 HlyD family secretion protein

MKRALWWLMLMGVFSMGISLEAKEYPEIVLEEKNLQPMGLKVVKLDKEIFSKGLPFNAYI
DFDSKSSV VQSLSF DASVVAVYKREGEQVKAGDAICEVSSIDL SNLYFELQNNQNKLKIA
KDITKKDLELYRAGVIPKREYQTSFLASQEMGLKVEQLES AFKSFGVDPKNPKGQYGFRI
VASDSGLLALAPKNVGEKILAF TSYVRISKSDDLIAQIKLPVGVSKSIKRDSPVYNEEGE
KIGKIQSVSVVLDKGSNTILATALLDEGN YHVGEMVEMYIQGSQPKGSVLIPSNALIRNG
KDYLVFVRTPKGFKPVVQVLEERSKIFIVNAQN LHPNDSVAVGSLIGLKGMINNLGEE

>PROKKA_01406 Outer membrane efflux protein

MRFNSV KRKVQRFC LGVTFLMSMFGAGMLS AKIFTLQEFFKEVEINSMELIGKKADFKSR
LNEQRSVN AWDFFPYIYNETSMVKNFQ GIVEAQPR TLLMVRPKLPWVSSLLSKSLSIKTIQ
YDKSYQLNKNLAFI GAKRLYLTYMMTKEKYQVYVQREANFYSQLKIAKEKVKAGSMSEKD
YINFNNSY LESKLAKTNVETKLIDLEKMLDTMLAIVEPVKEGAHFDTYLNHLHDVKVIGL
DFEYVRLEPEALKFKLDRSLYVDILDLTAKDYQVNAKLANRDVFNAFEFGIGSESYSNST
NLSIEVHIPLPVT PKNIYQKRKFLDLQSGTLAQNEVMKRNIRINANSYLNQLKTKEAYIE
TQKEAIANKKRLMEMGRIAYEAQKIGLFEYLIYQNSYMDALITLAEAKIEYIDISALLEE
TLGESLTRLGVLH

>PROKKA_01407 Glycine--tRNA ligase beta subunit

MHSD ELLVEILVEELPAQALLNEYKEMPKKLHALFNKRALEVGNIEIFYTPRRLCLFIKD
FPLLTQETKEEFFGPPVKIACNHQDKTOGLNALGLGFYQKLGLKDPKHFQTAFKNNKEVL
YHAKIHAKQPTKDLIMP IVEFLFLEGLNFGKSMRWGNVEKS FIRPIHNICVLFNGENFNGI
EVKEYGFKTKQATKVHRQEGFDFIQVDS PKAYFEVLEKNHVILDPKKREAKILQEIKELE
TKHHIIVEIDRDL LDEVVAITEYPSALLGEFDKAF LKLPSEIITTS MKENQRYFAVFSQK
SQEESPTLHNGFIVVSNAINKDKQKIILGNQKVLKARLSDAVFFYENDLKKPLDNAPLER
VVFVQGLGLTKDKMERESTIAQCLTQKYISSLNMPLEKALELVKRAVKIAKADLLSEVVY
EFSELQ GIMGYYYALKQENELVALSVKEQYLPASENAPLPSSVFS AIVALS LKLD SLFS
LFSVGKIPSGSKDPFALRRLS FGLLKIIAHYGLEFDL KADLKS LFEKVG VYQSFDLEVLE
KFLLERFHN LIDCNPSIIRSVLNTNERDIVKIIQKV KALKRFLDDPKNARKKELLFS AFK
RLANINKDRNPNESSGFSISL FKEPEEHALFEAFNAIKTS AFEGLD SKIEAYFSLHAPLE
EYFKSVLVMDKDIEIQKNRKNFLWG VYQSFL EIGDIKEIAI

>PROKKA_01408 hypothetical protein

MKHLTPLTHTIFKALLGTALSASLNLVAAESPTKTEPKPAKG VKNKPKSPVTKVMMTNC
DNIKDFNAKQREVLKAA YQFGSKENLG YEMAGIAWKES CAGVYKINFSDPSAGVYHSYIP
SVLKS YGHNDSPFLRNMGELLIKDDAFASEVALKELLYWKTRYHDNLKDMIKSYNKGSR

WEKNEKSNADAEKYEEIQDRIRRLKESKIFDSQSSNDQELQKSTNSNLDLDPIGNAMPQ
TLAAQKSQIEKSQIEETQAEKPQEMKEATSEQITNKPEKEKDKPMYLAQINSTDFTPAAK
SPKKPAKVSQKRSFKNNIKNNVKNNAKTASKKQEVCKNCSPGQRNAILANRITLMQEL
>PROKKA_01409 2,3-bisphosphoglycerate-independent phosphoglycerate mutase
MAQKTLTLLITDGIGYRKSDSHNAFFHAKKPTYDLMFKTLPYSLIDTHGLSVGLPEGQMGN
SEVGHMCIGAGRVLYQDLVKISLSLQNDLKNPAFLNTIHKSKVVHLMGLMSDGGVHSH
IEHFIALALECEKSHKKVFLHLITDGRDVAPKSTLTYLETMQNICNENIQIATISGRFYA
MDRDNRFERIELAYHSLMGLNHTPLSPSEYIQSQYDKNITDEFIMPACFKDYCGMQDDES
FIFINFRNDRAREIVSALGQKEFSGFKRQAFKKLHIATMTPYDNTFPYPVLPKESVQNT
LAEVVSQHNLTSQSHIAETEKYAHVTFFINGGVETPFKNENRVLIQSPKVTTYDLKPEMSA
KEVTLAVLEQMKLGTDLIIVNFANGDMVGHTGNFEASIKAVEAVDACLGEILSLAKELDY
AMLLTSDHGNCERMKDENQNPLTNHTAGSVYCFVLGDGVRSIKNGALNNIASSVLKLMGL
KAPATMDEPLF
>PROKKA_01410 Glutamyl-tRNA(Gln) amidotransferase subunit C
MQIDDALLQRLEKLSMLEIKDEHKESVKGHLAEVLGFVENIFALETSAKTDTKLCTPLR
EDEPKSQPNIAKEILSQNKHSQDHYFVVPKIIIE
>PROKKA_01411 Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
MNFQENLAAALDLEYLWHPSCSQMQEHQNFPIIPIKKAQGIYLYDFENDNAYMDLISSWWVNL
FGHNNAYISQQLKNQIDNLEHVLLASFHKKPIITLSQRLCQLTGMNKCIFYADNGSSCIEI
ALKMSYHAHFLKNQTRRKRLFLSLSNSYHGETLGALSVGDVKLYKDTYTPLLLKNLTTPV
PKNDNEIENSLNALKRLLDKHGGEICAFIAEPLLQCAGNMHIYSAYLKQVVLCKQKNI
HIIIFDEIATGFGRTGSMFAYEQCEIEPDLCLSKGISGGYLPLSTLLTRNEIYNQFYAPY
EENKAFLHSHSYTGNALACACANATLDIFEKENVIEKNKALSGFIFNTLQNALKPLIEQQ
VVSNLRLHGMVFAFEVFIQTKERLSLAVFKKALTKGLLLRLPLNNTIYLMPPYIITHEEIK
KAVAGLVETLDELKRG
>PROKKA_01412 Peptidyl-prolyl cis-trans isomerase D
MIEWMQNHRKYLVVTIWISTIAFIAAGMIGWGQYSFSLDSDSAKVGQIKISQEELAQEY
RRLKDAYAESIPDFKELTKDQIKAMHLEKSALDSLINQALLRNALDLGLGATKQEVAKE
IRKTSVVFQKDGVFDEELYKNILKQSHYRPHFEESVERLLILQKISALFPKTTTTPLEQSS
LSLWAKLQDKLDILILNPSDVKISLNEEEMKKYYESHKKDFKKPTSFKTRSLYFDASLEK
PDLKELEEYHKNKVSYLDKEGKLQDFKSVQEQVKHDLMSQKANEKALRSYIALKKANAQ
NYTTQDFEENNSPYTAEITQKLTALKPLEILKPEPFKDGFIIVVQLISQIKDELQNFNEAK
SALKTRLTQEKTLMALQALAKEKLKDFKGSVGYVSPNFGGTISELNQEESAKFINTLNFN
RQEKKGFITIGNKVLYQITEQNFNHSFSAEESQYMQRLVNNTKTDFFDKALIEELKKRY
KIVKYIQ
>PROKKA_01413 Cell division protein FtsA
MEHKEIVIGVDIGSRKICAIVAEFKEGILRIIGTAHQDSKEINSKAIKRGRINSLAHASN
AIKEVINSAKKMAGLNADEDNRNPISSFRESYYPKTKAIVSFSGAYTESIRDITGVASTK
DNVVTIDEINRAINNACAKAGLDNDKHILHALPYRFTLDKQEVNDPLGMSGTRLEVFIHI
VYTEKNNIENLEKIMIQSGVEIENIVINSYAASIATLSNDERELGVACVDMGGETCNLTI
YSGNSIRYNKYLPGVSHHLTTDLSHMLNIPFPYAEVVKIKYGDLSFKYDTETPSQNVQIP
TTGSDGHESHIVPLSEIQTIMRERALETFKIIHRSIQDSGLEEHLGGGVVLTGGMALMKG
IKELARTHFTNYPVRLAAPVEKYNIMGFEDLKDPRFSVVVGLILYKAGGHTNYERDSKG
VIRYHESDDYTRTAHQSSPTPHIHSSPTERNLSDLKTPNTPLNTAKNDDFLPIKPTEQKG
FFKSFLDKISKIF
>PROKKA_01414 Cell division protein FtsZ
MVHQSEMENYNIGQASIEEVSDPAYKGAKIVVIGVGGGGSNMIKHLVEYGVHQDVTPIAT
NTDGQHLKNNPAPVKILLGKESTGGLGAGGVPDIGRKAEEESANEIREAIKDAKLIVIST
GLGGGTGTGATPTIVKIAKEVGALTIAIVTKPFKYEGSQKSKKAEGLKELEQSSDSILV
IPNDKILLTMKKNASTKECYKEVDDVLVRAVSGISTIITKPGDINVDFSDLKSALGFKGF
ALMGIGEATGEESAKLAVENAIQSPLDDASIDGAKSIIVFFEHPDYPMYAYSQACISI
QERANQDQDVVKFGQHTSESIPIDHVRVTIIATGAERNSGRAGLESIAATPSQPVVKQTRKV
GNSDFLRIPTEEEELSTPTAIRIQD
>PROKKA_01415 hypothetical protein
MRSVAKHSYSETKKDKIKRDDKSKSASFGFSTYESIARTYKVTEQKDIIANSVLAKLKAN
LIENGQRKIRGTRGGWENWSKWVPGYSYDLNASFTMDSVLDNTYKPKPLKTNQFDTNAGY
CSINEQDLLNQFINVSAKLSHDTLLAIRQTENIDLVIYALIKDTTYSNLNKRYDFIFEIDK
AVRISIS

>PROKKA_01416 hypothetical protein
MNTTALSMSRLYDRMLNCPVAVISGCRTYNDKAIADYGLKEATKILELNATIEGEEKKEE
VEALLKPFMLTNADNEKRTKWLRRKKLDLYKGYIGYKIVKGYREAGMPQANIETSFICFE
LKPFHNKNQFKDIIVNLGREFNQDSIIINSPKDKRGAVFTLICCTTHYPYMLDGGKKQDLE
NPLLN

>PROKKA_01417 hypothetical protein
MSSDFKNLFIGANNSKRKTSNLLDSEKPPINSHSLSSLIDARL

>PROKKA_01418 Small-conductance mechanosensitive channel
MDEIKTLLVDFFPQAKHFGIILIKAVIVFCIGFYFSFFLQKKTMKFLSKKDEILANFVAQ
VTFILILIITTIIALSTLGVQTTSIITVLGTVGIAVALALKDYLSIAGGIILIIILHPFK
KGDIIIEISGLEGKVEALDFFNTSLRLHDGRLAVLPNRSVANSNIINSNNTACRRIEWVCG
VGYGSDIELVHKTIKDVIDTMEKIDKNMPTFIGITDFGSSSLNFTIRVWAKIEDGIFNVR
SELIERIKNALDANHIEIPFNKLDIAIKNQDSSK

>PROKKA_01419 hypothetical protein
MIVSFFAGPCYYVLLAMEYQIAYEHPGELISTIGFVALAALVYYLWGWKWERLLWGAPSNQ
EQQLSNQGNQKQ

>PROKKA_01420 hypothetical protein
MNVLIRLCFILVIEFFGAKKTLNATAIPSLNFGSFSMPITANLSDGALNIFKWFEKHPSV
GVKVGRLANQDDTIFTLVFIVIVVAIIALIAIFIRSILLNTIFVGSIGSLWLYMVGFFY
FYGVPFLSYLSGCYESFSFSACYPDSLQLLPTLMQYSPIYSIIKLLAHFNIEITIRLSFL
LFGCV

>PROKKA_01421 Tyrosine recombinase XerD
MSDCKMNRVSRKLFDNISFLHYKLKTMIRIQSVNDLQLVLKWQDRVLECCQSLIALKELN
RKLYNQGVRRHTIMMQLFLFFFEYFDNRIKLKSLHNLAEQVIDFLFGLTKNRKPSSMAKY
VMYLRQFFDYLDKRKNYSFEFELKNLSFAKKETHLPKHLNKNDFKAFVQALLKYHPKTSF
EKRNCQICLLILVLGGLRKFEALDLELKNIALENNHYRLLIKGNKNKERYAYIEKEFLQVP
LNAWLSDTKRKLSFKGRFVFKKAKNNTTQKTCSLKGFIAKIFKLSNIDVKSGLGLHLFR
HSFATFIYDATQDLVLTSRALGHSSLLSTKIYIHTTQEHNNKVALVLGGLLRNEKSE

>PROKKA_01422 hypothetical protein
MKLISFITPKISNLEIYLRNALDYCLTQMKGNEWVFDEVSLIPLIEELKDKKKEITHSLV
LSKMSLGAVIKLIFCYKLEGVILDLRAYRFRAYYHENKDTLLIKNRKQNLSNYAKAHIAL
NLLWTIRNRAYHWENLLKIQPNRPRITTYFTGLKDNDRAKMPMNISVEPSKIVLFLDDL
IKSIGNKDLENLSGL

>PROKKA_01423 hypothetical protein
MAKTLKPNLDKDELNTLYKANLAYAKNTHEHYFKFKKDLDYKLFNPSIMHEQGSISFVGG
QGAKRLLDILYKLAFAKSNKIALDRHYAKMFLQVVARTLRKNINILEEQGFIEVIKKGQ
RYLYVYLKDYRELEGYNYVGANQKNIPSPFFLQIMRFLEKFAKEIEKVKINNKECVMHI
PSQELMQRVNNVVLKFTPISNPNTTYTSLSYKDLTNKDIPSNLCFYQTAIVKKNLLKALD
KECVGEPINNSPKVSNYFQESA

>PROKKA_01424 hypothetical protein
MTPSTTLNICFNKKNLKLILQIDFSQMNTTQEKFLADLFEKALQKNLQVYWITTTETKN
ELTREEFSNLLKENND

>PROKKA_01425 hypothetical protein
MIELILHKKSIQIDETLLNVKEHLEKFYSNKEQETIAKTLESQTELTC SYLWDKDFSLL
KHLNSLKNQIIHLTQVVNEENLEFDKELVIYHLDKFLNQNNTYKVLAKFCVLKKKGTLHE
KIKAF

>PROKKA_01426 TrbC/VIRB2 family protein
MKKLRHFRKLI AFLGFSPLLLQADMTTFFNSIEQQLTSPTAKGILMVIFLGLAIFIWKNL
DRWKEILMTVLAIAIGAAIFFKAPALANWFMSIF

>PROKKA_01427 hypothetical protein
MQLVGISVSNLKEISSKEKFLWLNKSFLLSGFVPFIMIPWLDILNSFVLYVCFLIFSI
AEFFDEDISDILIAHSKI KTKANSFYA

>PROKKA_01428 hypothetical protein
MQKEVLVEKPNPLTISYNLEQAISSLKSSVNALMKKEAHLDAYILVTNIRNIVDELHKEV
VLANQSNKNTPKRKRKSS

>PROKKA_01429 Type IV secretion system protein virB4
MLESALKYCKEKAIDLLAGFVPKTYSMAQECNIGLYDDTFIITKQENLVGIISLQGLSY
SNLMQKDLEGYFDTRQNVLNTISKDIQLRIVAKRRKEFINQSPNIDNIYAKAIITQFESK

EIYKTEYFLVFETITSNVKSFFEKKKLEMTTSINEELEESSKEDKQENENSSNETHSNTS
SKDKKNKFKKKITFSATSKRALLIQTIERVKNALKEFKPTLLNSKEVLNFYAEYINGKY
IAFNPKLKRLSDSYIASNVHFKKDYFVIEFQNTFCACVGKAYESEESSIPISTLLH
TKIELDLIFHIRSLGQFESLNLKTKKKLTLSKIVKADIDNYIELVQANRLSMQECALNL
VIRAKSKAKLDSKLEILSLLNAGLGSAETIGLKPSYFSFFPNNANINPRMRHQTSQV
IASLILFEKNNTGFRANSWGDMPLSVFKNLDHSPYLFNFHNQEVKHKGVLAHNVARVVGH
TMIIGATGAGKTTLISYLMMSALKYSNIDILALDRLNGLYSFTKYFDGIYNQGENFHINP
FSLEDSATNRAFLHLHFYAQMAKVDSYDDHKDKVEDRTALLNAIDTMYRNYNDEVKQAKFS
NQELPLPFDLKEFVNIAIVKTNADILDSSFEDYKSSLFSSRMSLDLDFKTRISTINTDSIL
HNDDDAGLLAYYVFHKMIDRALRINRGFLCFIDEFKSYAQNEIMNKKINEIITQARKANG
VIVLALQDINQLSEVRNAQSFIKNMGLILYPQRNIDTKDLNDKFGIRLSDEKHFLENT
AVNEYKVLNLMNDGSSNIIDVSLSSLGNYLQIFSSNSSMVEHIDNLIKHYPKTWREVFV
SNKHENFDDKKHLEKVLK

>PROKKA_01430 VirB8 protein

MSNLQELREHLKELENSFEIGSFTKENIKEYAKCFFMSLSMFLEEQEKNNQEEFLEQDTK
ENQEELIKSIQTSIAKNOELEKISFEKWENKIQERVLPKLKRIVTHKLQESITSSINTQL
ESFKKDELDSLSSVFEIQRKNTQIAYRLAIGGLMGIIALSIAIFIMMPLKENTPYFIDFAN
SDKHFVAVVQRADTKVDYDEAFRLNLVGSYITARETINHIDDKIRLNETIREQSSEEVWKT
LEQLVSGKGSISYNSNMDREIKIINISYKQKGQONIAVADIVAKVFDKGYLISEKRYRV
SLIYHFKPLIQFDYSSMPKNPTGFIVDKYSLSEIASIKALDKTYKKVERPHSKIEYKKQN
PTNPQNNPANMPNNKEEQ

>PROKKA_01431 Conjugal transfer protein

MIRAFRGISFTLSLLAPLLAENNIQNYNEPMDYSKGDVSLQPFNNQTNQETQPSTNYNTQ
PNSMQTPMPNQNPYSQPTTPIQNNQPNNTQTQTMLNNAKQFTNNALNNTKQITITNATNYV
NDKLTTPKPNDAPISAFSANSPYANKENSKQTQKSDEVLSDDNSDEDNSLKPMTNFNAIQA
NFFAKNRSIKDNTHFIHYTIGDTFNIRLRYAMTTTFIFDEPIAQVVLGDDIGFSTKMLGE
DKEHLISNILLIKPLQIGVDSNLTIIGKSGKVYSFYMFSTTYTSSKNPALMVYISSKEYF
KNFTTTTKEINKENKPTHSTQNKDIKTSHNVPKKITKKSPNIKEQNPLTKSKNPPTTTKIA
LANAPISLNRFKSPSYPPSNALANSRKASENFKVISNQTKTKEHRHSQDEMKYVDYSAK
IKDDGKFIKIGDNVNHIIYINKKKMEYGYVIIDKKKRKYCLWMCKKVVKSKYKDDLPTQI
FNDEQFTYFKFNRSNARSKFPVVYKVIDGYDNPVNSRVVDYLIAEDVSNQWNLKLGKAY
LCIEKIAKRAR

>PROKKA_01432 Type IV secretion system protein PtlG

MKQSLKKGIGLVGGFVVLMAAGLLANKGSNNNNKEEIFQVSESKFPLSDYFFEEVKEEP
KEKAKETPEVTTQEPTTQTSNPNSVSANPNIKDYSHLNNNLNQNNNTITHLQQVLQSPSK
IKPKPKHTKEQLEFLNTRLKPLEPTXPATKPDTEYGVDSFTNLKYKDIDGTNEHKLLRTIT
ADRMIPAFILITPISSQIAGKVTAQVESDIFASMGRAVLIPKGSKVIQYNNNNNQIGQYRL
NIAWTRIITPQGINIMLTNARGADVKGYNGLVGKYISRNFQRYGIPLLTNTLSNGLLIGI
TSALANRQNRKGDGNPFFGDYLLMLTRNTGMGINQVINQMLRQYGAQNPIIIIREGSRV
FISPNDIFIPKPKDGEILAEFFKEKKPKIKTKEEINNEEEI

>PROKKA_01433 hypothetical protein

MKKKSNVFKILMSLCIALSNSHALVIQTMQDWNLRKNSEFKALVVVGKQVQENTIAMGNY
KVSKNRANLSISVTHIKGNDYKTLQSPNPAIKEIKSKLVKKGTRIVLGGENDSEMSNA
IGVSKSAYVSSKNGKGFSGSGASGMGYASGYGGTSNNAGSNGTSANGVNGTSGNNGAKGE
NGSSGANGANGTSGYQGVGSNPFPIAGSGNGSSGSSNSGYTPFMSGGGGIGGMGGGFIP
FPYSPGLQNGSGANGINGTNGINGTSGANGSNSANGGTASADSHYTNEFCTAPIRNGNTM
NIDIVDRNGSCFKMQAFRNDNVQCQYNYDFENMKAIKQTQFYFINRENKSINIGGCVLDYG
EQFEFPMYEDANKCKLDTTNNKGYGSGMASFFQTEILFRGLDNLIHVAKPCTDFANVQEQ
LVKYDNDPSTRVQRVIDQYYINPITKEKVFISHDVVSPLKFSYQTYACGKWQFDNAKLE
AYRPTQIRIFDVTSNQYYNVTGCDYTSDMGKVPKIIQPYTKINVESSVKGEDVGDLDTIS
GSNYDLQQYTTFEIQEMSLNSSQWITTSYCDGWSWTKTHSYTGYSQGNLRTLAKWVTR
YQTINQGTITIGYIRPRNEGSDDAVYDTYKYYVSQAQKTLKRFPSIIYVNTPNLSDEYRK
NFELTVLSQKLNLNNGVKQLINTQEFTNWQATYYRPNNGNGLCMSYNFSSGNAWKPKTKG
KYCSWAQTPNANYSCASHNNWVPNN

>PROKKA_01434 hypothetical protein

MNTNLKKALFLTTLGYLQANESFLAIMPDENKQEDISQEQKLEHKMEQKIAPQNTLKKQQ
SAPLQSAELLQSQVKREVKEQKKSQKHLVAMQTKNGFFLGSEIAITGKLTQSLSLISNA
TDDLITYSHTSSNANFNMGVLMGYQHYFGATKKHGIKFSMHLYGGTPSDVSTDIKIYKSK

VSTNYTPLNFGIDVKYLFDFLNVIKKNRHTLGLSVGFGWRMQYYFAKIDPIKTSHSEFIT
NKPKDIFNHGFYPTIGLHYYLNHHQFEVNYRFGGAINYVSVVPSKIVVQNGNFKGEVLFS
NFLTKATNSSYIAFNAYLF

>PROKKA_01435 hypothetical protein

MSDEITQENELEINSNNQNQEPKEVEKMPLNNIQKAKELKNNANLIVRRTDELDKAINKR
ESLQREFERRIKHLDNKIETLSNNIEELKRK

>PROKKA_01436 Type IV secretion system protein PtlH

MDLDKLDYRALRNAILRLLPYLD SGITELIMNKEKEIWLYKLNGVREKVF DENLDKAFI
LGFGEQLASFRDLFFNANYPTLNTSIPTSRYRVSMNHFAISADNELSLNIRVPSDKKFDL
KAFKLSSTCQYDYEYLKNLMIEGKNLLISGGTSGSKTSFLNALIEFIPKHTRIISVEDSE
ELDLRAFENHKSLLVDKTESKFTYENALNMAMRMSPDRLMVGEIDTRNSMLFLRFGNTG
HKGMVSTLHADSVHGVIEAIALNLQMNKSGLDVNVAKKFFKSSVDIVVQIVLDKATNTRY
IQEILPAKDLRDSL

>PROKKA_01437 hypothetical protein

MKKYQNKRRKLGKRINCFISNECKDILENYLKVKKLSVTQCIENLVLGLDPSKETLKTKEF
YESLFFKVISINQHVYRELIANSNNLNQLTKNTNIAIKYNKPFYQLQNLAFLEEIFKVM DTL
NKNVEECKKISLELLLQIYKEQKSQTEFNAISKILNALNEKQSTTEPKKDNA

>PROKKA_01438 Conjugal transfer protein TraG

MMQMVFVKAFCVLF SILLIPLFFILANYTTFDALKSLRQAYFFSQSVFVGLYHGASIFDLK
FEVYLTMLISLMPFVATIYINFPKTTTETSHGYARWANVKDIECFKIFSKEGFCKVVHRLG
VQFDNGFILGKFGFPKLRNVCYDKPLGAMIVAPPGAGKTACVALPNLLTLPNSCIITDIK
GELRDKTAGYRQKFLNNRILIFNPYGDDNTCYFNPFDKRIVEKMTFAEQLRHVKA VG DGI
FVDEEDHWVSKAKELFVFFALLQVVTKGHSSFYDVSIAPANDYAPLIHPKSPYYKQLYQH
DKKTGEVILDPQTNAPMKNPQANVLKLFLNLQVADQKYIDMNDEKNYDPREPEPPYGT KGA
LDEIIRTDARSWANTPDDEFGSIMSSFKRFRMYVYKDPKVREATSKMSFDYEELRTGNISI
YIVIAQIDIGTLSSSLVRAFLESI AKNLMVKESSKPEERIFI IADEFVRFGKLPF LLEMPA
LCRSYNV VPLFITQDYAMIRKYYSDDDLKILKGVVHYNIVFKMNSAEDAEIVSKEVGEFT
RQSKNYSTEKGQLVFGGSSSYSHEGRNLLTAQDIMNINSDEVIVIVTGAKATPLK LKANY
WFKDKELLKRANLPIDLEVERQRVEEPIQPTTPQELENSNLKESEKDNES PITLENANEN
IEQGNHNEIDEILKKPLSEISMEEKRALFKKMQQGDEESEQEVTQSTQS

>PROKKA_01439 DNA topoisomerase 1

MNSVIIIESPNKVAKIREITGAKVFATIGHFMQLKSYDESN GFKPTFDYDQEKKKHIFE
MIEACKNKKVYIATD TDREGYAIGYMFYQKIKNVASSIYRAEFFEITPSGINKGLQ NALL
FENTNKQMYQSALARRVADMLLGFTLSPYLGKALGQMKGCSVGRVQTPCLKLIVDRDREI
EKFKALPENЕКVSYQIQAKINDSANREVTIKHCDEKGEEIKFNDKEEALKLFESLKD NKA
CLLKD LKNSV VETPKPKPFITSTLLEKASSMLGLSIIIEVQSLAQNLFEAGLITYIR TDAE
SLSVEFLDETESFYAPIYKDLYLKREYKAGKQSQA EAHEAIRITHPHTTEDLESIVY NAN
ITNQDALKLYQLIFERTIESQGKNAIYDKQDLLFKIKNEYFKCSVKGLKSAGFLAMFSKK
ELENDESND DDKNKEKEQNAQFNLKIDDVLSLNDLVLATIKRNAPSAYKEAD FVKLENK
GIGRPSTYASYLPTLVKREYISISQDKKHIITPTHKGKRVVEVFENAYQFIIDLT YTKQM
EEMLDEIVENKSSYVDFISNLNSKCPKIEKLERNDDEIKPSSEGQITYIENILRDLQ LNL
SEEFKNYKEDNRVAKAFLDRYIKEHEFFKKN SKKASSSNNDENR PATPKQISFAEMLAKK
HNVKL PKGFKYSMKVCGDFINEYHKK

>PROKKA_01440 hypothetical protein

MLYIPPTSKEVYVSSIVALNIHSPQGTGDWHSSYTLMENAFDDIGVYIYGEKQAHNTNKL
LGNLGIIDGTARLNKMGYPKHTPTYIAEHPRACVDCLYVSVLQGTGKLGVMLDEWFPSI
EDKESVYALIEVMKPKLNKQERENLDKWIARNPIIE

>PROKKA_01441 hypothetical protein

MDSKKPYRVSLSEQALNHEKLMRAIVKNLADTPMVLKGGTALYLG YGLNRFSEDLDFDC
HKKINLLGRVKS AIPNGIILNDIHIKKD TDSVGRYMVRYATKDNKEEQTLKLEISYRDAS
KESENVIEGMRIAKIERIIDNKL CACFDGEHTRTKARDLFDLHFLAKHYEEHFNL DLAS
RLKDFS KDPDKLVSDYLV DVKLDALLNQIMDLEETALELGVMAQLIHKKLEKQSHSLNAL
QEQQGYSNNDNSLDNSNENTYTPKRRR

>PROKKA_01442 hypothetical protein

MKTMNLNEFFTHKVIYKDTPLKFKDTLEQEISQASLVEKLILANILANMVFAKISNENAP
KILISRLMCKFSPI DHESTIPSDFKPIDEEYEDDLEWLNEEKEDRLFNY YLFLNGIKES
DVEEVFNESVEIYDECLIEIAQNVLKDKFSYDIDLLQVLVKGYAKEIREFLSYKPTKEIK
DFKDKDTALYISLGKDYDKEKEPFSKKLQQCFKEILES KGI

>PROKKA_01443 hypothetical protein

MDLQQIDELEKKFEEQEEQAQDTPLKQEPSTKEVKIPKKRGRKKSLLDEDDKKKSFNIAFS
PCVIKELDEFLLFEGSFKETRSAFIEEALIRHLKHRKNTQEQLKLLKQLERLQNKEKGNNE
NNELE

>PROKKA_01444 CobQ/CobB/MinD/ParA nucleotide binding domain protein

MTICIANEKGGSGKSTLCLNLAVQLLKDNKEVVLDLTDSDQKSMETTFATIRAEKERPTFSL
FNRSSGFSDTLKQMVSKYENILIDTKGEYSKETQKAMLLSDIVLVPTTPSQDLDEVLANM
LERIEQLQELNENLRALIVINRMPTIPTLKERQALIEFIKENNPSDRITLLESSLSERIV
YKRSVSEGLGVIEYSDKKAINEWVNFYNELKSHLEKEKRHAF

>PROKKA_01445 hypothetical protein

MTQDFLELKEGKYPMNEWQMIEFMQMGVEPKSEQETKIELEHSLNEVKNKEH

>PROKKA_01446 hypothetical protein

MKAWNEMDIKEQKEVFARFLASEIAKSLESGLEFKPNNRAYNGNAGNAYNGLNSLILDAK
QHENGYESNVVWGLDDALKLGANPKEVEHIKNNTKSKNNPNGFYDKASIAIYIRDYEMRYV
KARDEQDNLIPLKDKEGNLKHYSNGEVIYENEKVPQRDKEGNIKYMQDGKTPFYEFKKEK
IDIEPTLEIKNLNVNVFQTLDKTKLKELDPKTLRTQYISNTFSMNNQNLVIYDLSKHLN
EKQYNKVLVDYVNYGATHHEKRKQHKIYQEPQQQIAPNTDNNAKENNTKEPNNNEMNMAE
FSKMLEMAQNNPQMLAMLQQTNLKGAEQQKYDHYFANDEVVLTQSKGRGR

>PROKKA_01447 hypothetical protein

MQDFINGFLKAWKAWSDDGKDAFLESFLDEVKNSKNDPTNFLGILEQEIKSLSQEGLKDD
ISSMLEVIESMKNEISNTKDNENVSDFFNAFQQEINEEQSISNDNPSEFVESLEKVIVET
QNANKHEIQNSSENEVELTEEKSFLEMSSEDEYYEYQQEQYIQMNEYDGIARVSACLEQQE
LEREEAELESKALQDYEENQIQRAGEISIAEHENKLEKEELERYALEGENKALEQYYESQ
NPDTLYEQQVFQDQEKELLNQQGLELETLRDNNNTKQGLDNESATLLESLOVIRDETNEVI
TFLNPNELELLRPTGENTTSKIASIALEESNKLNEKEPNTQEQQEYQAQKEQEKKTPSNDI
DLKDKYASLSHFFKQTELDIKQKRIKTKLDFKESFLNYAKNHMNEDELKVLLVVSQKEPT
KLTKEHINLIKDCLEQAAGNKHMQDKGLQLSVKSILGHQLEKEHHEVLKDYLAKSKDSLE
PMDNRVYEVMRKHLQQAQKTLKLNLDSATQSVGVNAIKEQNKANKNSEQPKNSQNEPRQE
TTNAQTNSTATQQENTKQNAIEQNGTTQAKEPKKTLHPDEPWLDYDPAHKGLQERQKE
ESQEKQSQSNSDEPWIEFGKKEQARAKAHYQAMLEKEKAKELAKEQNNAQKEVKKEMPTI
DYGTYQNTLSKSRRR

>PROKKA_01448 hypothetical protein

MSYIPNLTALEPLHEILLDNGYVINKNKHSKNNPCLKHENEESLVIIFKNQNKDGSISYTY
KETHTDKVGNIITFCKDRNISVEDLLAGKLEGYRNKKDTLQARDNSSENNEEIQKIINEF
KNLKPYPDLQATLIKKRGIDTKLLEPYKEHLKTDNFNNLILATYLAFFENKNLNVIPHIQC
GINKRLNTPSLDKEGNIRDKPLKSIAQSGSKGIEVLFPNNLSLVKNVIVTENIFDSLAYL
ELQGLEPKESVLISTAGQFNAQKLELFLKSFFKQLKGRQQGAYNHYLKQEEQWQELVRQG
RASDDFNSVVVETTYTDIIKNYQREKNALIYNKQVERTREYRKPKPVNKPQDSFNVILAFD
NDIKGKGKYEKCEGILYALTQQFPTIYTPFSKDCNDDLKLAHIIENKAINIDTMAEFLES
SLEKLKDNYSSTQEKENIMDKLEQIDSIKPFNERLKGILENAKENLQTSQCVKGRGR

>PROKKA_01449 TrbL/VirB6 plasmid conjugal transfer protein

MNFFDTLMGMFVEPSQKVAKSLAEHVGSFFHAQLILNTIITILFMILAYKRVKEGDMFEF
KTAIGVVVFVIFVGFINWGIRNPNDFTYFINTIFYPAEKLAIIAQSLNDGLEIPTNAN
PSETFNIGNLVSSAYAMIVNLWDNAFDGISMFNWLTMIPIIMFFLVILGELLFLGLLLI
IVLLVTAEIFTWSALGLIVLPLGLIPQTKGMLFSYLKKLISLTLYKPCMMLVAFFNYGII
YKVNALIPTKHEVTQGFYGNADKMANEGKIIDAFGNVLKGDWNSYIAHSSIVGFLTIIIVL
GSVICFFLVKRPVDFINNIFGTSGGVGAVTEMMQKIGMTIGGAVFGGSAVMVANQVKQAY
QSAGGGLAGLQAGAKALGLAGLSGGANAVNARSAGVRHFVASVKSFGFGFDNDKNNK

>PROKKA_01450 hypothetical protein

MKKSIFSALEFALLCCKSVFAEGETPLIVNDPETHVSQATIIGKMVDSEIKRYEEIISKAQ
AQVNQLQKVNMMINTTNSLISGAITLANPMQVLQNAQYQLQSIRYNYENLKQNIENWNT
QNLIRNKYLLQQCPWLNVNALTNNKIDNLKLNLLITKNGEQTQTARDVQNLIQSLSSSG
YKNMQSLAGGLSGRAWGEMLCMVNDSNYESEQALLATGNNSEEQKRRFLLRVKKKVNDN
RQLKKKLDPFLLKRLDVLQTEFGVTDPTANHNKQGIYYCTENKETGKCDPIKNVFRTRRLD
NELEQEIQTLTLDLTKAPNKDAQSQAYANFNQRIKLLTLKYLKEITNQMLFLNQTMAMQS
EIMADNYFRQNNDFGEEENHIDKQLTQKRINERGRARIYFQNPNVKFDQFGFPIFSIWD

>PROKKA_01451 hypothetical protein

MIKLNLNKTLLKTSLLAGVLLGATAPLMAKPLLSDEDLKRVKLHNIKEDTLTSCNAKVD

GSQYLNGGWNLSEKFPQEYREKIFECVEEEKHKQVLNLINKEDTKDKEELAKKIKEIKEK
 AKVLRQKFMAFEMKEHSKEFPNKKQLQTMLENAFDNGAESFIDDWHERFGGISRENTYKA
 LGIKEYSDEGKILAFGERSYIRQYKKDFEESTYDTRQTLSAMANMSGENDYKITWLKPKY
 QLHSSNNIKPLMSNTELLNMIELTNIKKEYVMGCNMEIDGSKYPIHKDWGFFGKAKVPET
 WRNKIWECIKNKVKSNDNTTTKIGIVWKKNTYFISHH
 >PROKKA_01452 hypothetical protein
 MQDLQDFKNDITLILSKDRLETYDNLEQYKENLKLISLITPKISNLEIYLRNALDYCLTQ
 IKGNEWVFDEVSLIPLIEELKDKKKEITHSLVLSKMSLEAVIKLIFFYKLEGVALDLRAY
 SLKAYYKDNKDTLLIKGRKQHLSTNYAKAYIALNLLWTIRNRAYHWENLLKLRANNRPRIT
 TRFIRELEKPTSKSFNFGIMPKNIVSFLDDLIKSIGNKDLEKLSSL
 >PROKKA_01453 hypothetical protein
 MSFAPMLLATINNSIGNKDKHVSLEYLIGLFMDKKTTNLSNTDKYIIGTIQTEALEQEIE
 WFSQDYHIPMENILHVLSINPYQ
 >PROKKA_01454 Serine/threonine-protein kinase CtkA
 MSFLSIPINEKSLSFIKNNFQAQIQALEQFYPIIKNALNLGLNPSSIKFGGGTALSMYYF
 QHRLSFDIDLFDVNDACQLGFFSPKLWIDNCSHFDSTRYIDQHNHIGVTNKDNIKIDVLVD
 YASNEGVDNSKKIFAFDIYIESLENIIAKKITFRKTDNKTRDIFDIAVALHNDNNLFDK
 LLSSQKITKQDILLTLQNSLQELDKKRYSEIDMVEPISHYKDLSSAYSDFLIDKIDQIKT
 ATYSNSHDNADDLDNSNEYTPTPKRRR
 >PROKKA_01455 Tyrosine recombinase XerD
 MPHNLTKSQRELCNLKAFLYTKAKNFTPIQDVKDMALILDTQDKILKCHNIEQLKQLC
 HILYNQGIKHTIMMQGLFLFFNYFKDNLKLSFRMLSEEQVINFLFELAQNRTSSMAKY
 VMYLRQFFDYLDKRRYSFDFTLKNLAFAKTKESLPRHLNDKDLKSFLKTLLGYKPATSF
 EKRNCICILLIVILGGLRKCEVLNIELKHIQVEEQNYSILIQGKGRKERKAYIKKSLEPS
 LNAWLSDEYRLKYFNGAYLFKKDKQKAQNSLTLYNFIPLIFKLAQIKHYKQYGTGLHLFR
 HSFATLIYQETQDVLVTSRALGHSSLLSTKIYIHTTQEHNNKVALVFDSLIENKK
 >PROKKA_01456 hypothetical protein
 MGKEFFNELRSEFAYRLNDHLLSEEYLYVNEPKLKELDNIKQQLQDLEKEEKALEQIKSP
 QDEWDLNKALQSEYLQELKDKNKAKALDIQNNHSTPLKQKISEFKIALFNHKDTSDDKEE
 QLDIDRIDKRKPVSEHLKNTNKHLEYELLGFYQKELDKKQNHSAFKNFALNGLDRDFER
 ETKGFSVLKKKEMLLSKLERLDKRLLDKNSHLLLAQLRNVVRTKQNTQYNTLTNPFLAK
 ALELPKDKRPTLKTFFKNAYFSARKYQFMLESFKTKQNDPTYNLNDNAYELVSKQLQDYQN
 TMLLLAKERLLFLEQDLKQKEEEFERAKEHYVKSLSKHYRETSFLFPKEKQGFLKQIKQFSK
 ISKDILYTCNEIIGANRFLTHYDNLNLEKVLHAKDTKLEQKEIQAITKEPNSDEPWIEF
 GKKEQARAKAHYQAMLEKEKAKELAKQQANTLHSNELDDDPKAHAGLQNDNTNFKGRNR
 >PROKKA_01457 Transposase IS200 like protein
 MSKAVLYKSNHNVVYSCKYHIVWC PKYRRKVLVGAVEMRLKEIIQEVAKELRVEIIEMQT
 DKDHIHILADIDPSFGVMKFIKTAKGRSSRILRQEFNHLKTKLPTLWTNSCFISTVGGAP
 LNVVKQYIENQQNSNRPKQKEKWKSVDNLQTKAL
 >PROKKA_01458 putative transposase
 MLITYKQKLYKNDKNRRIDTLLRHYGVFYNHCIALHKRYRRLFKKYLKLNLDLQKHTTKLK
 KTHRYAFLKTLGSQTLQDLTERIDKAFKKFFKKQAKLPRFKKVANYKSFTFKSQVDKKTG
 LNKGVGFEIKDNVVSFNGYSYKFIKTYALIGKVTLTIKRDNMGDYFLCLVCKLENNPNK
 QTACDKSVGFDFGLKTFLTGSNDTKIESPLFFSKYLPKIKRLSKNLSKKVKGSNNFKKAK
 KKLAQLHQKIKHLKTDFFYKLALKLSRKYQSIFIEDLNMKAMQKLWGRKVS DIAFSEFVK
 ILENKAHVVKIDRFYPSSKTCNSCLSVDENFNKDIKKLGKTDKEREYYCKYCGLELDRDL
 NAAINIHRVGASILGVEFVRPT
 >PROKKA_01459 Relaxase/Mobilisation nuclease domain protein
 MLDRAILDTKLDSRPIDLSFLDDILKHMQKPRFSDSDKVDKRLLIKFNQFNPNSKFIKSN
 LVGTKTASSVLIKNIGQMKCSHLENALNYALENSETAYNEMFLECDKQFILESWLNDFDL
 TKDYNETMHLVFSIKDKPNEETMQGLLHSTWESLKIRLPEYKFALVPHAHQDHAHIHCFI
 NKTNQLTRRRRLRFKGHEDFDILPIAKARGF
 >PROKKA_01460 hypothetical protein
 MTYLEKIITYLSLVIFMLPLGFFQQTRGFLVS YVKKIISLTFYMPLLLLLVLFNSFALQYA
 IKVGSNEIVAKFGIIVVIGISLTFIQKVPENAINAIFGTQGGLTDAKS FVYQGVQIASAG
 AGSIAGSLKSVGRSAFGRTL EAYKDAKSTINSTANMRDMPGRC
 >PROKKA_01461 hypothetical protein
 MAAPLLALPFLSNPLVLGALAVIGTGAYLLYNKQDSL VVQADGLYSEILGFFISFSNKIL

KEIGEPLANVVQPFGMVLGVLLILLYSFKRYQNNDLFEIKTFLMLFVFGYLSLYHYAFK
SDGSSSGNGRSSFAFQNHVTEIFDAPANLLNAGISNVIKEYQTN SAGEYKKIDARHSITN
ANISFHVRQILSSLNKLYEDFIINNGLSLKTLLIAAVLLLVLGLGLEFFCYSKFSVMFL
>PROKKA_01462 hypothetical protein
MFKSRLNSWILLGILGILVVVFWVDVIKYKIEDLKHAHYLSQVKEREYYKNHIEEALKKD
SECFEKGDKVDCSAAMRIAAGERNRRMLEIK
>PROKKA_01463 hypothetical protein
MKALKTFLKKSFILLLAVALNHLHAVAMIVDNPTQNAWNGAKRAWDESKWAKHLATITER
IKLAQDTLDRANQTLNSINKVNDVLNKTNQFLTGSILSVNPMQYVEKIQSFSKQVQANA
ERIKENAQNYDIRNQIATKRISEKCP ELNWDVSQDASPT EKNLHQFFTSKGKESADTKAL
KDFANAIGNTQIGTATDLGAGLKG RALLEYICIQKGNLEAAKKIQLLDSQMTLALLNNDY
TAYEKLRAEKEELKRQIASNVYAKVKQLAVASQDRAFSQMDNELGVKTFGFNDENVKKG
YCKKENRNGKSECIPNMLNVNRLKMQFDELNLDSRDIAGKKGEAAAKVFNDYKHRFQQLS
VETALEIAQNLSFMNKTGLMAQMQSYAFKQQMGYFEDIIPADALKDDKEHQENLKQKQQ
EIEKVYRAKLDAYGFPNGSVGKASGVNSNSNNEAPSSDNIQSFNPY
>PROKKA_01464 hypothetical protein
MKNQHRNPTTKALMKTYPYNHFLFFCFILGAFLLGLLSPAYALSVITTK EIDANLPNGAI
ESRVVLGKRVEKVEAHGFYFRNNATNSIDIEITSLLRDNQSFLTSSAKTSLKIPNAKI
KKSTLLVLKGENAEVAKILGISKEEYQKLENIAQT NATNDPMYANTPFNNPNNSPNDNA
INGKDGANGSNGYGINGNDGINGGSGSNGNSNNNAVGSGIDTDGVLGVDGVNGSNSSSG
GSVGGYENNFTNHGSTNNNTGGYDNFNNSSSSGGGLGNGGFFPIPFNGDTNNSNPTNS
PTNGSSSNSAINPNSQENNYSSQYCKAPKLSPNNTMKLDVIAKDGSCISMNALRDDTKCA
YRYDFEAGKAIKQTQYYYVDRENKTQNIIGGCVDLQGAQYAMQLYKDDSKCALQTTSDKGY
GMGKTQTFQTEIVFRGMDNLIHVAVPCSDYARVQDRIVRYEKNDKTQTLTPIVDQYYNDP
SNPNKQEILNRGIATQLSSQYQEFACGQWEYNDAKLEAKRPTMLKSYNKLNGEWVEVTPC
NFEAGIKSGAVVSPYVMGPSSKVLSDTTTSHYFKIERKNYGEREQCQKPYGVNRCQPQY
FHTNPSITDWSATYKTTTTQTTPYLRPAQDNESPTTYNLITTQTTINRTQSVLKNELHL
SNDYLKYVEIHQGYKDNLDLKSQGYIDWSNYLYLNQNEGFCSQYNIWSESTKIGKNWYHW
HDNKISCAQRLLIYRPFQG
>PROKKA_01465 hypothetical protein
MIDKARKNLKDLKSIFLYWYEMFGSNDKSKDMMKAFMKDLAFAKDKNEMIKAQDNELVKE
FMQELNSLSFKEQLEFYQTNRDYDPSLDNDENSSELLNAFTGSLENLEKQEIALIRDQKR
ESGLQNNEDNKENGKLVPLDKKEASKEEELESKEKEPKETLGFNHS LGVSVLRSRVEEAK
NHFNEKEKGSNKRFSDFEFKKFLKDRAKNGMTNEEMQVLLLMINNKKSNLRPSHKEFLQE
CLIEAVLHNKNINKTTEYAIRKNIAHLPINKTESFLVRTYNHSLGILNENIAEKLKIPKK
EKEILQPLNANEEQALNTALENNEIDARTFKSWEFGSKEASKENASMSEELTQNNSKLSH
FNENSTTQNNIGQNNVKEPNNKTNNSKQNHSLSNLNLTSNTDYNKLVADM FENQMDANGN
YVGKKFSAQAYDNKTPYKVATGNASERSAYLYSKTKSKNVGMER
>PROKKA_01466 hypothetical protein
MANKERGHNALNGEPYCNLNDMLLDMEKNRLGFQSNAWVSLEEAKMLGASKEERDAIFKA
TQNK EISPVRLMFIKNKEPVPLVDNNGELVIDKNTKKPKHKQFAVIENGQKVFKPAYRDI
EPQAQFKFVYNVEMFPSINKEKIKPLNLDKLSNYSYETRLFHQKDYSQEKRDKTNIYED
LHRDLSPDNRNEALLERMRSYTLKNEKYNQLANETKQQTQTRSQSQSYQKKNKASSGI
ER
>PROKKA_01467 hypothetical protein
MNNLNNLSDEQINGMIDYLNILLEMGG LKQEPQNSNYSHNSEEDKAQKTPNQPDNLES
KETLLKAKSGQIN YVKTRIIQSLK
>PROKKA_01468 hypothetical protein
MQENENTLKQDAFYEAGFCDELSIETQH E EYIAQRDFAKSNPQGRRR
>PROKKA_01469 hypothetical protein
MQKKKTKNPQPNLFSILDKGDIATNNPIEESDKTNEIQEPLPYVVKTQINRASMINRNPI
EWAKYLSFEKRIYKDNSKEDVNFFFANGEIKESSRVYEANKEGFERRITKRYDLIDTIRNR
EFFPREIEILTYTNSLEELKEQGLEIQLTHHSETHKKTLENGNEIAKEYDHLKDIYQ EVE
RTKDGKLAREIIPSISSAEYFKLYDKLPFESTNNENTKLNTSDNLRDINNFEKAYKQRYN
KLVLSAAIIQAKKGSNERQNY SANNNNPINNTIETNTSNNI IQNNDNNHPNLSLADLEL
EQQNLGEPNEKERTNR TDEPSRARAGIPQEIHRRSEHGGRQEGMERS SDEDL SHQDSSLF
TEPREQGGTRGIYRSSDKQAVSEKPHRERDRIHEHVSRGDGV SARADARANSNGASSQAS
RMENGARSEEKGDNP SNERGIPQPPQSPSHQQNSSRDLGLSLSREQPGQTGRLRLFDHGQ

MGSLFPTDHENQRSQNNRELDKSSDRANENGDRSPRQNGSANQEGARGERFGIAPGSSNQ
SVLPLAQSRLYHARLSAPNGLRDLEKNRDQEGRLLSNLNHLESLLDAIRNNTIASEPDR
FRLLEAIQNNELPKDSIVGMQLLKDPPTTKIFYDKFQLKISPKKILEILENRIQKSIETTN
EALNAFNALDSQAIDLNAISNSVGLNPTQESKITDNSVELNNAQEQTAEQMDQENAPTT
LKQEIKEQETPITPAIPLNPKIDFKPSEEVLIKAKTRYKANIKAEILLKELQAKQEILK
GDYYATLKEQEILAQFSGWGGLESYFKKAQHPEEFKELNALLTKDEFRRAYLSARDAYYT
PKLVIDSIYQGLDQLGFNNDNHQKEIFEPSLGTGKFIAHAPSDKNYRFMGTELDPISANI
SKFLYPNQVIQNTALENHPPHQDYDAFVGNPPYGNHKIYSSNDKELSNESIHNIFLGKAI
KELKDDGIGAFVVSSWFIHHPENILGHLSLEKTRYSEINGEQIYKYELQALEDKSLDLS
QALNQAIEKLPKGVYQYHKTTLKTDALIIDANNERYQEVQKLIKNLERGELVKWDNLYFQ
LEQNNERGIFLKPTKINSKVQDSRLKAYFKIKDALNDLTSAELSPLSSDLELENKRVRLN
LVYDEFVKKFGYLNENKNRKDIKQDLYGAKVLGLEKDFEKEITPRSAKMQNI EPRQAQAK
KAQIFFERTLNPKKELIITNAKEALIASINQKGCGLDLHFIRDHFTTQSLETTIKELLEQK
LIYKDHKDNNGGYILANDYLSGNVKKRLKEVKEAINQGVEGLEANVKDLELIIPKDLKATE
IMANINSPWIPTQYLEEFELIELSANHYEKQYGDKMTDYQLGNLKEDIKVEHLNGTYEVFA
RNNELNELYGIRHKDRAHSYKAPFESLLNKVLNNKDL SVKYAQVDPNDPKKEIFITDEEQ
SNLSRQKAEELKEAFKDWIYKDYARRTHLEQIYNDTFNNSVLKTYDGSQLELEGFNHHIS
LRPHQKNAIFRTIQDRAVCLDHQIGAGKTLCAIASCMEQKRMGLVNKTLIAVPNHLTKQW
GDEFYKAYPNANVLVVD SKDTTEKERELLENQIANNNYDAVIAHTHELLSNPRGIIGE
LKEEELVNAEKNFERQELAYKNNPRETKKPNERAFKNKLDKIRTKYDAILEKQGSHIDIS
QMGI DNLI VDEAHLFKNLAFETSMEKIAAGLGNQQGSNRARDLFIKTRYLHQNDKKIMFLT
GTPIANSLSEMYHLQRYLTPDVLKERGLEFFDDWAKTYGEVVNDFELDTSAQSYKMVNRF
SKFSDVQGLSTMYRAFADIVSND DILKHNPHFV PKVYGDKPINVVVKRSEEV AQFIGVAL
ENGKYNEGSIIDRMQKCEGKKSQKGQDNILSCTTDARKVALDYRLIDPNAKVVKFEFSKSY
AMAKNIYENYLETNATKGTQLGFIGLSTPKTHSQKVSLEVLDNAHETENKNPLDEAQELL
ESLSSYDENGNIAPSKKELENELKEKEAKGVNLDEEIAKGCLFDVYS DVL RHLVQMGIP
QNEIAFIHDAKTEEQKQDLFKKLNRGEVRVLLGSPAKMGVGTNVQERLVAMHELD CPWRP
DELLQMEGRGIRQGNILHQNDPENFRMKIYRYATEKTYDSRMWQIIETKSKGIEQFRNAH
KLGLNELEDFNMGSSNASEMKA EATGNPLII EEVKLRAEIKSEEAKYKAFNKEHYFNES
LKNNA SKLDY LKQELKDLETQSSV IIPHTHEIKLYDLKNEESKDYELIKVKEVEPLKEN
ASMSEELTHKKLKEQNKQIAEQNKEKLDAIKKQFASNLNDLFFNEERDYK LLEYKGFVFN
AYKTKYQVEFSLSPKDNPNIAYS PNMVYKNDTANMFSSYNFCGEIKFDGFLKRLDNAIT
KLPEKIKELKNSIETTQENIVKYTRLVEQKPSYPRLEYLQALKWDHKT LIDDLAKMSKDR
DYKPVFNPKSKEVLEKMNAEKRASLVDES VT EEI KEQANQEAHMPMKKAANDDYDMGM
>PROKKA_01470 CobQ/CobB/MinD/ParA nucleotide binding domain protein
MIITIANEKGSGKSTLCLNLCVQLLLDKKDIAALDTSQKSLEVFNNIRSETSLPNFTL
FNRTGNITDTLQMMDKYEYILIDTKGEHSKESQRAMLLSDWVLIPTTPSQLD TAVLLDM
LERIRDIQALNENLKACIVMNRIPTIPTLKEKKALIDFINQNNANESVFLMDNLLSERIA
YKRSVSEGGMVMEYNDNKAKNEWFQFYDELSGYLGKK
>PROKKA_01471 hypothetical protein
MEFKNTKKDR LSDLENRFENANKHETNKHEKEDRKAHTLYISEKVMNSVEEYLN EFGAF
RENKSVFVQDAIVFYLEHKKKEMQMLLDKASKL
>PROKKA_01472 hypothetical protein
MKKTRLLKLRLASLACLMGLGVSGCAFLDKQILNDHLTKAKNNPKYDCQKEMGSFPKKYD
GINQCLKAQEELIEPIITKKIDQYQCGDFTNEGLKDKCFKRNDAYLNTLLTPIIQKQERR
FSCSDFHNPQLQE QCKDKTNAYEKQKDQQRLINLVQLEAFEKEYAQYQSYIIPYFTKEC
VKNAPHLANKERLCQKEVHEKFHDPYSSSKELS VKSAISFCIKKVDAKLEKAALMNGVFI
SPYKKSTHCQORTHLDNKS LKEIALDMNP KLEKQSPFIDANKLAMQSAGLLRKNKDALIAF
ATDICMERNEHKKGEFISLKESCTQSQA KIYNNKERFDKFIQDYQKDLKTCLLDTSNTKE
EMEQNVSQCQKEQLRDDNKGLGFTLEELVKKYAK
>PROKKA_01473 hypothetical protein
MGKDFESDELFDY EIIKPKKTLKIQYTYAKRYYKEVEKFAKNLTQLTQEEFMRLREPQKQ
VVIKNIGNMTRLHSKRAMDYTA KHGELVRDEFFNEVNYNDIAEQWNEQFEK LLENKSRVK
NCALHLVFSIDENCNEKNLKAELSVYQTLTNTLGYDYPFIMKLHTHQNNPHAHVIINKT
NKITNKQLRFNSKDSCKEFYHTLRET FKYDLFANSKGELQYSNTPNIYTAIKN IETELDA
LEKQARNNSFRHENYFYKVFGSATSQIDSLKKRENALLDHLDSLKNLLEKTHWEKENFT
PPINEKELNRQLKEIKWLNKESLTPKN TYKKTQKL VVCKSRLIKDYLYTTKKLFTTQKKI
IDLEKDYKNLKV LKEEFSKDLETDLSSHKKRFELYTKLKSMSKVFTSKNIVKNLEKIALD

FKSDRHSISQRAFEFFRYMNYQNLSLIDKGNMFLVAKFFKDSALLVNIARFEMKKIDDSV
KNSNPQDNLLDKQAWLNLLLEHLKRLEENYCFAKKRKEFLETRAMELSKDLKFLTQANEN
DLPIYERGQDKI IKRCEKSLNFLQKELQCFKTLKLSASIALENLQNNHQITAVTQDTQE
NTNALKKIPLKDFNKTNELISPNNYEMDF
>PROKKA_01474 hypothetical protein
MQIYKPYCYIISILAIPFIFLGVSHYFLDFPSFLTSTYNYLKKIILNLHNFPHFKFEIYFS
FFLCVITAIFDSSLIF
>PROKKA_01475 hypothetical protein
MTKETDIDLIRNYLEKMASIIKNSQSIINAPHFLKDENQAKILQENLKENFTLLMEYVSE
VWVKSEEEFEFMKDESFENDFMHFMSLLDSKIKVE
>PROKKA_01476 hypothetical protein
MKNPRLYRKSQKEGAKRMEFLLDVRCVQILNDLKTRHNKSYTKIVENLIILNSQISFYKQ
QKITSTLGLSLAMLYSSLNATCSNFNQIAYHLNNAALLGENVIHLGLLQDIEIQLKAWSEK
TKILNLAIKKSAFIVAHCLKGDKKIFKGMNL
>PROKKA_01477 Type IV secretion system protein VirB11
MISNLLKSYLDDFMPILSQPNLNEVVFNKEKEYFLHRPKEKVKCFNKKFTNDYLLVFCEQ
LAIFRNQKFTLTKTPKLNSTLPYTQIRINALHPSIIASSNISINIRVPSNFKFEINAFKLS
EKCLAKNITYDFLLNLVSVGNILISGGTASGKTSFVNSLIENIPKNERVVTIEDSPELK
ISNEDQVNILVDKSGSGFFTYEDGLNAAMRMSPDRLLLGEIDHTNTALFLRLANTGHSGM
ISTLHANSVVDAFIAICQNINLNKGGKPTPKETLLDYFCTGMDYVIQIKKKGSDRIIDDV
LNVKSEFFKKELML
>PROKKA_01478 hypothetical protein
MKEKLDKLLKLELNELDKLDLEKELSNLKLTSHKIYQEYLLEKHEICEKNLEIKANNKRI
SKIHHLYSIAKRIEDKREKERIELQKKMLRDFDNHQGEF
>PROKKA_01479 hypothetical protein
MIDENLNCDKKLAHLNEELPILEAPKKDKPKPKPRNSKIKDLDFYASYVSKNVEIQKKFD
ETYRLCEQKQQLNNLLKELENLRKQL
>PROKKA_01480 hypothetical protein
MKLNPEEELLNTLEESKDAEYVELCKKVVEAQMELRKVMDKRLNETKESRSRMQARITE
RLVSEILNTEEHFSKKIIQNEKPLEIIFDIERIYNND
>PROKKA_01481 Bacterial conjugation TrbI-like protein
MIGTLHNKYWERYGIPLSLSTLSNGLLIGLTSGLTEAMKNKKGGFNQNYFGDYMMQMTR
QTGISLNNIIAQIMRDQIRIKPIITIREGSHIFISPNTDIWFPIPKNNEVLAKFFNEEKE
QNNDTK
>PROKKA_01482 Bacterial conjugation TrbI-like protein
MGFFRVIRNAPIVSVFIVLMLIFCAWFFLYDDSSPKAVRNIVKSKFPISDYLYKKADKKK
IPPQNNVIKQEIKEEEIKKCKGQITDNKGNLIDDCNIIATQKQVNNLNLNPKDILKTK
PTVASNPSNYNTANEEETRLVILASRISSQKETQPPTSINKSVSNTKSKEKQELEKEYG
SSGFKNFKRKNIASSENKLLRTITADMIPAILITPISSEIGGSKIVAQVESDIYATMGR
AVLIPKGSRAIG
>PROKKA_01483 hypothetical protein
MVGNYIIAEDVSKKWLTRSGKEYVCVRRDKRKYQKSKDFMRLKRLLEQDEAMKSKRNNMH
LIEDTTPQPKNAISVTELLAQIKNQTPNEKCRALNEEEIKTLEEVKNIIKAPKPLQKDLK
NKLKKHTIKTHYCVPIPIYIERPKPTDGLEKNKKGLNNTKKQNQANQK
>PROKKA_01484 Conjugal transfer protein
MKKFTLSLFLYCALLSAEEDIFRNNTNETDPTSSFEHGKESNNLIPAKSDSLESFKEQEN
KEKAKQLMDLKALQSVYFSKSRKLQDNNFNVLYVAGNTNKIRLRYAMTTTFIFDNDPIIY
VSLGDPSPGFELTYPTNDHYDLSNMLVIKPLLIGVDTNLTVVGASGTIYTFYLFSTYTSK
FQSYFSVFVSNKRAIGKLNILSKNELEKREQEQLAQTTETNTNSPNKQDKQKLDYRKINYE
SKEIDDGKFIRIGDEVNHIFIEKAKINRGYLQKPKHKRTWWSLWLYKNPAMMRI
>PROKKA_01485 VirB8 protein
MRDASVIYHLEKKLGDYIFYVACFFFGTTALLIILMTVLLPLKQKEPYLVQFSNNKENFA
LVQKADSTITANKALIRSLVGAYVLNRESITHIEQHEKMRQNTIKEQSSNEVWYEFEKLI
AHYDSIYTNPLLTRKVKIANIYLDKDLAYIDIEVSLSYHSGELESKRYKVVMSFEFKKQE
INFDSMSLNPTGFMVTGYDVTEIAILKDLDEKNKVKDDGVKSRIIHTTEKKDPHMSQYKDV
KEQ
>PROKKA_01486 hypothetical protein
MDLWGDLDKENEEQTLLEEIENELKQLENKENLKAADNNTFKEENQDTKENQPNDLFLSLP

LPTQTTINGIKEFVEEPIVETEEKETPQNEPIQEKKKNF

>PROKKA_01487 DNA topoisomerase 1

MKDSVIIIESPNKVEKIKQLTGAEVYATIGHFTNLVGVDIENNFACKFEIKEDKAKKINF
LINACRGKKVYIATDPDREGYGIGYQFYEKIKNVASEIYRAEFHEITPSGIESGLKNAVL
FSRSNTNLYQSFLGRIASDQVVGFALTSYLRKSLNLSELSTGRVQTPALALICQRDQEIR
DFDKLDDKEKVDYQIQAKIVYNEKEAIIKHVRANEKNELVDFKFKDKNEASQFLEDLKD
LGSM SVLVS IKESLSNKEPKKPFTTSKLLSQASKALKIPTKEIAQLAQKLFEVGLITYHR
TDSEFLSPEYLKEHEAFFEPIYPSVYQYREYKAGKNSQAEAEAIRITHPHALKDLEKVC
SNAKISEELALKLYQFIYANTICSQSRNALYNQYDCIFKIKSESFKLSFKLLKEKGFLEI
EELIQGKEELNKEEQESEIENFSLKENDSVPLKEVFIKKIEKPSPKPYKESAFIPLLESE
GIGRPSTYASFLDLLLKRKYVSIIDTKTNAITPTSQGLEVISFFKKDKVEVDFIALTSKDKS
KLGSTTKQFEECLDLIMRDEASYEFKMLEVISKLKSTAKFYQAQADNNMPTDKQLELID
KICKDKKLQKPSQEILQNKEKCSDWIADHVKEKPSQKQLNLVKKICEECKLAMPTNKILN
DKFAISKWIDEHKKTKK

>PROKKA_01488 Type IV secretion system protein PtlC

MRQEYTQEEISNPFALVINLWEQGVENVYQNFYYLIFETKNDSIKGYLERFKKKITTNE
LENIQENDKQDEVKYQENYFIGFDNFNLLDKSKILDNIINN VKNMLSGLFVEKIDANSL
NFYAEYINGVPSEYTFARGRLHDGMINSDVHFKKDFFTHIHNGKEIFKRKFISIKTYDIDK
IVSTALSSVLHLSLEFDVILNIDNIPKAKAEKRIETKRKRANKSIRVEIDELNDLIKTDR
VLMQEISLNLVHAKTKTDLDSACIEITNLLKQKGIVSTQESIGMLPMFFSFFPNRNRLN
FRKRLSSQNIASLIVLEKQNCGFSRNSWGDRLHTIFRNQDKSPYLFNFHAYEAKRKNDM
VSGHTIIFGGTGAGKTTLIEFLITNCFKYEDLSILALDRNNGMRVMTEFLDGQYNDSDNF
YINPFSLKDTSANCTFLVSWLAFMLNIDEDTKDEKEKKTLSALSKTIRVCYNNITKQGGA
IFKL RDFKDTLERDYLD SKDILEKESSKDLYKHSTDCLDFAKKLSVIDMDTLSSSKKDFN
LAVLYLFYKVMNRAKLENKPFYIFIDEAKSVIENPIMLAKIKDTLAQARKLNGVLTALFQ
DINQLDGVGAKSIIENAAQVILYPTNNFEKLESYGILLSPIEKSFLNKTPLNARQILVK
NLLTQSSVFLEINLEKLNSTNKKFLRAYSSSASSVFEEMLKKDNP KDYKEIYLTKE

>PROKKA_01489 hypothetical protein

MLEKIFNSLCSLVPKRYSMKEETNIVGIYNEHYLLSEKSNLVGALKLEGISYLSLDDKEI
AQKFNERILALNEIVEECILKS

>PROKKA_01490 hypothetical protein

MALHIVCVESINIREISKRETFLGLNLDSWIVSFLMSAWLLSFQPLYAMGLMASCVFVFY
ILEFFDEDITAILHALFTLRTGARNYYA

>PROKKA_01491 hypothetical protein

MFKQIRRMMSLA ILMPSFLLAAPDYKQKFTQILD FISISNDFIKAIGGLIIVGT CIYAYKNW
DRLGEIGWKCVGIIIIITAAISNAKTL SQWLF

>PROKKA_01492 hypothetical protein

MAYIANLQFLPLDEILLSNGYKHKVEKSSKNNPALYNEYDTIKGTLIVSRKPDGSYHYFN
THNDEDKGTIIAFCKNRGIDFNDLIKNRDDLEISDKPNHKNYNSKPYTIITKEQEAERQK
VVKQFEKLRYYDIGSSDLFHTLLDSRSLDKSLLKPYNHLLKEDEHANLCVPCYLVNDDGN
LIQSGYNKRLSKPFTMQGATNPTKHLMQAGSIKGFV LAPQNIKNIQNIIVAESIFDGLS
VLEMQWLDPNQTTIISTIGNFTEERMQKFVDKFLNTINTYKCKEYNEYHKEIDRYNKQLE
DYENYTNFQKRYEKWRAKELAKHQNNPKKVPNNPIFSPIRDKNNLIPNPICKPSALALRL
KEPKIPNLSLNVILAMDNDEQGRKFNAILEKIFYAHTKEIPNIYTPISKDANDDLKIAKA
LGLKQISNRILQDFLFDAKEMQNP TTTTPSQKSILANQLQKLQDLMPKPKLLQGVFHGYQ
QDHGFGSKYVANSVYYTNNQSQNKGHER

>PROKKA_01493 hypothetical protein

MQFPLKKDLKISQKYVIPKGFDP TNKEYKNLLVKDMPKSVYREIEIAMGLLFEINEVETA
IEVMKFLTQAMDRVLYDARARHYEMRLAKEIKKS RFG

>PROKKA_01494 hypothetical protein

MQERVFKRKVLDANILKEMHANNVCYSKHSKDRFIPFIFDKFGYVECKIFKKILNFP SNT
TFFGGTGCKKLMELLC EIVIDSRSSKIALNRHYALTRLQWCDRTL RHNLKILERIGFLTA
FKNKKGYIFLSMHDFTKIENYEHLLNGESNLPNSFFLGICGYLKKLKFKKLKDRVFKLLN
KHGAFFLKIPKHFQM QNFNNIFLEFVSVNNSCLSYRLTYEQLVGKKNSKYQVLLPTS NCK
KEYP

>PROKKA_01496 hypothetical protein

MAQLEDLKAHEKYNLLLC LFKSDYYNDPTNTDLEGKPFKQACEECTLTFFRPRMDKVAQE
EGFQKVRKADYAVAIVYEKSTKESIVQRSVFALKIDHDTIGTDE

>PROKKA_01497 hypothetical protein

MELLDLDGVIEKGVFEIPSYQRGYAWQKEQLKDFWNDLEHVSKLGDKFHYMHSLTTLRELE
NEFESSAFEIIDGQQLATSLILLGLLAKITQHKDPKYYSMNLEPILSYKYYGLSEAFRA
ITEEEKDLKAFKTSFYAKNLIDAYEFFQEKISDTPVGTLEKMF DALTKKMLFSVVELNDN
RIDPFSSFETINNRRGKDLSTLELLKNRLHFVSHKICDEEDLENLQQEINDTYTRIYHDLR
FFKDDHLEGFLKHFFVAYYYGENSKFKERLLNTAFDAHKKYHSSYDEYEKINDLLLYLSYS
SKVWVFLHMPNDKELDKELEIEITPKMCGLLDKMRRLNALSENAFLPLLLSLLTIQLAVR
SGSERHYTTKELEGLLEYLERFGFLIYGVAGRDTAKNEWIESAFEAFQAYKEGKGNIAIE
DLPTLEKNFFNKHNNSGLEFLEESIHSKKKPRKWYEWGKALNYLLYEYELHHPETTLNF
DSSIESIEHILPQKPDQGYSAKEKSWAKNPHIMHALGNLLLI AKNANSSLSNKPFE EK RK
QYLKGSYSEKEVAKNASFGITEIQERSEKLLD FLIARYRIAE LVGESAIKAFKNAFLKDI
E

>PROKKA_01498 hypothetical protein

MGFQENENKLKVGALVKATINNKVVEAKVISIGFNRVTLRSEKGNEATYAFNSDKFLKWFK
EVPLNEVAKNHAEKSGADLLKNVKIVTSGPSVKDRTTTAKGKENRFKLA FG FREANEEGV
SVSELMIRDYTLTERKRRLGVLMSPLFSGNGSQISALIITALANAEGFSKHSDAEWVKM
IESRNEELCEVESFDNLDRVALTLYCNVIKYAEGREEFQND FND FSPDGFWAKYLPKNK
NEALFVAQLLCDGGINKYGLSCAGLTENLLKDIFNNIGLATAKEVDEHLANMDKEEAE LK
E

>PROKKA_01499 hypothetical protein

MLINAVIEKDENG YFAFVPFLKGCVSQGKSYEEALRNIKEAIELYLGDLEADELAF LSKK
NSVIAPIEIAFA

>PROKKA_01500 YcfA-like protein

MPELPRLTAKEAEKLLLQNGFVFSRQKGSHRIYVKDKIRQVLPFHSGEILHPKIVKEIME
NILK

>PROKKA_01501 Biosynthetic arginine decarboxylase

MQEVHDYGIKFWSNNEFKIEKGLVKVCHGKNPSLLEIVQSVRDKGYRGPLLVRFP HL VQK
QIKSLF DAFSSAIKEYQYSGAFKAVFPLKVNQMPSFVFPLVQGA KGLNYGLEAGSKSELI
IAMS YTNPKAPITVNGFKDKEMIELGFI AKSMQHEITLTIEGLNELKTI IAVAKQND FVA
CPKIGIRIRLHSTGTGVWAKSGGINSKFGLSSTEVLEAMRLLEENDLLEHFHMIHFHIGS
QISDISPLKKALREAGNLYAELRKMGAKNLSVNI GGGLAVEYTOHKHHQDKNYTLEEF S
ADVVFLLREIVKNKQEI EPDIFIESGRYISANHAVLVAPVLELFSHEYNEKSLKIKESNN
PPLIDEMLDLLANINEKNAIEYLHDSFDHTESLFTLFDLGYIDLIDRSNTEVL AHLIVKK
AVQLLYVKDHNDILRIQE QVQERYLLNCSFFQSLPDYWGLRQNFV MPLNKLDEKPTRSA
SLWDITCDS DGEIAFDSTKPLFLHDIDIDEEYFLAFFLVGAYQEV LGMKHNLFTHPTEF
SVVFDEKGDYEVEDICEAQ TILDVLDLDYDTKEIERLLKQKIEDNNQLDMEEKKEIMGR
LYVMLSENGYLRTIS

>PROKKA_01502 GDP-mannose-dependent alpha-mannosyltransferase

MVIVLVVDSFKDTSNGTSM TAFRFFEALKKRGHV MRV VAPHVDNLGSEEEGYNLKERYI
PLVTEISHKQHILFAKPD EKILRKAFKGADMIHTYLPFLLEKTAVK IAREMQVPYIGSFH
LQPEHISYNMKGWFSWFNMMLFSWFKSSHRYIHHIHCPSKFIVEELEKYNYGGKKYAI
SNGFDP MFRFEHPQKSLFDTPFKIAMVGRYSNEKNQSVLIKAVALSKYKQDIVLL LK GK
GPDEKKIKLLAQKLG VKTEFGFVN SNELLEILK TCTLYAHTANVESEAIACLEAISVGIV
PVIANSPLSATRQFAL DERSLFEPNNAKDLSAKIDWWLENKLERERMQNEYAKSALNYTL
ENSVIQIEKVYEEAIRDFKNNPHLFKTLS

>PROKKA_01503 fatty acid biosynthesis transcriptional regulator

MQESVVHV DYSLETCKNFKPSVGT ELVVLEKDIAHARFKGNESMVYEENFVHAGFVLIA
CNYAALCALNKRHSVVVSNNINFYAPLELNQEALIKAQVIQDGVKKAEIKIEAFVLDIQV
LEGMIEIVVFDKKPFKFNFKEE

>PROKKA_01504 tRNA (mo5U34)-methyltransferase

MLICNDKFNPKTLL EIMALKPWRKGPF EISQIKIDSEWDSSIKWDLVKNATPLKDKVVA
DVGCNNGYYL FKMLEYKPKSLVGFDPGVLVKKQFEFLAPFFDKEKKI IYESLGVEGLHEK
YPNAFDVIFCLGVLYHRKSPLEALKALYHALKIKGELVLDTLIVDSPLDIALCPKKTYAK
MKNVYFIPSVSALKWCERVGFENFEILSVLKTTPKEQRKTD FILGQSLEDFLDKTDP SK
TLEGYDAPLRGYFKMLKPSKL

>PROKKA_01505 hypothetical protein

MRLGVNEAVELSLGELQNTPSISYFNSIVLSLNKVQKGS L FVAKDHTLIPKALELGAYGI
LYAGEYPLSDRDVAWIKLDIEHSLNHLFKFCLLNERVVGVLLSPIELEIASKIIVSEFV

WCLKESLEDLFIIIEGCKIAFFDKLEWLHLFYKQERLKEDLKEDLKEDLKESRLIILNQSF
FCSALVYEKQEYEFKMPICIFLEPLKRVIQLCEKLQIEFDLNLGKKEYPLDHCKPFFVNK
NLEIAPYGATARVVVAEASKELFEMMLQKALETLSWGKIVVFCRKNSAAFFKKTNPYCYT
TQNNLKEQLKNLAFNFAFIYGISSYHLESLLNPPLFKKTPTLW

>PROKKA_01506 Methionine--tRNA ligase

MQKSLITTPIIYYVNDIPHIGHAYTTLIADTLKKYYTLQGEEVFFLTGTDEHGQKIEQSAR
LRNQSPKAYADSISAI FKNQWDFFNLDYDGFIRTTDSEHQKCVQNAFEIMFEKGDIYKGA
YSGYYCVSCESYCAISKADNTSDKVLCPDCLRETALLEEESYFFKLSAYEKPLLDIFYAKN
PEAILPIYRKNEVTSFIEQGLLDLSITRTSFEWGIPLPKMNDPKHVVYVWLDALLNYAS
ALGYLNDLDNKNMAHFECTRHIVGKDILRFHAIYWP AFLMSLNLPLFKQLCVHGWWTIEGV
KMSKSLGNVLD AQKIAMEYGIEELRYFLLREVPPFGQDGD FSKKALIERINANLNNDLGNL
LNRLLGMAKKYFNYSLKSAKITAYYSKELEKAHQILDNANSFVPMQLHKALEELFNVDY
FLNKLIAKEEPPWVLHKNNSEKLEALLSLIANALLQSSFLLYAFMPKSAVKLANAFNTEI
TPDNIYERFFKAKKLQDMILQDTEPLFSKMEKIEKTEKAGEASPEKNEKEKKDTKEKAPLK
QKNYISIEDFKKVEIKVGLIKEAQRIEKSNNKLLRLKVDLGEGLYLRQIISGIALDYEPESL
VGQMVCVVANLKPAKLMGEMSEGMILAVRDSNLA LISPTKEK IAGSLIS

>PROKKA_01507 Cyclopropane-fatty-acyl-phospholipid synthase

MISKFLLKSMFKQWKNGDYQVVFWDNSVYRNGEHS PKFTLKIHRPLKFS DIKKDMSLTIA
EAYMDGVIDIEGSMDEVMSLYLQTNYEHLHKHDNAKAIQKPIKESSNISKHYDLGNDFY
SIWLDETLSSYSCAYFKKDDTLHAAQLQKLDHTLKKLHLKPGEKLLDIGCGWGYLSVKAA
QEYGAEVMGITISSEQYKQANKRVQELGLEDKVTIKLLNYQDL DGRLYRFDKVVSVMFE
HVGKDNLPFYFKKVEVLKRGGMFLLHSILCCFEGKTNAWVDKYIFPGGYLPSLREVMSV
MSECDFHLLMAESLRIHYAKTLDIWRDNFNHNL DQVKRLGYDERFIRMWDLYLRTCASAF
RVGSADL FQLLLTNSVDNTFPLTKEYIYQ

>PROKKA_01508 Low conductance mechanosensitive channel YnaI

MRLLLWVVLVLSLFLNPLRAVEEHETDAVDLFLIFNQINQLNQVIETYKKNPERSAEISL
YNTQKNDLIKSLTSKVLNERDKIGIDINQNLKEQE KIKKRLSRSIKGDNFYTFMKDRLSL
DILLIDEILYRFIDKIRSSIDIFSEQKDVESISDSFLLRLGQFKLYTFPKNLDNV KMHEL
EQMFSDYELRLNTYTEVLR YIKNHPKEVLPKNLIMEVNMDFVLNKISKVLPFTAHS LQVS
KIALALMILTLLGLRQLITWLLAKFLDWIFEIMQRNKTMHVNVQKSIVSPVSVFLALFS
CDIALDIFYYPNASPPKVSMWVGAVYIMLLAWLVIALFKGYGEALVTNMA TKSTHNFRKE
VINLILKVYVFLIFIVALLGV LKQLGFNVSAIIASLGIGGLAVALAVKDVLANFFASVIL
LLDNSFSQGDWIVCGEVEGT VVEMGLRRTTIRAFDNALLSVPNSELAGKPIRNWSRRKVG
RRIKMEIGLTYSSSSQSALQLCVKDIKEMLENHPK IANGADSALQNVSDYRYMFKKDIVSI
DDFLGYKNNLFVFLDQFADSSINILVYCFSKTVVWEEWLEV KEDVMLKIMGIVEKHHLSF
AFPSQSLYVESLPEVSLKEGAKI

>PROKKA_01513 Neuraminyllactose-binding hemagglutinin precursor (NLBH)

MKKGSLAIVLGSL LASGAFYTALADGMPAKQQHNNMGESVDLHFHYPIKGKQEPKNGHLV
VLIDPKIEANKVIPENYQKEFEKSLFLQLSSFLEKKGYSVSQFKDASEIPQDIKEKALLV
LHMDGNVAILEDIVEESDALSEEKVIDMSSGYLNLNFVEPKSEDI IHSFGIDVSKIKAVI
ERVELRRTNSGGFVPKTFVYKIKETDHDQAIRKIMNQAYHKVMVHITKELSKKHMER YEK
VSSEMKKRK

>PROKKA_01514 hypothetical protein

MGMIILKN SAKSGIFWVRLGIDFKEKERKEVNEKR

>PROKKA_01515 GMP synthase [glutamine-hydrolyzing]

MILVLD FGSQYTQLIARRLRESGIYTEIVPFFESVENIQKKAPKGLILSGGPASVYAKDA
YKPNEKIFDLNVPILGICYGMQYLVDFFGGVVVGANEQEF GKAVLEIAQNSVIFEGVKIK
SLVWMSHMDKVIELPKGFTTLAKSPNSPHCAIENGKIFGLQFHPEVVQSEEGGKILENFA
LLVCGCEKTWGMQHFAQREIARLKEKIANAKVLC AVSGGV DSTVVATLLHRAIKDNLI AV
FVDHGLLRKNEKERVQAMFKDLQIPLNTIDAKGIFLSKLKGVSEPELKRKIIGETFIEVF
EKEAKKHHHLKGKIEFLAQGTLYPDVIESVSVKGPSKVIKTHHNVGGLPEWMDFKLIEPLR
ELFKDEVRLLGKELGVSQDFLMRHFPFGPGLAVRILGEVSESKIKRLQEADFI FIEELKK
ANLYDKVWQAFCVLLNVNSVGMGDNRTYENAI CLRAVNASDGMTASFSFLEHSFLEKVS
NRITNEVSGINRVVYDITSKPPGTIEWE

>PROKKA_01516 hypothetical protein

MSV FQMSL KCCVGLVLSMGVLLGDFKAFKVRVDKSLAPPFLNVLSLAFKQDMKKEIIFVF
TKSNKLSKKVLCDFDAFLLPEALMSGMPKKVLFHKEFLFQSKESKTLYAFSLIDSQYCSK
GGNYRYELEKLERWFVQKAPELAESHVRVDYKSQYDKTQTKKQK

>PROKKA_01517 Dimethyl sulfoxide/trimethylamine N-oxide reductase precursor

MSISRRSILTKIPIALASTNFLKATGVFEKVESIPHATHFGPFIAKVQNGVIKDIVPQKS
DYNPTMMLKAMVDRVYSDSRVKYPCVRKSFLENKKNHKELRGREEFVRVSWDVALDLAAK
KLKEIPKENIYNASYGGWGHAGSLHRCHHLAWRFFNTTLGGAIGTDGEYNGAAARINPM
IVGDMEVYSQQTTHEEMIKNCKVYVMWGADLLKCNRIDYFVPNHVNDSSYPKYKRAGIKF
ISIDPIYTETAQAFSAEWIPIRPNTDVALMLGMMHYLYTSDQYDKAFIAKYTDGFDKFLP
YLLGESDNAPKTLEWASQITGVSAEKIKELADLFVSKRTFLAGNWAMQRAQYGEQPDWAL
IVLASMIGQVGLSGGGFGFSMHYGGNAQASSGARIVPMISQGHNSVKSVIPASRVSEAIL
NPDKEIDFMGKKLKLPKIKMIYNCGADLLGHEADTNELIRALRTLDCVIVHEPWWTPTAK
FADIVFASTSTMERDDITFGGSYSKNVVYAMRKVVEPVYESKDDYEIFRQLALRIGGNET
EQKFTEYKSYMEDIKGLYEKSDGPTLKSFDQFWRDGFVEFEIPENARKFVRHAKFRQDPI
NNKLDTESGKIQIFSQKCADFKLADFKGHPAWFEPAEWLGSKMAETYPFHLISPHPKYRV
NSQLDNTWVRNVYKIQGREPVMINELDANKLGIRHGEIVEVFNARGRLLAGAFVTKNIRQ
GVLSIQKGAWYDPEDARVRNPRCNGGHVNTLTSLRPTSSMTQAISANTALVNIRRLRKYE
LVKPYHSISTPSIIGA

>PROKKA_01518 hypothetical protein

MDILDNLKAQAVQQNEREAEDKEKESKEPVVLEDLSALAWLELEEFSSRLSGLPKERILEL
VNLGKIKSKISHNKLIDASSGTNALIKKVENNLI SMDMNGCSLEPVFVEKTINTILNLH
DKVIGAKDETISAFKNENMFLKDALISMQEVYEEKTIDLLRDELNQAREEIEFMKRKY
RLMWGKVADMSSVNKK

>PROKKA_01519 putative cysteine desulfurase

MQAFLNRSFAPLLNPNESVLEQVKSSIIILKGVSYFDWGASGLASALVEKRVKSLLPYYA
NAHSIASKHAILMGMLLKECQEKLRSLNLSDGYCVLSAGYGASSAIKKFQEILGVCIPS
KTKKNLEPYLKDMALKRVIVGPYEHHSNEISWREGLCEVVCIPLNEHGLLDLEILEQTLK
KSPNSLVMSAASNVTGLLTPLKEVSSLCKKYKATLALDLANFSAHANPKDCEYQTGFYA
PHKLLGGIGGCGLLGISKDLIDTQIAPSFSAGGVIKYANRTRHEFIDELPLREEFGTPGL
LQFYRSALAYQLRDECGLDFIHKKENNLLRVLVHGLKDLPAINIYGNLTASRVGVVAFNI
GGISPYDLARVLSYEYAIETRAGCSCAGPYGHDLNLNAQKSSDFNAKPGWLRVSLHFTH
SINDIDYLLDSLKKAVKKLR

>PROKKA_01520 HIT-like protein

MNVFEKIIQGEIPCSKILENERFLSFYDINPKAKVHALVIPKQSIQDFNGITPELMAQMT
SFIFEVVEKLGIKEKGYKLLTNVGKNAGQEVMLHFLHILSGDKH

>PROKKA_01521 Phenylalanine--tRNA ligase alpha subunit

MHTLIERLEKVTNSKELEEARLNALGKKGVFADKFNQLKNLNGEEKNAFAKEIHYYKQAF
EKAFELKKKAIILELELEERLKKKEIDVSLFNTIKTSSSHPLNHTKNKIEFFTPGLGYKLE
IGSLVEDDFHNFSA LNLPYPHPARDMQDTFFYFKDHKLLRTHSTSPVQIHTMQEQTPPIKMI
CLGETFRRDYDLTHTPMFHQIEGLVVDQKGNIRFTHLKGVIDFLHYFFGGVKLRWRSSF
FPFTEPSAEVDISCVFCKQEGCRVCSHTGWLEVLGCGMVNNAVFEAIGYENVSGFAFGMG
IERLAMLTQCINDLRSFFETDLRVLESF

>PROKKA_01522 Phenylalanine--tRNA ligase beta subunit

MKLSVNDLSTFVDVPKDIACLCEDLSCGLEVESCIPIAPKNVVVGKVLEKAPHKNAEK
LSVCQVDIGKEVLPIVCGAKNVASNQFVPVALNGALIGSTTIKTELRGVESCGMICSSI
ELGFPKINDGILELDESVELVLGKELNEYAPFNTHVLEISLTPNRGDCLSVLGIAREIS
AFYHTPLKPIKALNFTPKSDWITLSAGENIESHLAYLVYNHSLKTPLNIKLSLAHNNAL
SENDLNNFIEFSTHFGSVIMNAYS LNKTPIDLSVKNDENNLESVYINHQRKSIIAIKHQD
QKDLSEYLLLEASYIDPISLSLKLHALKDKTLQKDNALIYRSARGSNPNLSDGLNFLSAH
LKATILESKQTEHSLKDRTLTFQLEDITEILGLAVEKEKIQGILKSLGFKVSVKEPNSNP
QILEVVAPNFRHDIKTIQDIAEEILRFVGDNLVSKPLNCVSSKNSNP HYDTHRFFENLK
HKALACGFKEVIHYVFYSKEKQKLGFEVLEDPLELQNPITTELNTLRSLVCGLLDASL
RNKNLGFKSIALYEKGSVYNSKREEIQKLGFLASGLQKKESYPDAKGKAWDFYSFAECVS
KVIGDFSLEKLTAQTPINHYPYQSAKIIQNHEIIGVIAKIH PKVIQELDLFESYYAEIDAS
KLKRPAMLLKPFISIYPSSVRDLTLIIDENTAFSKI KALKDAQIPNLSEILPLDIFKESH
NTIALSVRCVIHSLEKTLNDEEVNSVVQKALEILEKEFNARLKG

>PROKKA_01523 3-phosphoshikimate 1-carboxyvinyltransferase 1

MIELDINASDKSLSHRAVIFSLLAQKPCVVRNFLMGEDCLSSLEIAQN LGAKVENTAKNS
FKITPPTAIKEPNKILNCNNSGTSMRLYSGLLSAQKGLFVLSGDNSLNARPMKRIIEPLK
AFGAKILGREDNHFAPLVILGSSLKACHYESPIASAQVKS AFILSALQAQGVSVYKENEL
SRNHTEIMLKS LGADIQNQDGVLKISP LEKPLEAFDFTIANDPSSAFFFALACAITPKSR

LLKKNVLLNPTRIEAFEVLKKGASIEYAIQSKDLEMIGDIYVEHAPLKAISIEQNIASL
IDEIPALSVAMLFAGKSMVKNAKDLRSKESDRIKAVVSNFKALGIECEEFEEDGFYIEGL
EDISPLKQCFSRIPPLIKSFNDHRIAMSAVLTLALPLEIDNLECTNISFPQFKRLLNL
FKKGSFNGN

>PROKKA_01524 4-hydroxy-3-methylbut-2-enyl diphosphate reductase

MEIKMAKDYGFCFVGKRAIQIAEKNQNSLIFGSLIHNAKEINRLEKNFNVKIEEDPKKIP
KNKSVIIRTHGIPKQDLEYLKNKGVKITDATCPYVIKPPQIVESMSKEGYQIVLFGDINH
PEVKGVISYATNQALVINSLEELQEKKLQRKVALVSQTTKQTPKLLQIASYLVERCTEVR
IFNTICNATSYNQKAALDLSKEVDIMIVVGGKTSSNTKQLLSIAKQHCKDSYLVEDENEL
ELAWFKGKKLCGITAGASTPDWIIENVKQKISTI

>PROKKA_01525 30S ribosomal protein S1

MSKIADDQNFNDEEENFAKLFKKELEKEETLEKGTIKEGLIVSINENDGYAMVSVGGKTE
GRLALSEITDEKGQQLLYQKNDPIVVHVSEKGEHPSVSYKKAISQQKIQAIEELGENYEN
AIEGKIVGKNKGGYIVESQGVYFLSRSHSSLKNDANHIGKRIKACIIRVDKENHSINI
SRKRFFEVDNKRQLEVSKEELLEATEPVLGVVRQITPFGIFVEAKGIEGLVHYSEISHKGP
VNPEKYYKEGDEVYVKAIAYDAEKRRLSLSIKATIEDPWGEIQDKLKPGYAIKVVSNIIE
HYGVFVDIGNDIEGFLHVSEISWDKNVSHPSHYLSVGQEIDVKIIDIDPKNRRLRVSLKQ
LTNRPFDFVESKHQVGDVLEGGVATLTDFGAFNLGGVDGLLHNHDAFWDKDKCKDHYK
IGDVIKVKILKINKDKKISLSAKHLVTSPTTEFAQKHKTDSVIQGVVSIKDFGVFINA
DGIDVLIKNEIDLNLKKDEIKIGQEITCVVVAIEKSNNKVRASVHRLERKKEKEELQAFN
TSDDKMTLGDILKEKL

>PROKKA_01526 hypothetical protein

MRFKSVVAFISLAVALGLAYLFLSVKKEMPATSHAISQTNEGLSQTDTKSHDINLEENS
PDEISHNEKASHNQEVRRNALSQNLDAQEVINYPVIEHHFEIPFEEKKREYSKLIIDKDL
DYQWWCLKEILKKEQIDYAYDSTKNQPNLIYYLDKNKKERFLADLDYYKIRYHAVF

>PROKKA_01527 D-3-phosphoglycerate dehydrogenase

MYQVAICDPIHAKGIQILEAQKDIVLHDYSKCPKNELLGKLTLMDALITRSMTPITSDFL
KPLTHLKSIVRAGVGVNDIDLESCSQKGIVVMNIPTANTIAAVELTMAHLINAVRSFPCA
NDQIKHQRLWKREDWYGTLEKNKKLGIIGFGNIGSRVGIRAKAFEMEVLAYDPYIPSSKA
TDLGVIYTKNFEDILQCDMITIHTPKNKETINMIGAKEIERMKGKGVILINCARGGLYNED
ALYEAELETKKVRWLGDVFSKEPGIHNKLLDLPNVYATPHIGANTLESQEEISKQAAQGV
MESLRGSSHPHALNLPMAFDASAKAYLNLAQKLGYSFSSQIHKGVQCQIELSLCGEINQF
KDALVAFMLVGVLPVVGDKINYINAPFVAKERDIEIKVSLKESASPYKNMLSLTLNAAN
GSISVSGTVFEEDILKLTEIDGFHIDIEPKGKMLLFRNTDIPGVIGSVGNAFARHGINIA
DFRLGRNTQKEALALIIVDEEVSLEVLLEELKNIPACLSVHYVVI

>PROKKA_01528 3-octaprenyl-4-hydroxybenzoate carboxy-lyase

MRDFLKLKKHDELEIIDTPLEVDLEIAHLAYIEAKKPDGGKALLFTQPIRKEHDQIKTF
GMPVLMNAFGSFKRLDLLLKTPIESLQQRMQAFLHFSAPKNFTESLKVLDLWDLRHIFP
KKTARPKDLIIKQDKEVNLWDLPLVLTWEKDGGAFITMGQVYTQSLDHKKKNLGMRYLQV
YDKNHLGLHWQIHKDSQLFFHEYAKAKVKMPVSIAGGDLTYWCATAPLPYGIYELMLY
GFMRGKKARVMPCLSNSLSVPSDCDIVIEGFVDCEKLELEGPFGDHTGYTPIEPYPVLE
VKTISYKKDSIYLATVVGKPPLEDKYMGYLTERLFLPLLQMNAPNLIDYMPENGVFHNL
ILAKIHTRYNAHAKQVMHAFWVGQMSFVKHAIFVNEDAPNLRDTNAIIEYILENFSKEN
ALISQGVCDALDHASPEYAMGGKLGIDATSKSNTPHPTLLNDSALLALLQDKMPNIVLLK
QYYPHTRNPICIIISVEKKDKSVIELAKNLLGFEEYLRIVIFVEHTSNDLNNPYMLLRWIV
NNIDAKRDILTSKHCFFIDATNGKIMDKHFREWPTETNCSMEVIENLKKKGLLKDFETLN
QKFHLTHSFSTHKEDL

>PROKKA_01529 hypothetical protein

MTDYSQRIDTLIAKIEKARTAYSRRHHIVKIVAVSKNASLEAIQHYYNCSQRAFGENRVQD
LKIKMHSLEHLPLEWHMIGSLQENKINALLSLKPALLHSLDSLKLALKIEKRCEILGVNL
NALLQVNSAYEESKSGVMPEETLEIYSQISETCKRLKLKGLMCIGAHADDEKEIEKSFIT
TKKLFDRLKNASVLSMGMSCDFELAIACGANLLRIGSFLEKE

>PROKKA_01530 UDP-2,3-diacetylglucosamine hydrolase

MLETYVLKSGAVFISDAHFLPKSPHLIHTLKELLSAKPPQVFFMGDIFHVLVGYLPLDKE
QQKIIDLIALSEISQVFYFEGNHDFSMRFVFNFKVVVFERQNPVLFQHDNKRFLLAHG
DLFITKAYEFYITQLTSTWARFFLTFLNFVSFKTLYPFLKKCIYQKPIRLWELEPKELKA
FIEKRLKAYQNYIKDLNIDSIDGIIEGHFHLKSNAKIPLSAPIYRPLPSFYEEQSLFKVS
SSVLEPSQNKDA

>PROKKA_01531 Chemotaxis protein CheV
MAEKTANDLKLSEIELVDFRIYGMQEGVPYEGIYGINVAKVQEIIPMPTLF EYPTNLDYI
IGVFDLRSTIIPDLIDAKWIGIIPDKSKENEKIVIITEFN NVKMGLVHSARRIRRISWK
DVEPASFSASNSINKENITGTTTRIENDKTLILLDLESILDDLKLNEDAKNTK DTPKERFE
GEVLFLDDSKTARKTLKNHLSKLGFSITEAVDGEDGLNKLEMLFKKYGDDL RKHLKFIIS
DVEMPKMAIISYSSSKTLGSLIFL

>PROKKA_01532 hypothetical protein
MIFNSSICDNYS AERAKEMGAVAYLVKFDAEKFTEEISKILDKNA

>PROKKA_01533 Chemotaxis protein CheA
MDDLQEIMEDFLIEAFEMNEQLDQDLVELEHNPEDLDLLNRIFRVAHTIKGSSSFLNLNI
LTHLTHNMEDVLNRARKGEIKITPDIMDVVLRSIDLMKTLTIVTIRDTGSDTNNGKENEIE
EAVKQLQAITSQNLEGAKERTKEAPQKENKEENKEENKEEIKENKAKAPT AENTASDNPL
ADEPDLDYANMSAEVEAEIERLLNKRQEADKERRAQKKQEAKPKQEVAPTKEAPKTETP
KAPKTETPKAPKTETKAKAKADTEENKAPSIGVEQTVRVDVRRLDHLMNLIGELVLGKNR
LIRIYGDVEERYDGEKFLEELNQVVSSISAVTTDLQLAVMKTRMQPVGKVF NKFPRMVRD
LSRELGKSIELIIEGEETELDKSIVEEIGDPLIHIIRNSCDHGIEPLEERRRLNKPETGK
VQLSAYNEGHNHIVIKISDDGKGLDPVMLKEKAIEKGVISERDAEGMSDREAFNLIFKPGF
STAKVSVNSVSGRGVGM DVVKTNIEKLNGLIIEIDSEVGVGTTQKLKIPLTLAI IQALLVGV
QEEYYAIPLSSVLETVRISQDEIYTVDGKSVLRRLRDEVLSLVRLSDIFKVD AILESNSDV
YVVIIGLADQKIGVIVDYL TGQEEVVIKSLGYLKNTRGIAGATVRGDGKITLIVDVGAM
MDMAKSIKVNITTLMNESENTKSKNSPSDIYVLAIDDSSTDRAIIRKCLKPLGITLLEAT
NGLEGLEMLKNGDKIPDAILVDIEMPKMDGYTFASEVRKYNKFKNLPLIAVTSRVTKTDR
MRGVESGMTEYITKPYSGEYLT TVVKRSIKLEGDQS

>PROKKA_01534 Chemotaxis protein CheW
MSNQLKDLFERQKEANAGSKQEDNEEILQFIGFIIGDEEYAIPIILNILEIVKPIGYTRVP
ETPNYVLGVFNLRGNV FPLISLRLKFGLKAEKQNKDTRYLVVRHNDQIAGFFIDRLTEAI
RIKQTDIDPVPETLS DNNNLTYGIGKQNDRLVTILRVEEILKKDF

>PROKKA_01535 putative thiol peroxidase
MQKVTFKEETYQLEGKTLKVGDKAPDVKL VNGDLQEVNLLKQGVRFQVISSSLPSLTGSVC
LLQAKHFNEQAGKLPSVSFSVISMDLPFSQGCAGIEKDLRILSDFRYKAFGENYGV L
LGKGSLLQGLLARSVFVLD DKGVVVYKEIVQNILEEPNYEALLKVLE

>PROKKA_01536 Superoxide dismutase [Fe]
MFTLRELFPFAKDSMGDFLSPVAFDFHHGKHHQTYVNNLNNLIKSTDFEKNSLFDILT KSS
GGVFNNAAQIYNHDFYWDCLSPKATALSDELKGALEKDFGSLEKFKEDFIKSATTLFGSG
WNWAAYNLDTQKIEIIQTSNAQTPTVTDKKVPLLVDVWEHAYYIDHK NARPVYLEKFYGH
INWHFVSQCYEWAKKEGLGSVDYYINELVHKKA

>PROKKA_01537 tRNA (cmo5U34)-methyltransferase
MKDTLNFESLNKRFCFDEKVAHVFD DMLERSIPYYYEMLDLGAYFIAQNLKENVYPKPLP
KPLIYDLGCSTGNFFIALNRQIQQDIELV GIDNSMPMLKKAQEKLKDFNNARFECMDFLE
VEFKEASAFSLLFVLQFVRPMQREVLLKKIYNSLALNGVLLVGEKIMSEDRILDKQMIEL
YYLYKQNGQGYSHNEIAFKREAL ENVLVPYSLKENVALLESVGFKHVEAVFKWVNFTLLVA
RKT

>PROKKA_01538 Primosomal protein N'
MFYHLIAPLKNKTPPLTYFSKERHLKGALVNIPLRNKTLLGVVLEEVS KPSFECLELEKT
PYFLLSFQMELAI FIAQYYSANLSSVLSLFAPFKECDLVGLEKIEPVLNRLSQTQT NALK
ELQKHPASLLFGDTGSGKTEIYMHAIAQTLEQKKSALLLVPEIALTPQM QQRKRKFEN
LGLWHSKLSQNQKKQFLEKLYSQEIKLVVGTRSAFLPLKELGLIIVDEEHDFS YKSHQS
PMYNARDLCLYL SHKFPIQVILGSATPSLSSYKRFDKALVRLKG RYTPTQKNIILEKTE
RFITPKLLEALQQVIDKNEQAIIFVPTRANFKTLLCQNCYKSVQCPFC SVNMSLHLKTNK
LMCHYCHFSSPIPKICSACQSEVLVGKRIGTMQVLNELEGLLKGAKIAILD KDHTSTPKK
LHNILNDFNAQKTNILIGTQMISKGHDYAKVSLAVVLGIDNIIKSNSYRALEEGVSLLYQ
IAGRSARQISGQVFIQSTETD LLENFLEDYEDFLQYELQERCELYPPFSRLCLLEFKHKN
EEKAQQLSLKASQTL SLCLEKGVTLS SFKAPIEKIASSYRYLILLRSKNPLSLIKSVHAF
LKTAPNIPCSVNMDPVDIF

>PROKKA_01539 hypothetical protein
MSLENDLSLEITYLGKRYKISLNNTFSDEM KRTLKERFHNQELNALELLKDYLHESCQNEY
LHNELQKLLEKISSCSIT

>PROKKA_01540 hypothetical protein

MQSLSLNQLGTKIDELIEKIQKQEEELNALRQANTTLNAQNEEKDIQIAILYDELSAKD
KGIQGLYDKISDLLS
>PROKKA_01541 hypothetical protein
MQKSILKMTLLLVFLFLRNAVGLEDEKADPKSVQNTPKNLPPIQLRLNQVHEELIEMLEN
MGKGTQYEFPKIKEILEQSEEEWLKVAHEECVALVMLISPKASIENSPIYKNCYEAYVKQ
RIHDLYDFYIEGKKVKRIKKAHKQEAAIKQSQPLTKESPKSENKSLTKPSLKDAGIPK
GYLQIGAFLNAPSKDFLQTLKTFPYQIKKKDSLTHYLIGPYKTKEEALKQLENATKSFK
NRPVLVEK
>PROKKA_01542 hypothetical protein
MPYSFKKRFLIVYTLSTLLLVGVLLALFFFYAKNNLLENTQIRMQYTADAIKSLLELNN
ASSLEPLKILEERFKNTPFVLLDADNRVKFSNIGAFVASFKNDALIKTPYFALKKQG FYL
TDSAPTNRGLGVSKIIIAEEEIQKNFIPLYKMIGYVFLGTSLFVALIAGWLYKIQ
>PROKKA_01543 Protease HtpX
MLDIWIDMIICIFYLLFFTTPYIVGDILQLKFIRQKLCEKPVLLPQKDYEEAGNYAIRKM
QLSIISQILDGIIFAGWVFFGLTHLEDLTHYLNLSSETLGYLVFALLFLAIQSVLSLPISY
YTTMHLDKFEGFSKVSLSLFFKDFKGLSLTLGVGLLLIYTLIMIIEHVEHWEISSFFVV
FVMILANLFYPKIAQLFNQFTPLNNRDLESQIESMMDKVGFKSEGIFVMDASKRDGRLN
AYFGGLGKNKRVVLFDTLISKVGTEGLLAILGHELGHFKNKDLLKSLGIMGGLLALVFAL
IAHLPPIVFEGFNVSQTPASLIAILLFLPVFSFYTMPLIGFFSRKNEYNADKFGASLSS
KETLAKALVSIVSENKAFPYSHPFYVFLHFTHPPLLERLKALDYEIE
>PROKKA_01544 Release factor glutamine methyltransferase
MTLSQALNKAKKELSPKGFRGGLESEILLGFVLQKERVFLHTHEHLELTHKEEMCFFELV
EKRLNDCPIEYLLLESCDFYGRSFFVNEHVLI PRPETEILVQKALDIISQYHLKEIGEIGI
GSGCVSVSLALENPNSI HASDISPNALEVALKNIERFNLKERVFLKKTRLWDRMPTIEM
LVSNPPIIARNYPLEKSVLKEPHEALFGGVKGDEILKEIIFLAAKLKIPFLVCEMGYDQL
KSLKECLEFCGYDAEFYKDLSGFDRGFIGVLKSFLK
>PROKKA_01545 NAD(P)-specific glutamate dehydrogenase
MYIEKILQSLQKKYPYQKEFHQAVYEAITS LKPLLDSDKSYEKHAILERLIEPEREIFFR
VCWLDDNHQIQVNRGCRVEFN SAIGPYKGGRLRFHPSVNESVIKFLGFEQVLKNSLTTLAM
GGAKGGSDFDPKGKSEHEIMRFCQAFMNELYRHIGATTDPAGDIGVGEREIGYLFQYK
KLVNRFEGVLTGKGLTYGGS LCRKEATGYGCVYFAEEMLQERNSSLEGKVCSVSGSGNVA
IYTIEKLLQIGAKPVTASDSNGMIYDKDGIDLELLKEIKEIRRGRIKEYALEKKSAYTP
IENYPKGGNAIWHVPCFAAFPSATENELSVLDAKTLLSNGCKCVAEGANMPSSNEAVELF
LQAKISYGLGKAANAGGVSVSGLEMAQNASMHWPWSFEVVD AKLHHIMKEIYKNVSQTAKE
FKDPTNFVLGANIAGFRKVASAMIAQGV
>PROKKA_01546 hypothetical protein
MLYENPLNTLDGKAYFYQNL SFFKKILDFFKTILENDTIYHDNPFIFYRDLHEPLATIDDL
RVNYDDL RVNYERLLQNASPLLELSQNTTFKIYRKIYHKS LPLLHAIRRWVKKLGL
>PROKKA_01547 Glycosyltransferase family 10 (fucosyltransferase)
MFQPLLD AFIESASIEKMVSKSPPLKIAVANWWGDEEVKEFKKSF LYFILSQRYTITLHQ
NPNKPSDLVFSNPLGSARKILSYQNTKR VFYTGENEVPNLNLFDYAIGFDELDFRDYLR
MPLYYDRLHHKAESVNDTTAPYKIKDNSLYTLKKPSHHFKENHPNLCAVVNNE SDPLKRG
FASFVASNP NAPMRNAFYDALNSIEPVTGGGSVKNTLGYKVKNKNEFLSQYKFNL CFENS
QGYGYVTEKILDSYFSHTIPIYWGSPSVAKDFNP KSFVNVD FKNFDEAIDYVRYLPRTQ
TLI
>PROKKA_01548 Cytochrome c biogenesis protein CcsA
MKNLKSLLSFL LASFWVAIPLIALYALACAIATFIENDYGTSASKAIVYNTPW FNFLHAY
LLVVLIGTFINSKALERKRYASLFFHSSLILILGAITRFFGVEGLMHVRENSAQSSFE
SADTYLNITLNDTTKLSLKTPLTFYHSKR LRPIHATLDHKPLILEPLEIYKQNAIKKDDA
TILVLKATYNGVSRKFNLKNNRNEGIEESEMFKDDKLSLSFGSAYIELPFQIKLR FEL
ERYAGSMSPSSYASEVEVLKLDNTLIKPYRIFMNHVLDYEGYRFFQSSYDTDEKGTILSV
NKDPGKIPTYLGYAMLILGALWLLLDKNGRFLKLSRFLKSQMASFLLALILISPFTSSF
ANEAPIDMHGGS AKIERQSVENSANKEDSKSAILERLKR LREYSKDHLKAFQRLQVQDF
DGRIKPLDTISIEYIHKILRKDDFQGLNAMQVLLGIMFFPNDWRSVKMIYTSNKALR KLI
GTPLDESRIAFRDVFD SGRYKLNVLVEEVNQKSPNARNELDKDVLKVDERINLVYTLFSA
QFLRIFPSDKTTAWLSPIEAINSPNKEISSVATAFLKNIFSGFDDALKTNQWDKVEKTLK
DLSVYQQEHAKNLYLPSSKVDSEIFLNHTNFNRLTLPYIFLGLLLFIVVISSLVKNTIP
NIWLTKILYMAILLCAFAHSIGLILRWYVSGHSPWSNAYESMLYIAWASVIAGFILRSKL

ALSASSFLAGIALFVAHLGFMDPQIGHLVFVLKSYWLNHVSVITASYGFLGLCFVLGIL
 SLVLFILRKQGRFNLDKILSISAINEMSMILGLFMLTAGNFLGGVWANESWGRYWGWD
 KETWALISICVYALILHLRFLGSHNWPFIASSSVLGFYSVLMTYFVGVNYLSGLHSYAA
 GDPLPIPTFLYFLVAIPFALVILAYFKRHLSLPKLI
 >PROKKA_01549 hypothetical protein
 MHKKAIIVVRYLVFKPIKNKGYFMFSLSYVSKKFLSVLLLISLFLSACKSNNKDKLDENL
 LSSGSQSSKELNDERDNIDKKSAGLEDVFLDNKSI SPNNKYMLLVFGRNGCSYCERFKK
 DLKNVKELRDYIKEHFSAYYVNISYSKEHNFKVGDKNNEKEIKMSTEELEQIYAVQSTPT
 IVLSDKTGKTIYELPGYMPSTQFLAVLEFIGDGKYQDTKNDEDLTKKLKAYIKYKTNLSK
 SKSN
 >PROKKA_01550 Ferrochelatase
 MNLINEKLNNLEDNATKSPKEAVLLNMGGPNSLYEVGVFLKNMFDDPFILTIKNNFMRK
 MVGKMIVNSRIEKSCKIYEKLGKSPITPITFALTERLNKLDPSRFYTYAMRYTPPYASM
 VLQDLALKEVESLVFFSMYPQYSSTTTLSFNDAFSALKSLETFRPKVRVIERFYASTKL
 NEIILNTILSTLNNRKSQDFVLIFSVHGLPKSIVDAGDTYQCECEHHVSLKELMQQKNI
 PFKEVLLSYQSKLGPWKLEPSTEELIEKHKRSHIIYPLAFTIDNSETLYELDMQYRLM
 AERLAVKEYLVCPCLNDSIEFAQFI IERVKNLKE
 >PROKKA_01551 hypothetical protein
 MSLQALNTQDSSVMAQSLLVHAFFAALLALAFMINLYTLFKEKNFIQLNRKIYLVMPAIY
 ILLSIALLSGIFIWAMQQFEFSFSAVAMLLGLLLMLIAEIKRHKSVKFAITKKERMEAYI
 KKAKILYFLETILIVVLMGL
 >PROKKA_01552 Ribosomal RNA small subunit methyltransferase E
 MRFVYHPLAKEPVLKIEGESYTHLYRSRIKSASRLDLRNKDSLLYTYEHAEITKKHAL
 LRLVGGRPLEVMANKKTHLILSVIEIKSIEKILPFLNQLGVSKLSLFYADFSQRNEKIDS
 AKLERFQKILINSCEQCGRSVLMELEAFPNTKEMLNAYPKASVLDFKGETLPASADFEKG
 VIIGPEGGFSEQEREQFKEREIYRIPLDMVLKSESACVFVASIAQV
 >PROKKA_01553 Helicobacter outer membrane protein
 MLKLVSKTICLSLIGLFNPLEAFQKHQKDGFIEAGFETGLLEGQTITQNTQNTQ
 KIYENPLTHPQTKEQPKQNKSDTATPQSAYGKYIIPQSTILEKATELFTTDNIENGLTF
 YSQNPVYANMVNGSVTIQNFLPYNLNNVELSFKDAQQKVVNLGVETIPKQSQIILPASL
 FNDSEFEQADSFNYQQLQATATQFSDANTQSLFEKLSKITTNVTMSYENADTNNFKGNCH
 DCVSDFTPQTAEELTNMLMDMIAVFDSKSWEAVLNAPFQFSNSPSECGSDFPKCVNPFN
 NGRVAPIYEKYVLTPQSVIDAFRRTINLEVNILKSGFVGLGYELDDNDGNLGIASALNP
 EKLFGKTLNKKVDIVELRDIIEHFSHTKGYTHNGNMTYQVRVLCQENGGAIQECEGGKEEL
 VNGKEELKFTNGKEVKDQDGYTYDVCSRFGGKNQPAFPSPNSIYTNCAPAGLIGVT
 TAVWQQLINQNALPINFANLNSQANYLDASLNAKAFVGSVFNAFNQNF LTSSAQQSFRSP
 ILGANVKIGYQHYFNDYIGLAYYGIIKYNYAQANDEKIQQLSYGGGMDVLFDFITTYTNK
 KQNNPTKKVFASSFGVFGGLRGLYNSYYVFNQVKGSGNLDIVTGFNRYRYKHSKYSIGISV
 PLMQNDIKIASNNGIYANSVVLNEGGS HFKVFFNYGWIF
 >PROKKA_01554 Deoxycytidine triphosphate deaminase
 MGLKADSWIKKMSLEHGMISPFCEKQVGKNVISYGLSSYGYDIRVGSEFMLFDKNALID
 PKNFDPNNATKIDASKEGYFILPANAFALAHTIEYFKMPKDTLAICLGKSTYARCGIIVN
 VTPFEFEFEGYITIEISNTTNLPAKVYANEGIAQVVFLQGDCEMQSYKDRGGKYQGQVG
 ITLPKILK
 >PROKKA_01555 Biotin carboxyl carrier protein of acetyl-CoA carboxylase
 MNLSEIEELIKEFKASDLGHLKLKHEHFELVLDKESAYAKKNALNPAHSQASIQAPIMVE
 ASMPSTQAPVPMVCTPIVDKKEDFVLSPMVGTFYHAPSPGAEPYVKAGDTLKKGQIVGIV
 EAMKIMNEIEVEYPCKVVSVEVGDAQPVEYGTCLIKVEKL
 >PROKKA_01556 Biotin carboxylase
 MNKVNKENKKVEKKELSRILIANRGEIALRAIQTIQEMGKESIAIYSIADKDAHYLNTAS
 AKVCIGGAKSSESYLNIPAIISAAELFEADAI FPGYGFLSENQNFVEICSHHSLEFIGPS
 AKVMALMSDKSAKSVMKEAGMPVIEGSEGLLSYQEAEEIADKIGYPVIIKAAAGGGGR
 GMRVVEDKSKLKNLYLAAETEALSAFGDGSVYLEKFINKPKHIEVQILADKHGNVIHVGE
 RDCSVQRRQQLIEETPAVVLEEGVRKRLLETAIAKAYIGYVGAGTFEFLDSNMKDFY
 FMEMNTRLQVEHTISEMVSGNLNIEWMIKIAQGEELPKQESFSLKGHAIECRITAEDPKK
 FYPSPGKITEWIAPGGVNVRLDSHAHANYVVP THYDSMIGKLIVWGENRERAI AKMKRAL
 KEFKVEGIKTTIPFHLEMLENADFRQAKIHTKYLEENF
 >PROKKA_01557 hypothetical protein

MQTLFKEITPKRYVNGNEMKENS SNVLDQYFTKPSVALKCFQKACEVIKKYENLDEFIFL
EPSAGDGVFYDLFPKDRRIGIDIEPKRDGFIQCDFLNYKLPTHQKVICLGNPPFGHRGVM
ALEFINHARNCDFVCFILPMFFESQGKGSIKYRVKGLNLLYSERLEKNAFIDFKNKEVDV
HCVFQIWSKKYQNKSEFSWYKNRHKEPFGEYIKVFTVSLAKNRECGKEWIFNQKASYI
SSTFYKSTQIVENFEEVKYQSGIAVVFTSTDKVLNAKLKLFKEIDWTKYASLATNSCYH
LGKSHIFQALHDHLDLRLDN

>PROKKA_01558 hypothetical protein

MDLEQTFLKIIIEKKHKELNLGQDYNAIFSKIRDFEANAIGQIGEEFLKSVLNAIDEVIND
GIIHDEYDIMAKSGVSFEVKTARKGRNTNTFQFNGINPRYNYDFLICLVGEDQLLYRIF
KKDEIHYIHKKRKYFMKQNEFKKQLVPMNPDNQVNYKLTNLKELKEITNLIKELERVLG
LD

>PROKKA_01559 hypothetical protein

MSKISNNYNPSLMMRDYHTQRVSPHARKEENKEIQNLSENDEKIKLAKQAKQDNLAIGDL
ESRLKSLKGMDDAKELVGISKAYAHNNEKDRSDFEHFKSRLDKAIDSFNQKSGNDSLKL
PNNIDIDDTKALEKFSKSLESEKENIQNSLHQWKKQLAETNHLNKEYNTLDKTRLNAQKF
QDVHDTSKITPSRLQDLLA

>PROKKA_01560 UDP-4-amino-4,6-dideoxy-N-acetyl-beta-L-altrosamine transaminase

MKEFAYSEPCLDEEDKKAVLEILNSKQLTQGKRSLLFEEALCEFLGVKHALVFNSATSAL
LTLRYNFSDFDADCNEIITTPISFVATANMLLESgyTPVFAGIKNDGNIDELALEKLINE
KTKAIVSVDYAGKSVEIKSIQELCKKHSLSFLSDSSHALGSEYQNKKVGGFALASVFSFH
AIKPITTAEGGAVVTNDSLEYEKMKLFRSHGMLKKDFFEGEVKSIGHNFRLNEIQSALGL
SQLKKAPFLMQKREEIALVYDRIFKDNPFYFTPLHPLLKYKSSNHLYPILMRQKFFTCKKL
ILENLHKGRIQAQVHYKPIYQYQLYQQLFNTAPLKSADQFYRAEISLPCHANLNLESAQT
IAHGVLKTFEDLKIE

>PROKKA_01562 Ribonucleoside-diphosphate reductase subunit beta

MEVSRKKIYNPDSTESVNERKIFGGNPTSMFDLNKIKYQWADHLWKTMLANTWFAEEVSM
NDDKRDYLLKLSAEEKIGYDRALAQILFMDSLQTNLIDNINPFITSPEINLCLVRQAYEE
ALHSHAYAVMVESISANTEEIYDMWRNDMQLKSKNDYIAQVYMELAKNPTEENILKALFA
NQILEGIYFYSGFSYFYTLARSGKMLGSAQMIRFIQRDEVTHLILFQNMINALRNERADL
FTPKLINIEVIEMFKKAVEIEALWGDYITQGKILGLTSSLIEQYIQFLADSRLSKVGIKAV
YGVQHPIKWVESFSSFNQRSNFFEARVSNYAKGSVSFDDF

>PROKKA_01563 Protein-L-isoaspartate O-methyltransferase

MNSIKNHLMCEEIHKRFNLHPKVREAMESIEREVFVPAPFKHFAYTLNALSMQAQQYISS
PLTVAKMTQYLEIDHVDVLEIGCGSGYQAAVLSQIFRRVFSIERIESLYIEARLRKLTL
GLDNVHVKFADGNKGWEQYAPYDRILFSACAKNIPQALIDQLEEGGILVAPIQENNEQVI
KRFVKQNNALRVQKVLEKCSFVPVVDGVQ

>PROKKA_01564 putative permease YjgP/YjgQ family protein

MAVSQVFFSFFLVLFSSIVLLISIASVTLVIKVSFLDLVQLFLYSLPGTIFFILPITF
FAACALGLSRLSYDHELLVFFSLGVSPKMTKAFVPLSLLVSAILLVFSLLILIPTSKSAY
YGFLRQKKDKIDINIRAGEFGQKLGDWLVYVDKAENNSYDNLVLFNSKSLSQESFILAQK
GNINNQNGVFELNLYKGHAYFTQGDKMRKVDFEELHLRNKLKSFNSNDAAYLQGTDYLG
WKKAFGKNANKNQKRRFSQAILVSLFPLASVFLIPLFGIANPRFKTNWSYFYVLGAVGVY
FLMVHVISTDLFLMTFFFPIWAFGSYLLFRKFILKRY

>PROKKA_01565 tRNA pseudouridine synthase A

MRCFKVTIAYDGAYFLGYAKQPNKLGVDKIESALNMLGIKSVVVAAGRTDKGVHANNQV
LSFHAPKHWNAAKLFYCLAPKLAPHIVLKKLEEKNFHARFDAQKRAYRYLLTKNLKTPFL
APYIACGDYGLDALNTALKQFTGKHDFSLEFKKEGGATTNPNRIIFNAFAYKTSIMGHEC
VVFKEIGDAFLRSSVRLIIQACVQYSLEKITLAEIEMQIHNLKATIRTPIMANGLYLHRV
YY

>PROKKA_01566 UDP-glucose 4-epimerase

MALLFTGACGYIGSHTARAFLEKTENIIIVDDLSTGFLEHLKALEHYYPNRVVFVQANL
NETYKLD AFLNKQQLKDPIEAILHFGAKISVEESTHLPLEYYTNNTLNTLELVKLCLKHA
IKRFIFSSSTAVVYGESSSSLNEESPLNPINPYGASKMMSERILLDTSKIADFKCVILRYF
NVAGACMHNDYTTPYTLGQRTL NATHLIKIACECAVGKRKKMGIFGTNYPTRDGTICIRDY
IHVDDL ANAHLASYQTLEKNKSEIYNVGYNQGHSVKEVIEKVKEISNNDFLVEILDKRL
GDPASLIANNAKILQNTPFKPLYNNLDTIKSALDWEHLLRFQ

>PROKKA_01569 hypothetical protein

MKKFLYTLLALLLIGLLTTYLILFTEWGNKIIASYIEKKINPNERYLNVKTFKLRFNSLD

FEAQANDDSTLILKGDFSLLKQSVDLNYHIDIKNLYSFKEWIPYPLRGAIVTSGNIKGHR
KALVVQGVSNVAQSHTTYNALLDDFKLSHLSLNAQDANLEDLLYLINRPAYANAKVSLQA
DFNSLKPLEGHLILTANNALINNALINQIFHLNLKDTLVFSLSHSSDFKGNKAISDTTTLT
SPLVNFTALKSEYLFVSLKLNAPYTLEIPNLAKLQNIINHPLKGSITLKGAEQSPKLLK
VSGHSNLLDGALDFTLLNKDLKARFSNISTLKALDLFHYPKFFQSIADANLDYDLSTKQG
VLKARLKNNARFLKNAFSDFLYSISKFDITKEIYNDANLISQINQQRLSDLSLKSPKTQL
KIHNGLLDLNNTKQMDMLMDAEILKFVFKMKLQGNMHQPKFSLILNEKAIQQNLQQGLKEI
LKNDTIKKGLDHLLKDDKLKEKLEKGLKGLF
>PROKKA_01570 NADP-dependent 3-hydroxy acid dehydrogenase YdfG
MAHILVSGATSGFLEIAKAFLOKNHVFGTGRRKENLQELQLAYPKHFIPLCFDLKNKL
ETKRAIETIFSMTRVDALINNTGLALGLNKAYECELDDEIMIDTNIKGLLYLTRILP
SMIEHNQGTIINLGSIAGTIYAYPGGNVYGASKAFVKQFSLNLRADLAGTNIRVSNVEPGL
CGETEFMVRFKGDRIKAQSVYENTIYILKPQDIAHIVLWIYEQPLHVNINRIEIMPISQT
FAPLPCHKSP
>PROKKA_01571 Putative beta-lactamase HcpC precursor
MLENVKKSIFRVLCGLGALCLGGLMAEQDPKELVGLGAKSYKERDFSQAKKYFEKACDLKE
NSGCFNLGLVLYYQGGQVEKNLKKAAQFYAKACDLNYSNGCHLLGNLYYSGQGVSNQNTNKA
LQYYSKACDLKYAEGCASLGGIYHDGKVVTRDFKKAWEYFTKACDLNDGDGCTILGSLYD
AGRGTPKDLKKALASYDKACDLKDSPGCFNAGNMYHHGDGVVKNFKEALARYSKACELEN
GGGCFNLGAMQYNGEGVTRNEKQAIENFKKGCKLGAKGACDILKQLKIKV
>PROKKA_01572 KHG/KDPG aldolase
MQDKIIEILQISPIVPVVVIENIKDAVPLAQSLIEGGIPIIEVTLRSSCALEAIELIAKN
VPKMHVAGAGTILNPTQLEQAQNRGAEFILISPLTIKLEHAKKKDMPLIPGVSSSSEVMQ
ALELGYNALKFFPAEYCGGVKLLNAFNGPFGVKFCPTGGISADNMRSYLNLENVLCVGG
SWLTPKDLIQNKEWDKITEICKRALA
>PROKKA_01573 Phosphogluconate dehydratase
MPKHSLEQIKEKITERSKKTRELYLENIIFNPKNQPKIESLGCANIAHVTAASMPEHLKMPL
GSHKRKHFAIITAYNDMLSAHQPFKNYPDWIKKELQEHNAYASVASGVPAMCDGITQGYE
GMELSLFSRDVIALSTAVGLSHNVFDGAFFLGVCDKIVPGLLIGALSFGNLASVFPVSPG
MVSGIENYKKAKARQDFAMGKINREELLKVMQSYHDVGTCTFYGTANSNQMMEFMGLH
VANSSFINPNNPLRKVLVEESAKRLASGKILPLAKLIDEKSILNALIGLMATGGSTNHTL
HLIAIARSCGVILNWDDFDVSNLIPLAKVYPNGSADVNAFEACGGGLAFVIKELLKEGL
LFEDTHTIMDTETQKGMQNYTKTPFLENDQLVYKDAVNHSLNTDILRPVSEPFAANGGLK
ILKGNLGRAVIKISAIKDEHRKVKARIVFKTQSEFLERFKNKELEKDFVAVLPFQGPKS
NGMPELHKLTTNLGALQDMGYKVALVTDGRMSGASGKVPSAIHLSPEGALNGAIIKIKDG
DWIELDAPNNALNVLEKDFENRGINPLFLETLENLEKPTFGLGRELFTSLRLNANTAEFG
GMSFGIKV
>PROKKA_01574 Glucose-6-phosphate 1-dehydrogenase
MLDFDLVLFGLATGDLAMRKLFLVSLYEIYTHYGFKKDSKIIASGRKELSNEEFLLCEKT
QLHSREKGEEFLAHISYLRVRLDNPKDFEELSKIATKDKPLIFYFSISPSFFATTAQNLA
KNALNHANTRILILEKPLGHDLKTCKEIFQSIASAFFKEEQIFRIDHYLGKKGVQNIELRL
NNPILNILDWDQISAVEICVYETLGEVEERGEFYDKIGALRDMVQNHLQVLSLIATDLPND
LKDLRKEKIKVLKTLQPPKDFKKQVIRAQYQGYRDENKVHKESTETTFVAIKAFDTPKF
KGVPFYLLKHAKKMPRNQASVKIHFNAVNTLEFFLSQDKITLTLKDHQNPLILETHNKQEF
LQPYAKLLYDAIQNNHNNFAHQLELEASWVFIDTLIEGFMNNATPLYSYESHNLNELEFL
KPLYQ
>PROKKA_01575 6-phosphogluconolactonase
MGYQLFEFESLKDCHKALIERFKEFFNAALKKHHQVSIAFSGGRSPIGLLQKLSVLDLKW
HECLISLVDERIVDTSHKDSNTKLLHDYLLQNNALKASFTPLLPKKISNDTNALFNANQ
HFKQPHLAILGMGTGHTASLFPETSAFLNEEKENIVLTKPTNAPYERLSMSVNALENCE
RLFLSISGVEKRGVLEKALKENAPYSLPIARILHSQKVTTEVFYAKN
>PROKKA_01576 Glucokinase
MPKTETYPRLADIGGTNARFGLEVAPRQIECIEVLRCEDFESLSDAVRFYLSKCKESLK
LHPIYGSFAVATPIMGDFVQMTNNHWTFSIETTRQCLDLKLLVINDFVAQAYAISAMQE
NDLAQIGGIKCEINAPKAILGPGTGLGVSTLIQNSDGLSKVLPGEAGHVSFAPFDDLEIL
VWQYARSKFNHVSARFSLSGSLVLIYEALSKRKGLEKMAKLSKAELTPQIIISERALNGD
YPICRLTLDTFCSMLGTLAADVALTLGARGGVYLCGGIIPRFIDYFKTSPFRARFETKGR
MGAFLASIPVHVVLKKTGGLDGAGIALENYLLHDKI

>PROKKA_01577 putative formaldehyde dehydrogenase AdhA
MRVPSKGFAIFSKDGHFKPHDFSRHAVGPKDVLIDILYAGICHSDVHSAYSEWKEGIYPM
VPGHEIAGVIKEVGKEVKKFKVGDVVGVCVNSCKACKPCKEHQEQQCTKVVFITYDCLD
SFHGNPHMGYSNNIVVDENYVISVDKNAPLEKVAPLLCAGITTYSPKFSKVTKGTVK
GVAGFGGLGSMVAVYAVAMGAEVSVFARNEHKKQDALSMGVKHFYTDPKQCKEELDFIIS
TIPTHYDLKDYLKLLTYSGELALVGLPPVEVAPTLSVFDIFIHLGNRKVYVGLIGGIKETQ
EMMDFSIAKHNIYPEIDLILGKDIDTAYHNLTHGKAKFRYVIDMKKSFD
>PROKKA_01578 Glycosyl transferase family 8
MTSASSHSFKEQDFHIPIAFADFKNYLIPAGACIYSLLESIKANKKIRYTLHALVVGLN
EEDRAKLNQITEPFKEFAVLEVKDIEPFLDAIPNPFDEDFTKRFSKMVLVKYFLADLFPK
YSKMVWSDVDVIFCNEFSADFLNIKENDENYFYGVLEVEKHHMMEGFLFCNLDYQRKKNF
TLRMHDLKLGNEAKGELDTKWCWPNMKALGIEYCVFPYYTITKDFSNAYLNNENYKKTIL
EARENPTIIHYDAWWGAVKPDYPPGLKADLWLNALAKTPFMSDWSYLTGGGEIGGEKW
HHYHSIAAYHYFFPLWKAEEQIAHDAFKTFLKHIFLHIHEIPQNARRRLFKYCISIPKLS
FISKTLKILKLHALVKKILIQKLKLLKKS
>PROKKA_01579 hypothetical protein
MKAASFVYDLKFMSFNFNLVAPANNPYWNSLTQMGRMPQIGVQLDKRQALMFGAWFIQ
NLHTHYSYFYPYSWGVTMYQYIGKNLRFGLIVPRSYQIGHYPLSAFKKLFWFIDPTFRG
GAFQFKPSYDPNRWWNGWFEVVDWYGGRNWNNQPKKKNYDFDQFLYFVSSEFQFLKGYL
GLGGQLVVFHNASHHSMGDNYPYGGNSYLKPGDVTQWPNGYPYFSQKNNPQGGEIGKYS
NPTILDRVYYHAYLKADFKNLMPYMDNIFMTFTGTQSSQTHYCVRYASECKNARFYNSFGG
EFYAQAQYKGFIFNRYFNSKPKMHFYATYQGSlyTGLPWYRAPNFDmIGLYLYKNKW
LSVRADAFSfVGGGDGYHLYGKGGKWFVYQQFLTLTIDTRELIDFVKSkipK
>PROKKA_01580 Helicobacter outer membrane protein
MDKTTVKILMGMALLSSLQAAEAELDEKSKPKFADRNTFYLGVGYSAINTSFSTESV
DKSYFMTGNGFGVVLGGKFVAKTQAVEHVGFYGLFYDQTFSSHKSISTYGLELSGLWD
AFNSPKMFLGLEFGLGIAGATYMPGEAMHGIIAQNLGKENSFLQLLVKVGFRFGFFHNEI
TFGLKFPFIPNKRTEIIDGLSATTLWHRLPVAYFNYIYNF
>PROKKA_01581 Pyruvate synthase subunit PorC
MFQIRWHARAGQGAGITGAKGLADVISKTGKEVQAFASYGSARKGAAMMAYNRVDDEPILN
HERFMQPDYVLVIDPGLVFIEINIFANEKEDTTYIITSYLNKEELFEKKPELKTRKVFLVD
CLKISMETLKRPIPNTPMLGALMKVSGMLEIGAFKEAFKKVLGKKLTQEVIDANMLAIQR
AYEEVQ
>PROKKA_01582 Pyruvate synthase subunit PorD
MKDWNFEFEMGAVLFPFEKNAQSEMEKHNDERHYTEQSYFTTSVAHWRVAKPVHNNNICIN
CFNCWVYCPDAAILSREGKLKGVDSHCKGCGVCVDVCPNPKSLWMFEEQIEPATALTQ
WPQKQEKKKS
>PROKKA_01583 Pyruvate synthase subunit PorA
MARSIELQEIEVWDGNTASSNALRQAQIDVIAAYPITPSTPIVQNYGSFKDNGYVDGEFV
LVESEHAAMSACVGAAAAGGRVSTATSSQGLALMVEVLYQASGMRLPIVLNLVNRALAAP
LNIHGDHSDMYLSRDSGWISLCTCNPQEAYDFTLMAFRIAHEHQKVRVPTIVNQDGFCLSH
TVQNVRLPSDAVAYQFVGEYQTKHSLLDfDKPVSyGAQAEEEWHEHKAQLHHAIMSASS
VIEEVFNDFAKLTGRQYHLTKTFQLEDAEIAIFALGTTYESAIVAAKEMRKKGIKAGVAT
IHSLRPFPYERLGQDLKNLKAAILDKSSPAGAMGAMFNEVTSAVYQTQGTkHPVVSNYI
YGLGERDMTIAHLCEIFEEINEDALKGTLTHPTQQFVGLRGPKMSFF
>PROKKA_01584 Pyruvate synthase subunit PorB
MIKEVKTlKGFSQSAEKfQGSHLLCPGCGHGIIvREVLNAVDGPiVLGNSTGCLEVCSAV
YPHTSWDVPWIIHIGFENGSTAISGVEAMYKALANKGRYQGQKPKFVAFGGDGASYDIGLQ
FISGCMERGHDMTYICLDNENYANTGGQSRSGSTPLGASTSTTPSGSVSFGKKEKKKDIVN
IMASHGVPPYAQLSPNKWKDMNKKIKTALDTEGPCFINALSPCTTEWKfESNKTIELADM
AVDSLmFPLFEIfNGMELKITYRPRNIIPVRDYLGAQKRfKHLfKKENEHIIEELQKDVN
ERWEYLQRREEAKV
>PROKKA_01585 Adenylosuccinate lyase
MLERYANEEMKALWNEQTKFETYLEVEKAVVRawnKLGQIQDSDCEKICAKAAFNLERIK
EIEKTTKHDLIAFTTCVAESLGEESRFFHYGITSSDCIDTAMALLMTKSLKLIQKGVKNL
YETLKNRALEHKDTLMVGRSHGVfGEPITfGLVLALFADEIKRHLKALDLTMEFISVGAI
SGAMGNFAHAPLELEELACEFLGLKTRQYQ
>PROKKA_01586 Adenylosuccinate lyase

MPHKRNPILSENITGLCRVIRSFATPMLENVALWHERDMSHSSVERFALPDLFITSDFML
SRLNSVIENLVVYPKNMLKNLALSGGLVFSQRVLLELPKKGLNREESYSIVQENAMKIWE
VLQQGAFKNADENLFLNALLNDERLKKYLSEDEIKACFDYNNYYTKNVGAIKRVFE
>PROKKA_01587 Helicobacter outer membrane protein
MKRALCLVLGLFCALNAKGFKDVLTKGDYTFNKKVVSPIKRYADQSAFYLGGLGYQLGSI
QHNSSNLNLSQQFNKSQIVFSDGLSPVFKNSYVSNGLGVQAGYKWVGKHEETKWFGFRWG
LFYDLSASLYGQKESQSVIISTYGTYMDLLLNAYNNGDKFFAGFNLGIAFAGVYDRLSDAL
LYQALLQNTFGGKVDPNGFQFLVDLGVRLGSKRNQFGFGIKIPTYYFNHYYSMNNISNNS
GDVLKVLRFLEYGINSLLYQVDFRRNYSVYFNNTYSF
>PROKKA_01588 UvrABC system protein B
MPLFDLKSPYPPAGDQPQAIEALTKSLKNNNHQTLVGVTGSGKTYTMANIIAQTNKPTL
IMSHNKTLCQAQLYSEFKAFFPHNRVEYFISHFDYYQPESYIPRRDLFIEKDSSINDDLER
LRLSATTSLGYYDDVIVIASVSANYGLGNPEEYLVMEKIKVGEKRAYKSFLKLKLVEMGY
SRNEVVFDRGSFRATGECVDIFPAYNDAEFIRIEFFGDEIERIAVFDALERNEIKRLDSV
MLYAASQFAVGSERLNLAIKSIEDELALRLKFFKEQDKMLEYNRLKQRTTEHDLEMISATG
VCKGIENYARHFTGKAPNETPFCLFDYLGIFEREFLVIVDESHVSLPQFGGMYAGDMSRK
SVLVEYGFRLPSALDNRLPKFDEFIHKNCQFLFVSATPNKLELELSKKNVAEQIIRPTGL
LDPKFEVRDSDKQVQDLFDEIKLVVARDERVLITTLTKMAEELCKYYAEWGLKARYMHS
EIDAIERNHIIIRSLRLKEFDILIGINLLREGLDLPEVSLVAIMDADKEGFLRSETSLIQT
MGRAARNANGKVLLYAKKITQSMQKAFEITSYRRAKQEEFNKIHNITPKTVTRALEEELK
LRDDEIRIAKALKKDKMPKSEREKI IKELDKKMRECAKNLDFEEAMRLRDEIAQLRTL
>PROKKA_01589 hypothetical protein
MINYPNLPNSTLEISEQPEVKEITNELLKQLQSALKSNALFTDQIELSLKGIVRILEVLL
SLDFFKNANEIDSSLRNSIEWINNAGESLAKAMKEYEGFFSDFNTSMKANEQEVNTNLNA
NTENIKSEIKKENQLIETTTKLTSYQIFLNQARESATTQISNNQTQAITNINEAKNNA
NNEINTNKTQSLEAITQAKTNANNEINTNKTQSLEAITQAKTNANNEITANKTQAITNIN
EAKENDHYQNIDFFEIE
>PROKKA_01590 hypothetical protein
MVCLSVSFTNKPFLWLWFLVVGCYLVYEWQKKQKDFQSAKSLKFDSVSELEKDLKMEVTN
DEWENIKKLNENLKIFFNVENVSIIFLLVSYLIVGMRVTLSLIDNYGVLK
>PROKKA_01591 hypothetical protein
MLNFTPKSRYIYEKAIKEYYNIDLKPDPELLVRVPHNRLNNTTEINNLAASSNQGRFNSESD
HALAVLSHYEAKLKELDKKLDADSIYSLKNIVAKNLNFDQATHPNVTDNLALLMFNMMPR
TKTQGIELLRWQKEFSNDIKSYEKVKMFVDNAGSFHNLIHDMNFPKVS LNAYLSDIMD
RSFANLKNYQSTSESLKDLSEKFYKTSSLEMFEKSDQSVSDITEILGGAIARFARFDDPS
KALFEALKSDNIKKGLKEFKIADITKDMFDPKSKEFKDIDIYDFTHYLLMVNREPNNNP
TLNRLIQAVKGMQKESEKGIKKPKLETPESEWGHNYSEFKNDGLGAINKLLETKKGFVAGA
FHKEGLGDIDLWVGTPKTKDSNGYGLAHILEKREKQYKRLGLTNEQAKERTKELLRQIPE
VIDKDIKNDYIGHVTIRHNGIEVGLSSQKGNTPLNHYMITSFETDEKVLRELETIGTL
PNDYRDGINYSASNLNKPNTTKKLKDD
>PROKKA_01592 hypothetical protein
MKPAITKNLSQGLATTLGALKYQWTKFTLTGLYRNAPDRILGVKLPKALNEATAGAALK
YHLKRALERSHSISEFSKNLELSAKNAKFSNNTLKIIEELNNGVKQASEEIKEATKTSNL
VKSIREQDTRPFEMIEVKDEAFLOGIEEKLKDSATPLPKGMSVEEFKQTLSEVENKDRF
LEHLETRDNSQDRLKFLSLIEPILREPHIEIFIKDKDSTIAKKEYIKAFKDENDKTRLYML
ITQDNDTILRTFIPKLNERYMRSHVRDADIHSFIQPNRTAKSDNALSDVVVEENPTQK
PLIDQEDLLKRTLNNTTPKPTNLSPLEQANAELAKLESEALESEQEFKAKEQEAKR
KEALKKKLEHEHGNAGNIESQTKIEVGKDIPTQIQAQIPKSRVRLNEREIYDLDYAIVKA
KDLKPSFTTGGTQKRTDMNEE
>PROKKA_01593 Putative beta-lactamase HcpC precursor
MGYASKLALKICLASLCLFSALGAEHLEQKRNFYKGEAYNNKEYERAASFYKSAIKNG
EPLAYVLLGIMYENGRGVPKDYKKAAYFQKAVDNDIPRGYNNLGVMYKEGRGVPKDEKK
AVEYFRIATEKGYTNAYINLGIMYMEGRGVPSNYVKATECFREAMHKG NVEAYILLGDIY
YSGNDQLGIEPDKDKAIVVYKMAADMSSSRAYEGLAESYQYGLGVEKDKKKAAEYMQKAC
DFDIDKNCKKRNASR
>PROKKA_01594 Gamma-glutamyltranspeptidase precursor
MRRSFLKTIGLVIALSLGLSPLSAASYPPKNTKVGLALSSHPLATEIGQKVEDGGN
AIDAAVAMGFALAVVHPAAGNIGGGGFAVIHLANGENVALDFREKAPLKATKNMFLDKQG

NVVPKLSEDGYLAAGVPGTVAGMEAMLKKYGTKKLSQLIDPAIKLAENGYAISQRQAETL
KEARERFLKYTSSKKYFFKKGHLDYQEGDLFVQKDLAKTLNQIKTLGAKGFYQGGVADLI
EKDMKKNNGGIITKEDLASYNVWKRKPVVGSYRGYKIIISMSPSSGGTHLIQILNVMENAD
LSALGYGASKNIHIAAEAMRQAYADRSVYMGDADFVSVVDKLINKAYAKKIFDTIQPDT
VTPSSQIKPGMGQLHEGSNTTHYSVADRWGNAVSVTYTINASYGSAASIDGAGFLLNEM
DDFSIKPGNPPLYGLVGGDANAIEANKRPLSSMSPTIVLKNKVFLLVVGSPGGSRIITTV
LQVISNVIDYNMNISEAVSAPRFHMQLPDELRIEKFQMPADVKNLTKMGYQIVTKPVM
GDVNAIQVLPKTKGSVFYGSTDPKEF

>PROKKA_01595 Flagellar hook-associated protein 1

MGGILSSLNTSYTGLQAHQSMVDVTGNNISNASDEFYSRQRVIAKPQAAYMYGTKNVNMG
VDVEAIERVHDEFVFSRYTKANYENTYYDTEFSLKEASAYFPDIDEASLFTDLQDYFNS
WKELSKNAKDSAQKQALAQKTEALTHNIKDTRERLTTLQHKASEELKSVIKEVNSLGSQI
AEINKRIKEVENNKS LKHANELRDKRDELEFHLRELLGGNVFKSSIKTHSLTDKDSADFD
ESYNLNIGHGFNIIDGSI FHLPLVVKESENKGLNQVYFQSDDFKL TNITDKLNQGKVGAL
LNVYNDGSNGTSKGLQDYIDL LDFS FARGLIESTNAIYAQSASHHIEGEPVEFNSDEAFK
DTNYNIKNGSFDLIAYNTDGKEIARKTIAITPITTMNDIIQAINANTDDNQDNNTENDFD
DYFTASFNNETKKFVIQPKNASQGLFVSMKDNGTNFMGALKLNPFQGDASNISLNKEY
KKEPTAIRPWLAPINGNFDVANMMQQLQYDSVDFYNDKFDIKPMKISEFYQFLTGKINTD
AEKSGRILDTKKSMLTIKKEQLSISQVSVDEEMVNLIKFSQGYAANAKVISIDRMIDT
LLGIKQ

>PROKKA_01596 hypothetical protein

MVVLHSHLENALNQLKELIDLTERDIRDIKLAKHTEIFERNHQQLAIQAFEKEKANIDA
QILSLKNQFPNKEMSGLLDEKTSDFLNQMRESLLVLKEKNLIYSRMAFAVSEFYSSLIQQ
IIPHDTCDYKGSRHVGS HFLRVQA

>PROKKA_01597 Modification methylase HhaI

MGILTFMDFCSGIGGGRLGLEQCHLKCVGHAEINHEALRTYELFFKDTHNFGDLMRINPN
DLPDFDMLVSGFPCQAFSINGKRKGLEDERGTIIYGLIRILKVKQPKCFLEENVKGLIHH
KQQETFKTIIKALQEAGYTTYQIILNSTDFQLAQKRERLYIVGFRKDLKHPHFHPLGLAN
DYYFEDFLDADNECYLDVSNAIFQRYLHNPYHNHNVFLENILTLENAVLDTRQSDRLRYF
NVFPTLRTSRHGLFYTQKGKIKRLNAVESLLLQGFPRDLIAKIKNPNFKESHLLSQAGN
AMSVNVIAAIAKQMLKAI

>PROKKA_01598 Anti-sigma-28 factor, FlgM

MVNAVSSLAPVQSLGNYKRVEKNEKVENNEAALDRVAEIKKAIENNQYKINLHETSHKMA
QDLLGIS

>PROKKA_01599 hypothetical protein

MNSSNLKNWLFPTICFFLFCYILIFLMFFMFKSLQSQSFGSVAETGKKPITTTKKFGKEL
QKQISKIH

>PROKKA_01600 FKBP-type peptidyl-prolyl cis-trans isomerase SlyD

MQNHDLESIKQAALIEYEVREQGSSEVLDSNISKEPLEFIIGANQIIAGLEKAVLKAQIG
EWEEVVIAPEEAYGVYESSYLQEVPRDQFEGIELEKGMSVFGQTEDNQTIQATIKDFSAT
HVMVDYNHPLAGKTLA FRFKVLGFREVSEEEILASHHGGGTGCCGGHGGHGGKGGGCGC
SCSHG

>PROKKA_01601 hypothetical protein

MKRLFFIFPIAPFFLNGEPSAFDLQSGATKKELKQLQINSKNFSNILT KIHSQVEANTQA
QEGLRSVYEGQANKIKDLNNAILSQEESLRALKASQEVQANTLKQOSQTLEDLRNEIRAN
QQAIQQLDKQNKEMSELLTKLSQDLVSQIALIQKALKEQEKA EKPLKSSVPANKNPSLK
AEASKNQEKKTQEKAKVEFDKDL SKQKEIFQEALSFFRNKSYAEAKERLLWLEANSYRLY
YVRYVLGEVAYGEKRYREAIKYYKESALLNKKASYMPVLLWHTAWSFKKIKDDQNYKFL
NTLQHLYPSSEQAKMAQKILENKEKHHHAKP

>PROKKA_01602 Outer membrane protein P6 precursor

MKRSSVSFSLVAFLLVAGCSHKMDNKTAVAGDVSTKT VQTAPVTTEPAPEKEEPPKQEPAPV
VEEKPAIESGTIIASIYDFDKYEIKESDQETLDEIVQKAKENHMQVLLEGNTDEFGSSE
YNQALGVKRTL SVKNALVIKGEKDMIKTISFGETKPKCAQKTRECYKENRRVDVKLVK

>PROKKA_01603 translocation protein TolB

MRYLWFFLIYTI GLFATDKTLDI IKTIQKL PKIEVRSIDNDANYALKLHEVLANDLKTS
QHFDVSQNKDQGVINYAELKDKKVQLVALVSVALENGNKISRLKLYD VDTGTLKKTFDYP
IVSLDLYPFAAHNMAIVNDY LKAPSIAMWKRLIVFSKYIGPGITNIALADYTMRYQKEI
IKNNRLNIFPKWANAEQTEFYTYQYGERTPMILKYNIQKATHENIASSQGMVAVSSVSSD

GSKILMSLAPDGQPDVYLYDTHKKTKTKITRYPGIDVSGVFLEDDKSMAFVSDRSGYPNI
 YMKKLGLKESAEQLLYEGRSNESIDAYKDSIVYVSRENLFEGKTVFNLNLITLNSKYIR
 RLTVNGSNQMPRFMSMDGRNIMYIKKTPQEYAMGLILLDYNQSFLFPLKNVKIQAFDW
 >PROKKA_01604 hypothetical protein
 MSKSAIFVVSGLAFLLYALLLYGLLLGRHNKEAEKILLDLGKKNEQVIDLNLEDLPSEK
 KNEKIEKVTEKQGDFFLEPKKEEPKEEPESLEDIFSSLNDFQEKTDKNAQKDEQKNEQEEQ
 RRLREQQRLKQONQENQEMLKGLQQNLNQFTQKLESVKNKTLDLQVPKQDGVDEKAYQEWY
 AQIYQILYKGWRGVFYHKASVSALIMITKDGEFDYTILSYSDFKDYNKSVMTLLDDLKKV
 DFPYPGGNMISIKVNFTTKEEQ
 >PROKKA_01605 Biopolymer transport protein ExbD
 MNYDNYWDEDKPELNITPLVDVMLVLLAILMVTPTLTLYKEEIALPSGSKTARATQDKMI
 EIRMDKDAKIYIDSQTYEYNSFPDTFNLLSKKYDKDTRVSIRADKRLTYDKVIYLLKTIK
 EAGFLKVSLITSP
 >PROKKA_01606 Biopolymer transport protein ExbB
 MLDSIVYFFNKSQGVTTTLVLVWISLYLVMTLWVFLYKSIVLKIELKREMQSLSNILNGVQ
 DAPEHFMFNKKRNDETKRYSNELLQAWKHQVLKQSTTGLVVLSSIISSTAPFIGLFGTVVE
 ILEAFNNLALGALGQASFGVIAPIISKALIATAAGILAAIPAYSFYLLILKRKVYDLSVYVQM
 QVDILSSKK
 >PROKKA_01607 ATP synthase epsilon chain
 MALLKISVVVPEGEVYTGEVKSIVLPGVEGEFGVLYGHSNMITLLQAGVVEIETENQKEH
 IAINWGAYAEVTNERVDILADGAVFIKKESDDRDAISRAKLLEDASSDRLAVSSVLAKI
 ESL
 >PROKKA_01608 ATP synthase subunit beta
 MEGKIIQVLGPVVDVEFESYLPAlFEALDINFEVNGVQKSLVLEVA AHLGGNRVRAIAMD
 MTEGLVRNQAVKARGKMIIEVPVGEVGLGRIFNVVGESIDNLEPLKPSLTWPIHRKAPSFE
 QQSTKTEMFETGIKVIDLLAPYSKGGKVGLEFGGAGVGKTVIIMELIHNVAYKHNGYSVFA
 GVGERTREGNDLYFEMKEGGVLDKVALCYQMNEPPGARNRIAFGTGLTMAEYFRDEKGLD
 VLMFIDNIFRYAQSGAEMSALLGRIPSAVGYQPTLAGEMGKLQERIASTKNGSITSVQAV
 YVPADDLTD PAPASVFAHLDTTVLNRKIAEKGIYPAVDPLDSTSRILSPQMIGEKKHYEI
 ATGIQQVLQKYKDLQDIIAILGLDELSEEDKKTVERARKIEKFLSQPFFVAEVFTGSPGK
 YVTLQETLEGGGILEGKYDHIPENAFYMGVSIQEVLEKAKNMKNS
 >PROKKA_01609 ATP synthase gamma chain
 MANLRDIRKKIGSVKNTQKITHAMKLVSTSKLRKAEVARNSTRAYALKLDAVFDDVL SKM
 KNQGIEDIQSKYFRELERLEIKKVDIIFITADKGLCGGFNTNTIKKVLACTNEYKEKDIK
 VRLRGIGKKGNEYFSFNGIEVLDKINNLSSMPNYERVQEFMKKVVEDYLSGKTDKVI IHH
 NGFKNMISQEI RVKTI LPIGYKIIHQNPQPSQTETITSEPSGSEDEILD SLAEKYVEYS
 LYYALIDSLAAEHSARMQAMDTATNNAKDLVKTLTISYNKARQEAITTELVEINAGVEAL
 K
 >PROKKA_01610 ATP synthase subunit alpha
 MSQKLKEEIEISSVIEEKIKNFELDCDMAEVGVVSYADGVAKVYGLNGVMSYEVLEFETGD
 KGVAANLEEDSVGVIVFGFGNNIKEGTGVKRTKNLMKVPVGDAVVGRVLNALGEPIDGKG
 EIETNEFSLIEQKAPGIMDRKSVHEPLQTGIKAIDALVPIGRGQRELIIGDKQTGKT TVA
 IDAIINQKGQNVICIYVAIGQKESTVAQVVRKLEEYGAMEYSVVINASASDSAMQY LAP
 YSGVAMGEYFRDHARHALIIYDDL SKHAVAYREISLILRRPPGREAFPGDVFYIHSRLLE
 RAAKLCDEKGAGSLTALPIVETQAGDVSAIYPTNIISITDGQIFLETDLFYSGIRPAINV
 GLSVSRVGGAAQIKATKQVSGTLRLDLAQYRELQFTQFASDLDEASKKQLERGQRMVEV
 LKQAPYSPLPIEKQVVIYAGAKGFLDSVSVKKVDFEEQLHPFLEAKYPQVLEEIH TTK
 ALDKDLEAMLRKVLEEFKLTYS
 >PROKKA_01611 F0F1 ATP synthase subunit delta
 MQDLKVIANHYTKALKNHTKGD LALLEEIVVGLKNVAEAIKLHKLNQVLAHVSLKV KKEV
 VFEILEKITSIKACSVLKPMEVVLKNNRLDVLELITQDLSFDSKKTLEATLLVPQKLEN
 KELEEVRQKLQARFNAPVGITQDTWSKKGVSLSVSSLDLEIGFSKEDILKKIEKQVIQSI
 >PROKKA_01612 F0F1 ATP synthase subunit B
 MFLVKMVLGFLILLSPICATGLDVSQTDIIERSLNFLLFAGILWYFLAKKLRSFLRSKSL
 EISKRL EEIQAQLKVSKENKKLLKELEQA KEKAELIISDANKEAYTITQKYELQTKMDV
 ENLIKNSKALMDLEVKKIKRELVESVFKDLRESKKVSFNAQDCVNILKQRL
 >PROKKA_01613 F0F1 ATP synthase subunit B'
 MNISVNPYLMAVVFVFLLLWAMNVWVYRPLLA FMDNRQAEIKDSLAKIKTDNAQSVEI

GHQIETLLKEAAEKREIIAEAIQKATESYDAVIKQKENELNQEFEEAFKQLQNEKQALK
EQLQAQMPVFEDELNKRVMGLGS

>PROKKA_01614 Chromosome-partitioning protein Spo0J

MAKNKVLGRGLADIFPEINEVYEQGLYERANRVVELGIDEVMPNPYQPRKIFSEESLEEL
AQSIKEHGLLPVLVSENGRYHLIAGERRLRASKLAKMPTIKAIVVDIEQEKMRVALI
ENIQREDLNPLELARSYKELLESYQMTQEELSKIVKKSRAHVANIMRLTLSSKVQNALL
EEKITSGHAKVLVGLDEEKQELILNSIIGQKLSVRQTEDLARDFKINANFENKKHGFKET
QTLIAEDELERLNQSLWDHYKCLKATLKGKIVLRKYENSLLLEAFMKKMMS

>PROKKA_01615 Sporulation initiation inhibitor protein Soj

MSEIIAVANQKGGVGKTTTAVNLSASLAAHEKKILLIDFDPQANATSSLGFRDKIDYDI
YHVLIGRKQISQVILKTQMPFLDLVPSNLGLAGFEKTFYDSQDENKRGELMLKNALESVV
GLYDYIIIDSPPALGPLTINLSAAHSVIIPIQCEFFALEGKLLNTIRMLQKSTNPKL
KIRGFLPTMHVPQLNLTKGVLAELFKYFDSEFFRDSATGEYIMIPKSVKLAESPSFGKPI
LLYDIKSNGSIAYQKLAQSILQG

>PROKKA_01616 Bifunctional ligase/repressor BirA

MRQCEKRVFDSL PSTQTYLLEKLKNNELKAPILIVAKNQSAGIGSRGNIWESAKSALTFS
LALNASDLPKDLPMQANALYLGLFLKEVLKELGSQTWLKWPNDLYLGGQKIGGVLNVYK
GMLVCGIGVNRVSTKWACLDIGTSDDWIIIEGFLKKIEENLFWGEVLSKYALEFHRNDSFS
FHNDWGELVSLKDAQLLEDGRICIKGKIYDRM

>PROKKA_01617 Methionyl-tRNA formyltransferase

MRIVFMGTPGFAEVILRALVENKDIEVVGLFTQMDKPFGRKKELKAPETKTYILENHLNI
PIFQPQSLKEPEVQILKDLKPDFIVVVAYGKILPKEVLAIAPCINLHASLLPKYRGASPI
HEMILNDDRIYGISTMLMDLELDSGDILESASFREDYLNLDALSLKLAHMGADLLLSTL
KNFSSITRKPQDHMQASFCKKITKADGLVGFKDAKSLFLKSLAFKSWPEIFLENSLKLLE
VELVENEKSHKEGEILAIDEKGVLVGCLKGSVRIARLQAVGKKPLKAKDYLNKRLKIGG
ILA

>PROKKA_01618 iron-dicitrate transporter ATP-binding subunit

MGNKDHSKSGSSWHKWDLHAHTPYTNLNKEYKCSSEKFIQKLCDISKIDCIGLTNYFKFNEK
EFELKEKIEKRGIVFYVLEVRDYQNKEDQCLDFHIIIFSDKVTQQEIDNFLRNADANVG
GTEKKLADLEKDDFDKVVVNFQDLECLEKESLKLKGKYLGLFLSRGHGSSSRSSSNYEKI
VKKVHFLIHSSDKQENLKKDREYWLEYNKPLIQSSDAHEEEQIGKKYTWIKAECTFEGLK
QIVYEPETRVSIGEEKPQDPLYKIDSVGLNFDDEEVKTTNEQDDTPFCYAGFNETLFFSPN
FTCVIGGRGSGKSTLLQLIASAIKNKSFVKGLRLETIQKSLKIQSSIDIADNIEYLAQNE
VEEFATNVSKFTEAIFNRIDSKSGGKLKELERQIKESIEKFDEQIRYWQEKNKLEEQLKE
SEKIRKKYQSIINTFTDKNYLDKKDKLQKKRKALIDLEQSKEGFWTFIKELKRVVNFESK
ENMEEKNSYDKVYNQLKQDICKELERIDINIKNGCFKSEDEKIRTLKTEREVLQSQEIGEF
LKEKGVSDENIGDIRNANDHLANIKMDINGLEHEIENANKIENFSYEDMDKNIEEFKDQ
INEELSKINSFAFKKISKNHKEVKPITIKYRLNEDIFEVFEDFDKLVDKGFNTQRHQSKI
KEYLKEIELKNVTGMQHAEFIEELYNSIDNKKAIFYETLKDIFDREIHFQIYRLLILKHL
RNVEKYKIFKVRDYDKRALNETSFGQKCTAVLVVLLSLGNNPIIIDPEAHLDSALIANYL
VALIKKQKQKRQIIFATHNANFVLNADAELIIQLKNENNKIVAQSFMIESDAYKEDLLKL
EGGEKAFKDRERKYGITKDKTKNKE

>PROKKA_01619 hypothetical protein

MQENQTRPFICPKCQEPINVNEALYKQIEQENQSRFLAQQKAFFEKEVNEKRAQYQSHFKA
LEQKEEALKEREREQKAQFDDAVKQASALALQDERAKIIIEEARKNAFLEQQKGLELLQKE
LEEKSKQVQELHQKEAEIERLKRENNEAESRLKAENEKKLNEKLEMERERIEKALHEKNE
LKFKQQEEQLEMLRNELKNAQRKAELSSQQQLQGEVQELAIEEFLKQKFPLDSVEEIKKGQ
RGGDCIQVVHTREFQNCCKIYYESKRTKEFQKAWVEKLKSDMREIGADVGVIVSEALPKE
MERMGFLFEGVWVCSFEFEKGLSAVLREGVIQVGLAKKSQENKGDKNLLYHYLTSSEFSM
QVNAIIIEGFQELRADLESEKRAMARIWKSREKQMEKVFEGTINMYGSIKGIAGNAIGQVK
ALELGYDGEDLED

>PROKKA_01620 hypothetical protein

MIDQLQIVPKVQVLENNTMGRSLASALQKASIPRPISAHRVPLINHFLQVISQSQSDEN
KSIKPIAFGSLIKLNPNNENPLFAFIALSCTQFHYSGMRLVAPYENDSIIFELLHCIDF
QARNFERQQLVQLDPKEIKTQILERLKL

>PROKKA_01622 hypothetical protein

MLFSKDTDFVPVLEAAWEKGFVFIANIQEGPNLVPPDLKKSCGVRKRSVAEIVDNL PKN
QYPPKKKNFSPNEFPNNPFKDQLFKKN

>PROKKA_01623 50S ribosomal protein L19
MKNRYIQQFEDTQLKDKAMPTFKAGDTLRLGITIKEGEKTRTQYFEGVCIAIRGNGVDKT
FCVRKIGANNIGVEKIFPFYSESLASVEVLRVGRVRRAKLYYLRDRRGKAARIKEVRH

>PROKKA_01624 tRNA (guanine-N(1)-)-methyltransferase
MKFSVLTLPFQVLWPYFKDSILKRALEKNLFELEVLNLRGFSTNKYQKADHTLIGGGAGQ
ILDPEMVENALHSVKNPKHTIFLSAVGKPFKQIDAMRLAQKKHVVLVCGRYEGFDERSIE
LSADEVFCIGDFILTGGEGLGALCLIDSIARHIQGVLGNARSLENESFENHYLEAPNFANA
VFKSKEINKIPAPLEYSGKNHAKIKQLKLDLSKLRTKFYRLDLFKQHKS

>PROKKA_01625 Ribosome maturation factor RimM
MVSMLLVGRIGKSVGLNGGLKLHLESDFPECLKKGVKVSAPINAFSCTSSFKEYTIHSY
EHAKNLLFLETIQTPKAKELTNLGLFMSEAESKKLCVLKEGEFFYCDLVGLSVVEENEI
LGKVIEIQRISQTDYFMVETTLNLVEKGLAKIFLIPYRDFYIKEILLQDKKITTHNAKTL
LENS

>PROKKA_01626 hypothetical protein
MRESDDLNTPFQADCKDYSYCVATFLEKYLKKVVSFPQALSVEYTLLEDKVKQITIYTH
PSDMGHVIGKEGKMVSAIKAFVSGVKAKDGFYSYKIVVFANNDKNSDKNPHALGDQTP

>PROKKA_01627 30S ribosomal protein S16
MTVIRLTRLIRGKKKPFYRVVVTDSRKRDRGGWIESIGYYNPLSEPKDVRIDKERLNYWKG
VGAKMSERVEKLSQKA

>PROKKA_01628 Signal recognition particle protein
MFQALSDGFKNALNKIRFQDDEKALDRALDELKKTLLKNDVHHKVARELLKKVESQTKAS
GIGKQQFLDALEKSILLEILSTKGSSGFTFAQTPPTVVLMAQLQSGSKTTTTAKLAHYLKT
KNKKVLLCACDLQRLAAVEQLKVLGEQVGVEVFYEENKSVKEIANNALKRAKEAQFDVLI
VDSAGRLAIDKELMQELKEVKEILNPHEVLYVADALSGQDGVKSANTFNEEIGVSGVVL
KFDSDSKGGIALGITYQLGLPLRFIGSGEKIPDLDFVFPERIVGRMLGAGDIISLAEKTA
SVLNPNEAKDLSKKLKKQFTFNDFLNQIEKVKKLGSMSSLSMIPGLGNMASALKDSDL
ESSLEVKKIKAMVNSMTKKEQENPEILNGSRRKRRIALGSGLEVSEINRIIKRFDQASKMA
KRLTNKKGISDLMNLSQAKNQTPPKMR

>PROKKA_01629 Valine--tRNA ligase
MKQEPTTYQPKIEKIEKIYEICSHRGYFEIDGNEKIQEKGRFCLMPPPNVTGILHIGHA
LTLSLQDILVRYKRMDGYKTLYQPLDHAGIATQNVVEKQLLSQGIKKEDLGREAFVQKV
WEWKEKSGGAILEQMKRLGVSAAFSRTFTMDKGLQRAVKLAFLKWYEKGLIIQDNMVN
WCTKDGAISDIEVEYEERKGALEYIRYYLENQKDYLVVATTPETLFGDSALMVNPNDER
YKHLVGQKAILPLINRTIPIIADEHVEMEFGTGCVKVTPGHDFNDYEVGKRHHLEMIKIF
DEKGILNAHCGEFENLERLEARDKVVAKLKEDALLEKIEEHAHQVGHYRCHNVVEPYVS
KQWFVKPEIAQSSIEKIQQGLARFYPSNWINNNAWMRELRPWCISRQLFWGHQIPVFTC
ENNHQFVSLDTPLNCPCTCKSEKLEQDKDVLDTWFSSGLWAFSTLWGQEKSDLFNENDLK
DFYPNTTLITGFDILFFWVARMLFCSESLGELPFKDIYLHALVRDEKGEKMSKSGNVI
DPLEMIEKYGADSLRFTLANLCATGRDIKLTSTHLENNKNFANKIFNAVSYLKLKQEAFAK
DRERLNEYQTPLGRYAKSRLNSATKEARNALDNYRFNDATTLRYFLWGEFCDWFIEFSK
VENEAIDELGSLVKEALKLLHPFMPFISESLYHKLSENTELENTHSIMVMPYPKDLAQDEK
LEHEFEVIKDCIVSLRRLKIMLETPIIVLKEASVGLREAIENTERLQTYAQKLARLEKVS
VISYKPLKSVSDVGEFCQTHANLENLDLSPLVARLKKQLEKLEKEKLLNLHNENFVKNA
PKSVLEKAQEGKLTLEKESKIKQELDLLEQP

>PROKKA_01630 Flagellar assembly factor FliW 1
MNYFLKAPILGFEHINEVRLEKIDSLFSRLMGQTNSPMALDMVLNPNPYCLREYSFVIPKY
IELLLELDSSHKVEVYCVVVLQKNLEDSMVNPLAPLVFNSKNGFGAQVALSMMDYPDFGF
RDPLKSFVVKERERA

>PROKKA_01631 UDP-N-acetylglucosamine--N-acetylmuramyl- (pentapeptide)
pyrophosphoryl-undecaprenol N-acetylglucosamine transferase
MKFALTGGGTGGHLSIAKALAIELEKQIEAIYLGSTYGQDREWFENSPLFSERYFFNTQ
GVVNSKSFKKIGSLFLQAKAAFAKAKEILKKHQITHTISVGGFSAGPASFASLLNKIPLYI
HEQNAIKGSLNRYLSPKAKAVFSSYAFKDKGNHVLTSYPVQNAFFDFARTRTEIKHILFL
GGSQGAKAINFALLNAPKLTKQGIKITHICGPNSEYQVRFFYQELGLLDKIDLFAFHNN
IIEVMHRADLCVSRAGASSVWELCANGLPTIFIPYPFASNNHQYYNVLEFEKENLCYVVP
QNELLPKKLFVIRKLNQKDDQGNKNLTIISNLQQKIAKDGTKTIIETILRA

>PROKKA_01632 Helicobacter outer membrane protein
MQNFVFNKKWLICSSLLPLFFLNPLAAEDDGGFFMGVSYQTS LAVQRVDNSGLNASQDAST

YIRQNAIALESAAVPLAYYLEAMGQQTRVLMQMLCPDPSKRCLLYAGGYQQNAQNGDTGN
NPPRGNVNATFDMQSLVNNLNKLTQLIGETLIRNPENLSNAKVFNVKFGNQNTVIALPEG
LANTMDALNNDITNALTTLWYNQTLTNKSFNSGNSVNFSPPEVLQHLLQDGLATANNNQTI
CSTQNQCTATNEAKSIAQNAQNI FQALMQAGILGGLANEKQFGFTYNKAPNGSDSQQGYQ
SFSGPGYYTKNDNTTQAPLKALPAGATVGSNGQYTYHPSSAVYYLADSI IANGITASMI
FSGMQNFANKAARLTGTSSYSQMQDTINYGESLLNNTVAYGNFITNWWAPYLDLNNKGLN
FLPSYGGQLNGANSQTPQLTPQQAQQEQKVMNQLEQATNAPTPAQINRILANPYSPTAK
TLMAYGLYRSKAVIGGVVIDEMQTKVNQVYQMGFARNFLEHNSNSNNMNGFGVKMGYKQFF
GKKRMFGLRYYGFYDFGYAQFGAESSLVKATLSSYGAGTDFLYNVFTRKRGTEAIDIGFF
AGIQLAGQTWKTNFLDQVDGNHLKPKDTSFQLLFDLGIRTNFSKIAHQKRSRFSQGIEFG
LKIPVLYHTYYQSEGVTAKYRRAFSFFYVGYNIGF

>PROKKA_01633 *Helicobacter* outer membrane protein

MIKKARKLIPFFLIGSLLAEDNGWYMSVGYQIGGTQQFINNKQLLENQNIINSVTQSAIN
VAGPTTGLITLSSQSVIDALGYGVSNVTGNGQLEGISNILNQIGKRKDFYSSRQISSISQQ
IIGLKGSDDLKAHSSQITAKLLSNTQSAFDQGIALSTSISSINSNLNPSNNTQEVKAQL
QNTAQSMALLQQIEHSITKTTSTTYAQSLLSNLTDVNASSNNTAYVSALVNALNTLGV
GVFPTTTSTHVVLNPPGQIVFYFANSILGSTSSNSNNQQQYNNTLLMNTLQGTLSANNQI
NPNGCTNQVQCLEQFIQNLSPLTATPTSNNQANQQVQAIQKLQSVAINTLDNNAINNTT
YNLNNLHNALNFQAYESTIEQYNNALKQISWISFTEPKNLLKNTSNNYQIGTVTNAQQQN
ISAYDCTSATGSLSSNASSGISCSATSSTNTNSFDNSLVATSKVQTINGKEQIGVNSFN
LVSQVWSVYNSLKTSEENLQKNANILCTNGSQSGTSPCNSSSSGGLSISGNAQLQNILSST
NGTTTTTQAKNNASKKAMVVVNNEEEAKTTNLAQNSGPTTQSSNSTVMGALNTVLQNNIN
FQQSIQSAFQNOENNIQAWANAIYNTNGNQSQDMTLNTNQDLRIQLRANFYQLINTINQQ
APTDMSALIAQSQQTQQTSGATSNNACASGMNGSSSNWCYQQWSDSKAYYSGLQSALGY
QTQATTQSGSSGGNSIAYNVQQITLTSGGLLNNIISLKSFSNSGSGNGSSGGTNGTSQI
NTAYQMLTDASDGKLGTYSSSNSGNSNGYTPCNSGANGNSSNCYTPSKTQNTTTATTTTD
SNLQKVYNDAQKIANIIASSGNNKGVENGLKQFFKALKNSSSLSNLCNGSSNCSGGLIN
LLGAIPINGVSDTNLVLNLLTEFIKTAGFIQNNDSVSTSLTNAFQAITNAISQGFQALQ
NDISPNAILTLLQEITSNTTTIQSFSQTLRQLLGDKTFFMVQQKLIDAMINARNQVQNAQ
NQANNYGSQPVLSSQYAAAKSTQHGMNSGLGVSLGYKYFFGKARKLGLRHYFFFDYGFSEI
GLANQSVKANIFAYGVGTDFLWNLFRRTYNTKALNFGLFAGVQLGGATWLSSLRQQIIDN
WGSANDIHSTNFQVALNFGVRTNFAEFKRFKAKFHNQGVISQKSVEFGIKVPLINQAYLN
SAGADVNYRRLYTFYINYIMGF

>PROKKA_01634 Pyrroline-5-carboxylate reductase

MEILQFIGYGNMAQAILEGAHEILSKRFILEITGRNPEKIAPFLQEKNIAQIVPYKDAI
DIHQKFVFLFLFKPYNLKDFNYQQQAKSVLSALAGVGFKALSDAINSLHYLKCMFNIAASKF
ALSSTAVCEKSPMPLISQKALSIIESFGNCVRVGNEEQVDSSVATNGSALAFSLVASSL
KDAGIREGLNARDSLELVKMSFKGFAKLEKERPEMIIIEQICTPKGATIEGLSVLEKKGV
RGAFIEACQKSVKKMRL

>PROKKA_01635 Adenosine monophosphate-protein transferase NmFic

MHLDRQSLKAKHLIQSGLIDTIEVGTIKGLQEIHRFLFEGLYEFAGKIRDKNISNFRFA
NCLYLDLILPRIESMPQNNFNQIVEKYVEMNIAHPFLKG

>PROKKA_01636 Adenosine monophosphate-protein transferase NmFic

MLKKELKKIVLWDRIDKAAYLSAIKRSPVNDLEIKTLLKKHLSSNTNDPLTLIKGITQSY
YYEGL

>PROKKA_01637 Endoribonuclease YbeY

MLEIDNQTPLESDFLLLEKIADVLAQTQIIELVLVSDETMREINRDLRNCYATDVLSFP
LEAIPHTPLGSSVINAPLAQENALKLGHSLENEIALLFHGVHLHLLGYDHEKDKGEQRQK
ESELIEAFNLPLSLIERAQD

>PROKKA_01638 Flavodoxin

MKGIGIFFGTDSGNAEIAEKISKAIAGNAEVVDVAKASKEQFNSFTKVILVAPTAGAGDL
QTDWEDFLGTLESSDFANKTIGLVGLGDQDTYSETFAEGIFHIYEKAKAGKVVGHTSTDG
YHFEASKAVEGGKFVGLVIDEDNQDDLTDERISKWVEQVKSSFA

>PROKKA_01639 Inner membrane protein YohD

MQEALLRFQEGFKEWGYLILFLYSLGGGYVGIVIASILSATTHALDIKITILVAFLGNMV
GSGALVIFARYQKRDFLQYFQKHRRKLALASLWVKRYAFLMIFINKYLYGIKSVVPLAIG
FSKYPLKRFLWNLVSSFLWALIVGSVSFQASDWKTLYERLSHYTSFFVISFVLIALLI
WFLLRYSRKMGE

>PROKKA_01640 Cytochrome oxidase maturation protein cbb3-type
MNTEILTIMLVSVLMGLVGLIAFLWGKSGQFDDEKRMLESVLSDSASDLNEAILQEKR
QDPKN

>PROKKA_01641 Ferredoxin--NADP reductase
MDQEILDVLIIVGAGPGGIATAVECEIAGVKVLLCEKTESHSGMLEKFYKAGKRIDKDYK
KQVVELKGHIPFKDSFKEETLENFTNLLKEYRITPSYKTDIESVKKEGELFKITTTSTNTT
YHAKFVVVAIGKMGQPNRPTYKIPVALSKQVVFSSINDCKENEKTLVIGGGNSAVEYAIAL
CKTTPPTTLNRYKKEFSRINEDNAKNLQEVLLNNNTLKSGLGVDIESLEEDNTQIKVNFTDN
TSESFDRLLYAIGGSTPLEFFKRCSELDPSTNIPVVKENLESNNIPNLFIVGDILFKSG
ASIATALNHGYDVAIEIAKRLRS

>PROKKA_01642 Major Facilitator Superfamily protein
MLRKNILAYYGANFLLIIAQSLPHAILTPLLLSKGLSLSEILLVQTFFSFCVLVAEYPSG
VLADLMSRKNLFLVSNVFLIASFSFVLFFDSFIFMLLAWGLYGLYSACSSGTIEASLITD
IKENKKDLRFLAKNNQITYLGMIIIGSSLGSFLYLKVHAMLYIVGIFLIMLCMLTIVIIYF
KEKEGDFKSQKSLKLLKEQVKGSLKELKDNPKLKILLVGHLLITPIFFMSHFQMWQAYFLK
QGIKEQYLFMFYIAFQVISILIHFLKAKNYSQKIALSSLLVLLGVSPLLLSNIPYCFIGV
YALMVAFFAYMSYCLGYQFSKFVSKNNISSLSLLSSCVRVSVLILSLSSLELRYSPL
TIITMHFALTIIILFFFLYKAKPFDE

>PROKKA_01643 Glucose-6-phosphate isomerase
MLTQLKTYPKLLKHYYEIKEMHMRDWFSSDKERASRYFVQLESLSLDYSKNRLNDTTKL
LFELAKDCSLKEKIEAMFKGEKINTTEKRAVLHTALRSLNDAEILLDNMEVLKSVRSVLK
RMRAFSDSVRSRGKRLGYTNQVITDIVNIGIGGSDLGALMVCTALKRYAHPRLKMHFVSNV
DGTQILDVLEKINPASTLFIIVASKTFSTQETLTNALTARKWFVERSGDEKHIKHFVAVS
TNKEAVQQFGIDEHNMFEFWDVVGGRYSLWSAIGLSIMIYLGKKNFNALLKGAYLMDEHF
RNAPFESNLPVLMGLIGVWYINFFQSKSHLIAPYDQYLRHFQKFIQQLDMESNGKRISK
GEIIPYDTCPVVWGDGMINAQHAFFQLLHQGTHLIPIDFIASLDKKPNAKGHEILFSNV
LAQAQAFMKGKSYEEAFGELLSKGLDKDEAKDLAHRVFFGNRPSNILLLEKISPSNIGA
LVALYEHKVVFQGVIVDINSFDQWGVLELGKELAVPILQELEGHKSNAFYDSSTKHLIELY
KNYNQ

>PROKKA_01644 Putative outer membrane protein
MKKASQVLFFFGAFLSSSLQGFEAKLNGFVDQSSSTIGFNQHKINKERGIYPMQQFATIAGY
LGLGFSLLPKKVSDHVLKKGKIGGMVGSIFYDGTKKFEDGSVAYNLFGYYDGMGGYTNIL
QTDLSLETQNMKHKNVRNYVFSDAYLEYAYKNYFEIKAGRYLSTMPYKSGQTQGFQVSGQ
YKKARLTWFSSFGRAFAYGSFLMDWFAARTTYSGGFTKNDKGGYDSHGRKVLYGTHAVQL
TYKPHRFLIEGFYYLSPQIFNAPGVKIGWDSNPNFSGTGFRSDTAVMGFFPIIYPPMIVK
SNGSPVYRYDTPATQNGQNLIIRQRFDINNYNVSIIFYKVFNANGWIGNMGNPSGVIMG
SNSVYAGFTGTALKRDAATILLSGCGGTHFAKKFTWKFATQYSNSVVSWEARAMISLGYKF
NEYLSGSVDLAYYGVHTNKGFKAGENGVPVKNFPALYSRDRSALYTALVASF

>PROKKA_01645 Carbon starvation protein A
MQKSLVSLAWVFVAILGAICLGVLAHLKHGESINTLWLVASACIYSIGYRFYSHFIAIRV
LKLDSDRATPACIRNDGRDFVPTDKAITFGHHFAAIAGAGPLVGPILAAQMGYLPSILWI
LIGSVLGGCVHDFVVLFAIRRDGKSLGEMIKLEMGGFVGMIASLGILGIMLIIIIAILAM
VVVKALAHSPWGFFTAMTIPIAILMGLYMRFFRPHKILEVSVIGFILLIIAIYAGKYVS
LDPKLASIFTFDASSLAWMIMGYGFVASILPVWFLAPRDYLSFLKIGVIGVLVVAIVF
VAPPLQIPKITPFVDGSGPVFAGSVFPFLFITVACGTISGFHALISSGTTPKMLAKESDA
RLVGYGSMVMESVVALMALVCAGILHPGLYFAINSPEVSIGKDIADAASVISSWGFSISA
EEISEMTKNIGESSILSRTGGAPTFAIGLAMIVYHILGDPSVMAFWYHFAILFEALFILT
AVDAGTRTARFMIQDLLGNVYKPLGNLSSYKAGIFATLLCVAGWGYFLYQGTIDPKGGIY
TLWPLFGVSNQMLAGMALLLVTVVLFKMGRFKGAMVSALPAVLILFITFYSGILKVVPKS
NDSVLNNVSHVAQMQUIKEKIALTTDEKALKTLQKSFFNHDAIDAILCVFFMLVALLVLIV
SVRICSNAYFKNQIYPPLAETPYIKAA

>PROKKA_01646 Glutamine transport system permease protein GlnP
MALDWDFMFHSIPAFKLGLELTLYISFFGILLSLVGFLCAIVLYFKMRFISPVVYVYGE
IARNTPLLIQLFFLYYGLNEIGLSALECAIALGFLGGGYMSQSFLLGFKSLASIQRESA
LSLGFSPKMMYYIILPQSLSVSMPSIGANVIFLLKETSVMVGAIALTDIMFVAKDLIGIY
YKTTESLLMLSLTYLIALPLSVLFVILERSFKKKVA

>PROKKA_01647 Glutamine transport system permease protein GlnP
MGVLLELDNLKRLLEGFEITLLIALSSAVISIIVGTLLGSLMAFGSQIVVLACRVYLESI

RIIPLLAWLFIVYFGLASWFDLHISAVLASVIVFSLWGAEMMDLTRGVLTSSVKHQVES
ALALGMDSKKVIFNIIFPQSFLSLLPSSLNLFTRMIKTTALVSLIGAIIDLLKVGQQIIEL
NLLRMPNASFVVYGVILMFYFSLCYSLSLYSSYLEKKFQHIRG
>PROKKA_01648 Arginine transport ATP-binding protein ArtM
MSVILETKGLKKTYQNHVLVDGINFTLNKGEVAVILGPSGCGKSTFLKCLNGLEKIDGGE
ILFENTNLNNKTTNWNQMRQKIGMVFNQNYELFPHLNVLDNILLAPLKVQKRSKDEVISQA
VELLKRVGLEHKQQAYPKELSGGQKQORVAIVRSLCMRPKIMLFDEVITASLDPEMVKEVLE
VILELATTTGMSMVIVTHEMKFAQKIASKIVFFDSGKIAEENSAKEFFNHPKSQRAQKFLE
TFHFLGSC
>PROKKA_01649 ABC transporter glutamine-binding protein GlnH precursor
MKTNGLLKMWGLFLVLIALVFSACSDSHKEKKDALEVIKQRGVLKVGVSFSDKPPFGSVDS
KGKYQGYDVVIXKRMALDLLGDENKIEFIPVEASARVEFLKANKVDIIMANFTRTKEREK
VVDFAKPYMKVALGVISKDGVIKNIEELKDKEIVNKGTTADFYFTKNYPNIKLLKFEQN
TETFLALLNNKATALAHDNTLLLAWIKQHPEFKLGITSLGDKDVIAPAIAKKGNPKLLEWL
NNEIDSLISSDFLKEAYQETLAPVYGDEIKPEEIIIFE
>PROKKA_01650 hypothetical protein
MSYFYKRCLKFSLIGLLGLLSVQLDARSFVDGDLDIQKFSYEDSLLKKGDPNGVHKVQVR
DYKGMQEAIEHSEIRIALKPGVKKEVKKGKIYSAQINDGMCYAFRMLQTGDNTTGLDSK
EFPKQSREKKGRVITLIGKDEVPLYLILETDCQVGDIAKISLVGNFDGTGFLTEHKFKDAK
PIY
>PROKKA_01651 L-fucose-proton symporter
MVLSFEKLLNLDPOSSAHYLVYYWGGAMVGRFLGSVLMNKIAPNKYLAFNALSSIVLIAL
AIIIGGKIALFALTFFVGGFNSIMFPTIFSLATLNLGHLTSKASGVISMAIVGGALIPPIQ
GAVTDMLTATESNLLYAYGVPLLCYFYILFFALKGYKQEENS
>PROKKA_01652 L-fucose-proton symporter
MQKTSNTLALGSLTALFFLMGFITVLNDILIPHLKPIFDLTYFEASLIQFCFFGAYFIMG
GVFGNVISKIGYPFGVVLGVITATGCALFYPAAHFGSYGFFLGALFILASGIVCLQTAG
NPFVTLLSKGKEARNLVLVQAFNSLGTTLGPIFGSLLIFSTTKMGDNASLIDKLADAKSV
QMPYLGAVFSLLLALIMYLLKLPDVEKEMPKEETTQKKPVFAQTLFCFWGFRDLFLCGRRS
GDWIILGAKL
>PROKKA_01653 Putative permease YicO
MPLLKAIPSNAIYPVLVVGVLMFVSLEGVNFKDMAISVSTFLTVMMPLTFSIADGLAF
GFLSYSIIKLVQKDFKALNSGIIILCIISVSVFIFR
>PROKKA_01654 Putative permease YicO
MPVIITFFVTDLFDLSLGTLAGIGHKTDFFNDEEKNKELEKTLEADAVASLGSVVGSTT
TAFIESASGVEEGDAQGLQRFLPDYFLF
>PROKKA_01655 putative adenine permease PurP
MGFFKLKEHNTNIATEFRAGLTTFITMIYIVPLNALILSHANMPYEALLSATAIITILSS
VFNGLWANTPIAMSVGLGLSAYFSFGLVQGLKLPWQSALGIVALSERFL

**The Characterisation of Dynamin-Like Proteins in the Gastric Pathogen
*Helicobacter pylori***

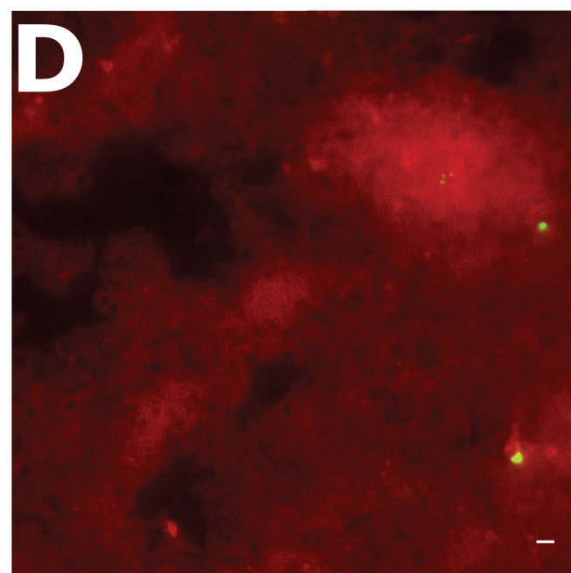
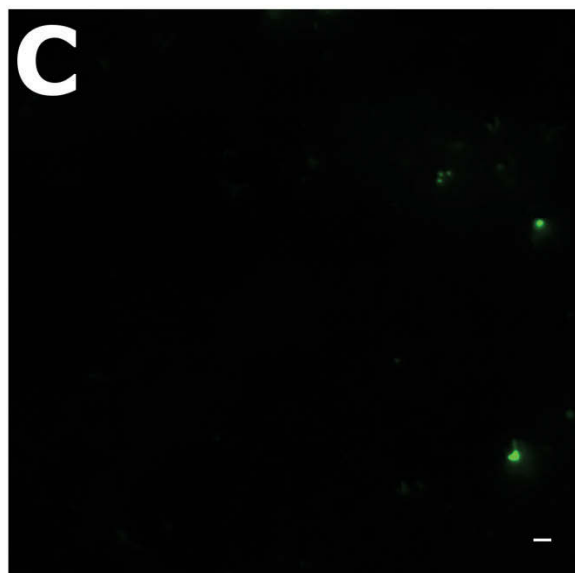
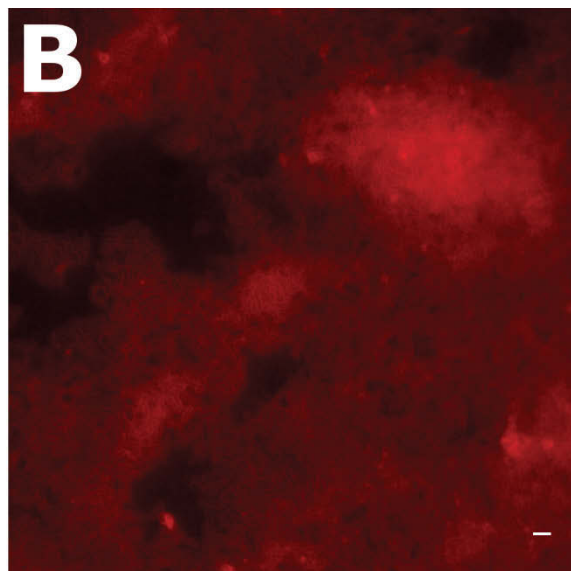
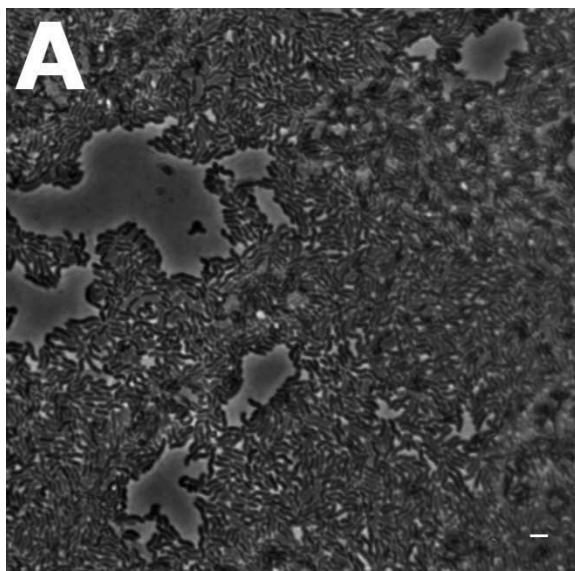
Emma Mary Dawson

**S6: Membrane Integrity Individual Filter Images from the Delta
Vision Elite**

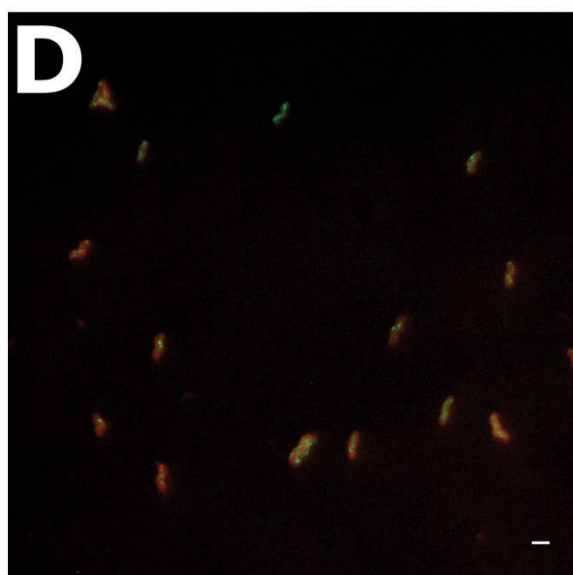
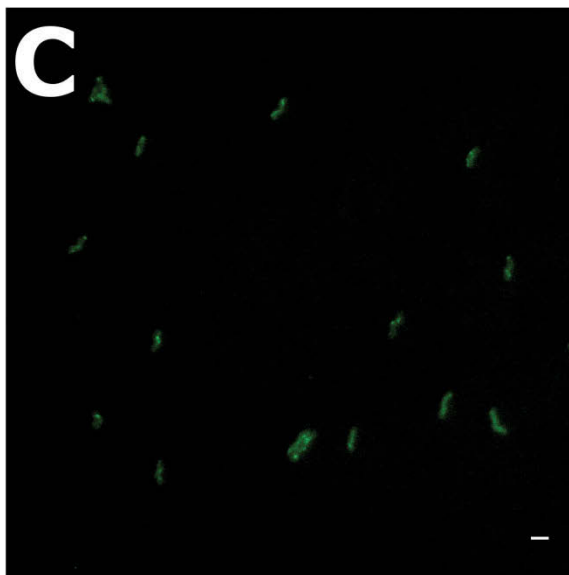
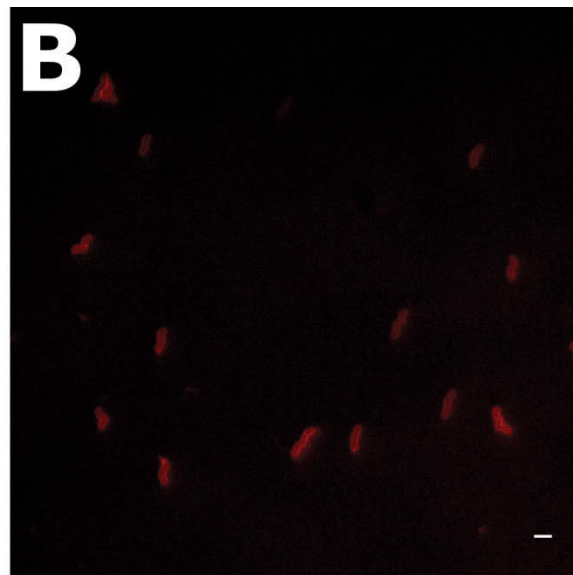
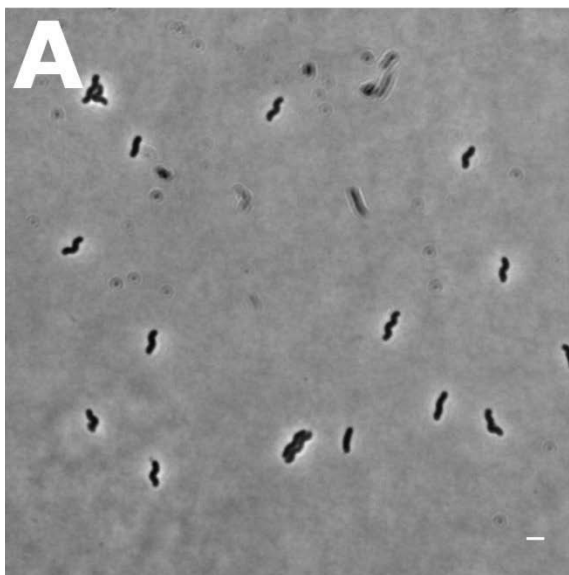
Below are the individual channels taken on the Deltavision Elite in the investigation of membrane integrity. The strain name is found above and the corresponding filters are listed below:

- Panel A – Phase Contrast
- Panel B – TRITC at 300 ms (FM4-64 fluorescence)
- Panel C – FITC at 500 ms (Sytox Green fluorescence)
- Panel D – Overlay of TRITC and FITC

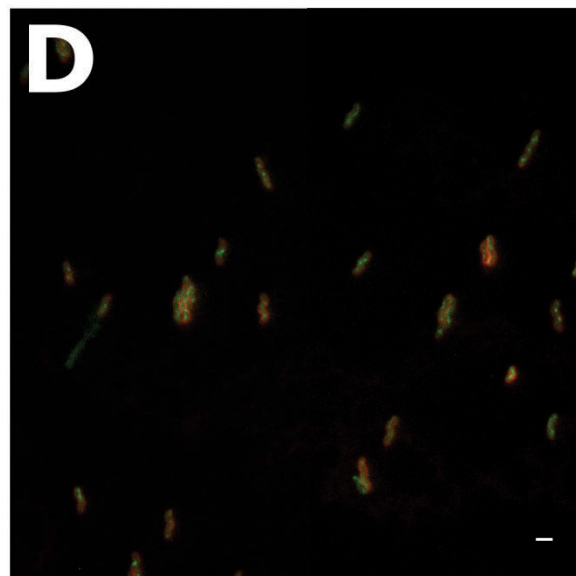
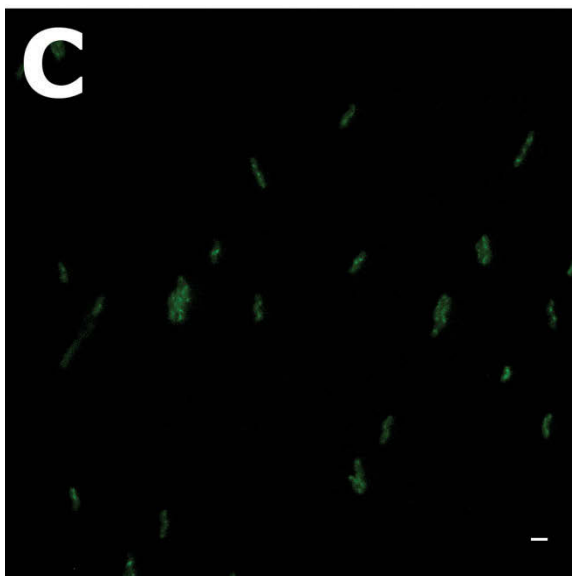
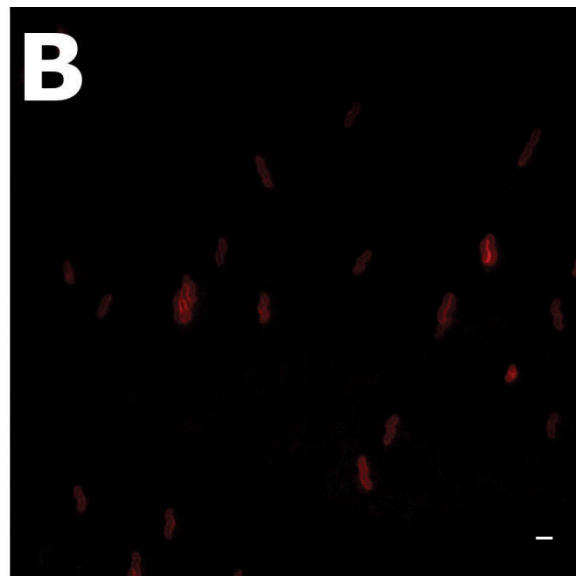
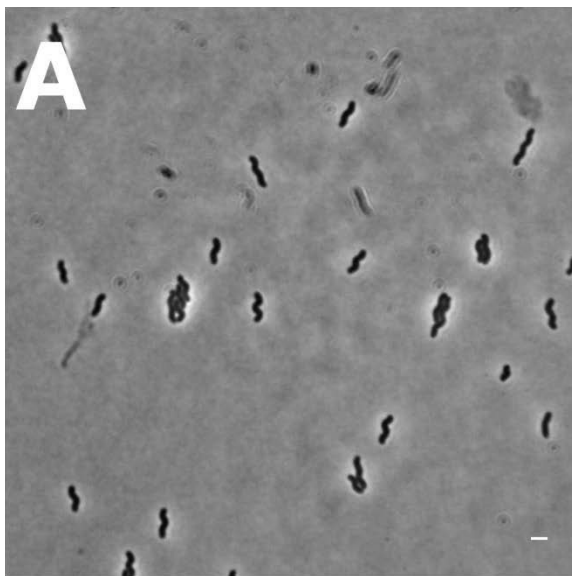
B128.7.13 WT



B128.7.13 Δ *dlp* Clone 1



B128.7.13 Δ *dlp* Clone 2



B128.7.13 Δ ::*dlp* Clone 2

