The development of proteomic techniques to study the Australian Paralysis Tick, *Ixodes holocyclus.*

The application of proteomic technology to an organism with poor bioinformatic information.

Matthew P Padula.

Doctor of Philosophy, 2008.

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I 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.

Signature of Student

_____

Acknowledgments

This work was started in February, 2000, performed on a part-time basis and submitted in November, 2008. During that time I was a laboratory manager for the Immunobiology Unit (2001), then one of the laboratory managers for the Institute for the Biotechnology of Infectious Diseases (IBID) (2002-2005) and finally a technical specialist for the Proteomics Technology Centre of Expertise (PTCE) (2006-present). For 18 months in the middle I was assisting in the relocation of the Faculty of Science from Gore Hill to the Broadway campus. As one would expect, over this considerable period of time, I met and had dealings with many, many people. Some of these deserve special mention on the following pages.

Associate Professor Kevin Broady: The man who took a chance on employing a 22 year old honours graduate with no experience in July 1997. I will be eternally grateful for the opportunity I was given and the $1 million worth of equipment I was given to learn my trade. However, he could have warned me away from ticks and onto snake venom instead. It is unlikely I would be in the position I am if he did warn me, so I'm glad he didn't.

Professor Bob Raison and members of the Immunobiology Unit: These people provided an environment conducive to performing quality science. I should especially thank Peter Hains, K.L. Sung, Jessica Chung, Darren Jones, Josh Moses, Susan Lemke, Joyce To, Blair Nield, Lin Bozduganova, and Roie Gorman.

The Members of IBID: Especially Mike Johnson, Scott Minns and June Treerat, without whose help I would have burned out long before this work got finished.

Associate Professor Ben Herbert: The most important man in this list, who rekindled my love of working in the lab and working on proteins. If Ben had not taken on the role of mentor that I desperately needed, this thesis would not exist. Nothing I can put into words will adequately describe the effect Ben has had on my life and my work. In

echoes of 1997, he handed me $1 million worth of equipment and gave me the means and the confidence to use it. Now I need to repay the faith by taking it to the next level of understanding. He is my boss, my mentor, my colleague, my equal and my mate. Now that this is over, I look forward to the return of scientific discussions in breaks between runs on downhill bikes. And finding a way to get him to read my work faster.

Cameron Hill: Cameron has taught me more about electrophoresis than I thought could be learnt. He also reminded me about the power of logic and its application to the scientific problem. Without his input, this work would not be of the quality that it is.

The PTCE alumni: All of the people who has used our lab and allowed me to gain more knowledge for my own organism by working on theirs. Special mention needs to go to our current group of students, Michael Medynskyj, Jerran Naidoo and Michael Sivell. You guys make coming to work fun.

The Mass Spectrometrists: The people who were readily willing to answer my dumb questions as I was learning how to use the complex beast that is my QSTAR Elite. Paul Haynes from Macquarie University, Mark Raftery from the Bioanalytical Mass Spectrometry Facility at the University of New South Wales, Michelle Colgrave and Jose Meza from Applied Biosystems. Special thanks go to my friends at the National Institute of Health in Bethesda, Maryland, Al Yergey, Nancy Vieira and Peter Backlund who not only put up with my endless questions, but also educated me in some of DC's wonders I wouldn't have bothered with.

The Extended Family: My parents, Hilda and Mike, who supported me both emotionally and financially because I loved what I was doing. My brother Anthony needs thanks for random discussions over beer that he may not think were important but the logic presented shaped some concepts here. My other parents, Olga and the late Milorad "Rado" Brusic. Rado put me to work as a steel fixer after my honours was

finished and reminded me that there were very hard ways to make a living and taught me a lot about taking things at face value. My “brother” Nikola and his family, Jo-Anne, Dylan and Lee for the many camping trips and caves that have been and are to come. The only “sister” I will ever have, Helena, who showed me that knowledge has been around for a really long time and you can learn anything with bit of passion and determination.

The Mountain Bikers: I thank anyone I have ever ridden a bike with. Scott “Spanky Ham” Minns for laps of Yellowmundee. Gaz Theiss and his wonderful wife Yvonne for being the awesome and generous friends that they are. The Southside Mafia, especially Mike, Bert, Nutty, MattO, Sarah, Jamie and Emerson. I’ll be down for a shuttle soon. Craig Flynn and the lads of Central Tablelands MTB Club, I will renew my membership this year. Kerry Moroz at Phantom Cycles for supporting my bike habit.

The Committee: My wife Venesa and my children Mikaela (The Bot) and Natasha (The Tooshk). Venesa pushed me when I didn’t want to be pushed and held me back when I needed to be held back. Rather than take my diatribes, she made me realise how dumb and unproductive it actually is. The cliché says that children remind you that there is something more important than work. My children make me realise that I have to keep learning so I can keep up with their endless and insightful questions about how the universe works.

Table of Contents.

Figures and tables.

Abbreviations.

1-D	One-dimensional
2-DGE	Two-dimensional gel electrophoresis
ASB-14	Amidosulfobetaine-14
BLAST	Basic local alignment search tool
CDS	Coding determining region
CID/CAD	Collisonally induced/activated disocciation
DTT	Dithiothreitol
ESI	Electrospray ionisation
EST	Expressed sequence tag
ETD	Electron transfer dissociation
IEF	Isoelectric focusing
IPG	Immobilised pH gradient
LC	Liquid chromatography
LC/MS/MS	Liquid chromatography coupled with tandem mass spectrometry
LDS	Lithium dodecyl sulphate
MALDI	Matrix Assisted Laser Desorption Ionisation
MCE	Multi-compartment electrolyser
MES	2-(N-morpholino)ethanesulfonic acid
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MudPIT	Multi-dimensional Protein Identification Technology
MW	Molecular weight
NCBI	National Centre for Biotechnology Information
NH_4HCO_3	Ammonium hydrogen carbonate
PBS	Phosphate buffered saline
PDB	Protein Data Bank
PIR	Protein Information Resource
PRF	Protein research foundation
PTM	Post-translational modification
QTOF	Hybrid Quadrupole Time-Of-Flight mass spectrometer
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Swiss-Prot	Swiss Institute of Bioinformatics protein database
TBP	Tributylphosphine
TFE	2,2,2 – Trifluoroethanol
Tris	Tris(hydroxymethyl)methylamine
UTC7	7M urea, 2M thiourea, 1% C7BzO

Abstract.

The Australian paralysis tick, *Ixodes holocyclus*, is representative of the majority of organisms studied in biology in that the bioinformatic information available (genome sequence, annotated coding regions and protein sequences) are far from complete. The study of well characterised model organisms has shown that proteomics and its associated technologies are able to isolate, identify and characterise individual protein isoforms at femto to attomole amounts of sample. With these model organisms, this can be achieved in either unpurified or partially purified samples (shotgun proteomics) or by high resolution separations using isoelectric fractionation and two-dimensional gel electrophoresis.

In a poorly characterised organism, this is not the case. The work presented in this thesis applies proteomic technologies to characterising the tick proteome in a hypothesis and non-hypothesis driven manner. In the non-hypothesis driven approaches, fractionation and separation methodologies were applied to determine which method or combination of methods provided the greatest number of protein identifications. The results of these studies showed that the resolution of protein isoforms provided by 2-DGE is invaluable for characterising proteins from *I. holocyclus*. This is because the homogenous protein spot can be excised from the gel and characterised by *de novo* sequencing of MS/MS spectra with the knowledge that all peptides are from the same protein. However, successful *de novo* sequencing is reliant on good quality MS/MS spectra, which is partly reliant on intensely stained gel spots, which is determined by the amount of sample loaded onto the gel. It is well documented and demonstrated in this study that overloading of 2-D gels with samples containing high abundance proteins, tick cytoskeletal proteins in this case, can cause spot resolution problems. Fractionation of the sample using a Multi-compartment Electrolyser and equalisation with Proteominer partially addresses this issue, but further refinement is necessary.

The optimised sample preparation methods were then applied in hypothesis driven experiments to characterise specific protein subtypes using Western blots and a novel fluorescent zymogram approach. The analysis identified a number of proteins that will need

further characterisation, using molecular biological and recombinant protein expression techniques, to determine their suitability as vaccine candidates.

1.1: Overview.

The Australian paralysis tick, *Ixodes holocyclus*, is of particular veterinary and agricultural importance on the eastern coast of Australia (Roberts, 1960). This importance is due not only to its ability to transmit infectious pathogens as the majority of the 820 tick species found worldwide do, but also because of the production of a potent neurotoxin in its saliva evident after the 4-5th day of feeding. This neurotoxin causes ascending flaccid paralysis, respiratory failure and death of the host (Albiston, 1968). A number of studies have tried to characterise the neurotoxin and produce a recombinant vaccine against it with limited success.

The original aim of the work presented in this thesis was to extend on previous work (Masina, 1999, Thurn, 1994) to purify and characterise the neurotoxin of *I. holocyclus* and produce recombinant vaccine candidates that would provide protective immunity to immunised animals. Following many technical issues and a lack of engorged ticks from which to isolate the toxin from, it was decided that the neurotoxin was not a good vaccine candidate due to the fact that the tick would still be able to feed for 4-5 days prior to the significant production of the antigen to which an immune response would be primed. A more appropriate antigen would be present at tick attachment and the onset of feeding. This would result in an immediate host immune response that would greatly shorten the attachment period and possibly reduce or prevent pathogen transmission. Unengorged female ticks, which are much easier to acquire and do not have the issue of contamination of massive quantities of host proteins, are the starting material to find and characterise these candidates. However, as will be demonstrated, the adaption of technology developed for use in human serum proteomics has useful application in the proteomics of engorged female ticks.

The work in this thesis uses methodologies and techniques that fall under the broad umbrella called proteomics. Proteomics is the study of the protein expression of the

genome of a particular organism. Nearly ten years has passed since the last published work on proteins from *I.holocyclus* (Masina and Broady, 1999) and technology has changed immensely. Nanoflow chromatography coupled to mass spectrometry has made it possible to identify and obtain peptide sequence information from femtomole amounts of protein and two-dimensional gel electrophoresis has unsurpassed resolving power of complex protein mixtures when performed by skilled operators. This work has utilised a number of novel and evolving technologies that have application to many blood feeding organisms other than ticks.

1.2: The Australian Paralysis Tick, *Ixodes holocyclus*.

Ticks are blood feeding arthropods and are second only to mosquitoes as the main vector for pathogenic organisms that affect vertebrates (Ribeiro, 1995). They belong to the class Arachnida, which includes spiders and scorpions, and the sub-class Acari, which includes mites. Ticks have evolved to require vertebrate blood for nutrition and egg development, but by evolving the mechanisms that allows uninterrupted blood feeding, ticks have had a co-evolutionary relationship with many pathogenic diseases (Andrade *et al*., 2005). However, ticks are not merely vectors for pathogens. Tick saliva is a potent pharmacologically active fluid that directly affects the host's haemostatic, inflammatory and immune responses (Ribeiro, 1995). These effects allow many days of uninterrupted feeding.

Ixodes holocyclus is native only to the eastern coast of Australia (Roberts, 1960). This tick is of veterinary and agricultural importance due to the morbidity and mortality of companion and commercial animals due to blood loss during tick feeding and the production of a neurotoxin in the tick's saliva around 4-5 days of feeding during rapid engorgement (Oxer and Ricardo, 1942). *I. holocyclus'* distribution is limited by its susceptibility to slight changes in environmental conditions, particularly humidity. The other main factor determining the paralysis tick's abundance is the presence of its natural host, the bandicoot. It has been observed in bandicoots and hyperimmune dogs, both of which are continually fed on by great numbers of paralysis ticks, that immunity can be acquired but this immunity can be lost over several months if the animal is kept tick free [http://www.peg.apc.org/~ullavet/tick.html]. These observations point to the viability of using vaccination as a strategy to prevent tick paralysis. Previous work with gluteraldehyde-treated, partially purified toxic fractions reinforces this observation (Stone *et al*., 1986) however it also showed the impracticality of using native toxin preparations due to the large numbers of ticks required and low yield of toxin. Neurotoxin production occurs during feeding by adult females in this three host tick (the lifecycle is shown figure 1.1), although infestations

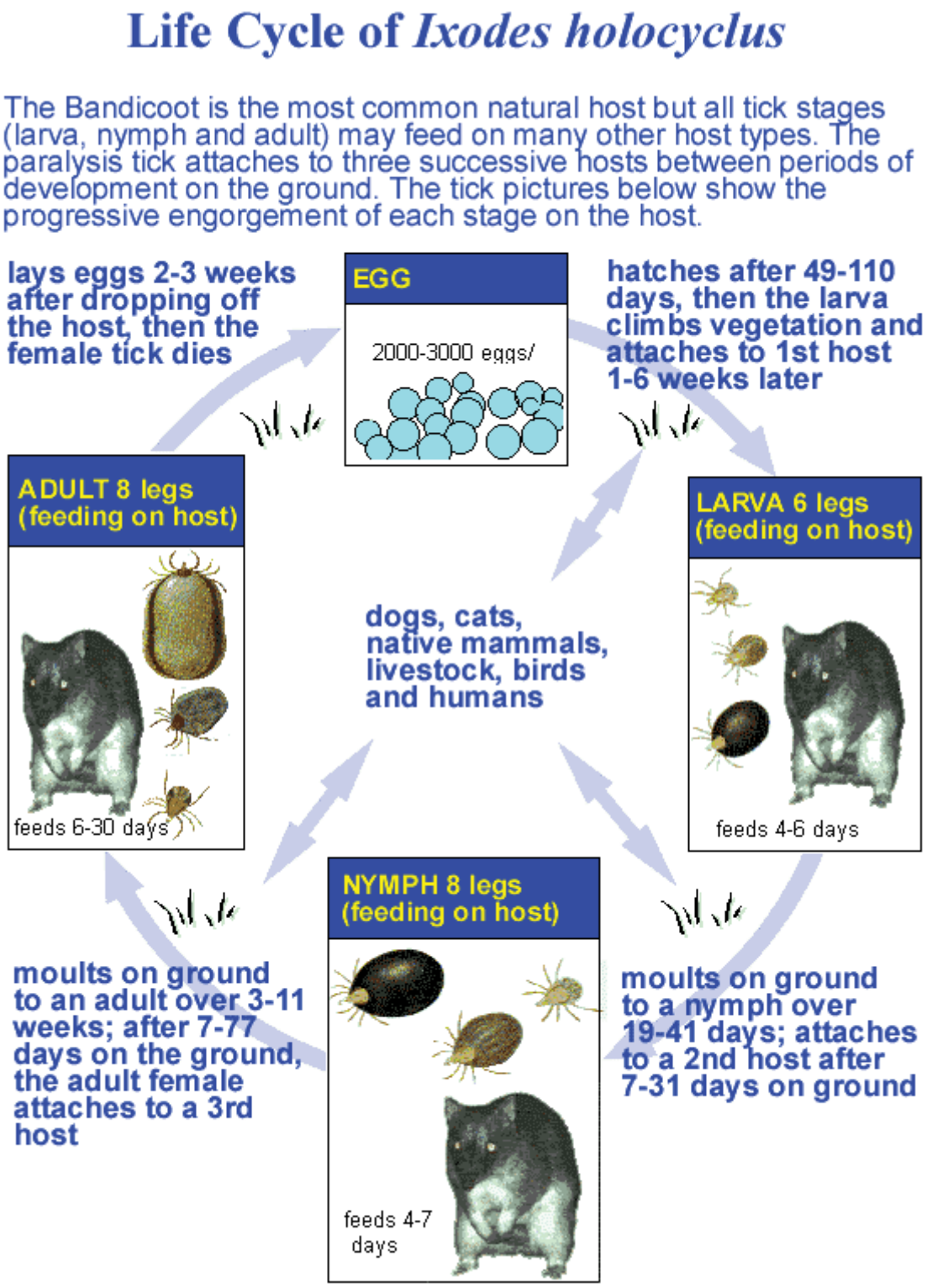

Figure 1.1: Diagrammatic representation of the lifecycle of the Australian paralysis tick, *Ixodes holocyclus* (http://pandora.nla.gov.au/pan/14045/20020920-0000/members.ozemail.com.au/_norbertf/index.html). This species of tick has a three host lifecycle where each feeding stage is followed by a moult before a new host is found.

of nymphs causing paralysis have been reported (Oxer and Ricardo, 1942), but only rarely while there are no reports of larvae causing paralysis. But as mentioned in the overview, developing a vaccine preventing paralysis but still allowing the tick to feed and possibly pass pathogens to the host may not be the best strategy.

Feeding begins by the tick lacerating the host's skin, tearing tissue and capillaries to create a feeding chamber. It is into this chamber that saliva is injected containing proteins, prostaglandins, nucleotides and nucleosides (Andrade *et al*., 2005). These molecules locally modify the physiology of the host, allowing the tick to remain attached and taking a blood meal with the side effect of producing an environment at the feeding chamber that allows the transmission and establishment of infectious pathogens. In *Ixodidae* ticks, feeding occurs in two stages (Cupp, 1991). The first, slow phase lasts up to four days with the female's body and organs developing for the final stage of rapid engorgement while the molecules necessary to control the host's response are also produced in this slow phase (Bowman *et al*., 1997). The second rapid engorgement phase lasts approximately 24 hours and is characterised by a 10-fold increase in weight of the tick. This increase greatly underestimates the volume of blood taken by the tick as the meal is concentrated by removal of fluid which is taken up by the salivary glands and returned to the host's circulation (Cupp, 1991; Beckman, 1989).

1.2.2: Host Haemostasis.

The vertebrate haemostatic system is composed of efficient and sophisticated processes whose main function is to stop blood loss through platelet aggregation, the blood coagulation cascade and vasoconstriction (Ribeiro, 1995). Thus to facilitate the taking of a blood meal, arthropod saliva contains anticoagulants, anti-platelet factors and vasodilators (Ribeiro, 1995; Stark and James 1996; Champagne, 2005) with numerous different molecules in each category. These activities have evolved independently in different tick species and upon comparison of the sialomes (cDNA

libraries) of different species this independent evolution is clear (Mans *et al*, 2008). Targeting tick anti-hemostatics maybe an effective vaccination strategy and they may also be useful therapeutics (Maritz-Olivier et al., 2007), but the independent evolution of the sialomes may prevent a single, all encompassing solution.

Arthropod blood feeders penetrate the skin of the host, lacerating capillaries and vessels to release blood. In the case of Ixodid ticks a feeding pool is created where the blood collects and is taken up through the hyperstome. Vasodilators in the saliva are thus essential to maintain the flow of blood and prevent vasoconstriction by activated platelets (Champagne, 2005). Vasconstriction can occur when arachidonic acid is released by activated platelets and converted to thromboxane A_2, a powerful platelet aggregating and vasoconstricting factor (Riberio, 1987). Serotonin is also released which acts in concert with thromboxane A_2 (Weigelt *et al*., 1979). In the saliva of hard ticks, prostaglandins E_2, $F_{2\alpha}$ and prostacyclin dilate host blood vessels, antagonising thromboxane A_2 (Ribeiro *et al* 1995; (Bowman et al., 1996). These prostaglandins occur at several fold higher concentrations than necessary to dilate vessels (Shimokawa *et al.*, 1988, Williams, 1978). Arachidonic acid is necessary for the production of prostaglandins, a molecule ticks are unable to synthesise and must obtain from the blood meal (Bowman *et al*., 1995, Madden *et al*., 1996). The vasodilators from soft ticks are yet to be characterised, but apparently they are not prostaglandins (Astigarraga et al., 1997).

Removal of vasoconstrictors is another way of achieving vasodilation and a major protein family in both soft and hard ticks involved in their removal are the lipocalins. In soft ticks, lipocalins function as platelet aggregation inhibitors (Waxman and Connelly, 1993) as well as anti-complement factors (Nunn *et al*., 2005) and toxins (Mans et al., 2002d). In hard ticks, lipocalins scavenging of histamine and serotonin have been described (Paesen *et al*., 2000, Sangamnatdej *et al*., 2002). Interestingly, the lipocalins

of different ticks reflect their host specificity. In *Rhipicephalus appendiculatus* three histamine binding proteins are produced to counteract the inflammatory mediator (histadine) of its host, cattle (Paesen *et al*., 2000). In contrast, similar proteins in *Dermacentor reticulatus* has the histamine-binding site modified to a high affinity site for serotonin, which is the mediator of inflammation in its rodent hosts, which *R.appendiculatus* never feed upon (Champagne, 2005, Sangamnatdej *et al.*, 2002).

Platelet aggregation (Figure 1.1) is the host's first mechanism to prevent blood loss (Andrade *et al*., 2005). Following activation by a stimulus such as ADP, collagen or thrombin, platelets aggregate, promote clotting and release vasoconstrictors to form a platelet plug (Champagne, 2005). Ticks and other blood feeders produce a number of molecules to inhibit different parts of this process. The most widespread inhibitors amongst blood feeders are salivary apyrase or ATP/ADP diphosphohydrolase. ADP has a key role in platelet activation, and it is not surprising that most blood feeding arthropods possess salivary apyrase that hydrolyses ATP and ADP to AMP and orthophosphate thus preventing ADP's activation role (Andrade *et al*., 2005). AMP also has vasodilatory activity (Collis, 1989) and thus it's production contributes to a successful blood meal. Prostacyclin in tick saliva inhibits platelet aggregation and is present at concentrations far higher than necessary to cause platelets to disaggregate (Radomski *et al*., 1987, Ribeiro *et al*., 1988). Tick saliva can also prevent platelet aggregation by inhibiting binding to fibrinogen and other ligands necessary for aggregation. *D. variabilis* produces a 5kDa peptide called Variabilin that inhibits ADP-stimulated platelet aggregation (Wang *et al*., 1996). Savignygrin, from *Ornithodoros savignyi*, inhibits platelet aggregation without inhibiting platelet activation or shape change (Mans *et al.*, 2002c) while Moubatin isolated from *O.moubata* inhibits collagen-stimulated platelet aggregation (Waxman and Connolly, 1993). The inhibition of platelet aggregation also prevents the recruiting of fibroblasts to the feeding site (Kramer *et al.*, 2008). The saliva of *D. variabilis* has been demonstrated in vitro to suppress fibroblast monolayer repair and inhibiting fibroblast migration while having

no effect on cell number. The possible mechanism of this inhibition is the saliva's ability to suppress ERK activity, a critical part of the signalling pathway involved in regulating cell motility in almost all cell types involved in the dermal response to injury (Kramer *et al.*, 2008).

The blood coagulation cascade (Figure 1.2) in vertebrates is the final phase of the haemostatic response and is launched by various mechanisms after injury to blood vessels. The result is the production of active thrombin that cleaves fibrinogen to fibrin which polymerises to form a clot and provide rigidity to the platelet plug. Salivary anti-coagulants from blood feeders target specific proteases and complexes in the cascade blocking or delaying clot formation until the blood meal is finished (Riberio, 1987). Hard ticks, due to their long feeding times are exposed to an environment where coagulation is imperative to prevent host blood loss during feeding and thus have evolved redundancy to inhibit coagulation at a number of steps. Thrombin, being the ultimate protease in the cascade, is an attractive and obligatory target for blood feeders to inhibit (Macedo-Ribeiro *et al*., 2008) along with Factor Xa (Champagne, 2005). Both hard and soft ticks share proteins with common folds such as the basic pancreatic trypsin inhibitor/Kunitz family (BPTI/Kunitz) which are thrombin, factor Xa and platelet aggregation inhibitors (Mans *et al*, 2008). Boophilin, a protease inhibitor isolated from whole engorged *R. microplus*, inhibits thrombin, but also inhibits serine proteases such as trypsin and plasmin (Macedo-Ribeiro *et al*, 2008). Boophilin contains two Kunitz domains and other salivary proteins containing Kunitz domains have been found in *I. scapularis* (Ixolaris; (Francischetti *et al.*, 2002, Valenzuela *et al.*, 2002)), *I. pacificus* (Francischetti *et al*., 2005), *Amblyomma hebraeum* (Lai *et al*., 2004), *Ornithodoros moubata* (Orthordorin; van de Locht *et al*, 1996) and *O. savignyi* (Savignin) (Mans *et al*., 2002a) indicating that inhibitors from this family may play a general role in anti-haemostasis (Macedo-Ribeiro *et al*, 2008). The Factor Xa inhibitor, tick anticoagulant peptide (TAP), found in both *O.moubata* and *O. savignyi* are also members of the Kunitz family (Gaspar *et al*., 1996, Waxman *et al*., 1990). Comparison

of anticoagulant proteins in hard and soft ticks has provided more evidence that blood feeding evolved independently and the timing of the separation of the two tick families coincides with the radiation of birds and mammals. It is likely that speciation on different hosts drove the separate evolution of blood feeding (Mans *et al*., 2002b).

Some tick anticoagulants inhibit the formation of complexes necessary in the coagulation cascade. Ixolaris is thought to bind Factor X or Xa via its second Kunitz domain and when this complexes with Factor VIIa/TF complex, Ixolaris' first Kunitz domain occupies the Factor VIIa active site (Francischetti *et al*., 2002). A similar protein is found in *R. appendiculatus* (Limo *et al.*, 1991) and inhibition of Factor V and VII in *D. andersoni* is expected to function through the same mechanism (Gordon and Allen, 1991). *I. scapularis* saliva also contains a metalloproteinase that acts against gelatin, fibrinogen, and fibronectin and thus can dissolve fibrin clots and thus reverse coagulation (Francischetti *et al*., 2003). In contrast to the mentioned anticoagulants, calreticulin can inhibit a number of coagulation factors by chelating calcium (Dai *et al.*, 1997).

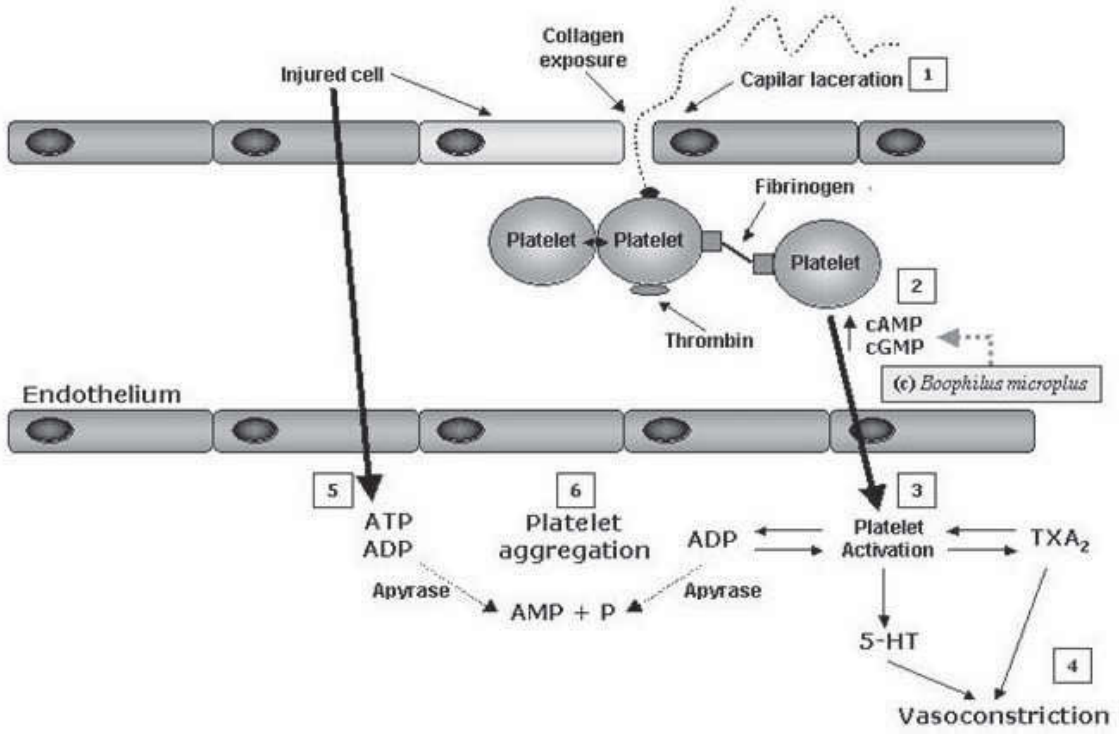

Figure 1.2: Vector's saliva acting on platelet activation and aggregation. (1) Blood feeding vectors induce vessel laceration and tissue injury resulting in collagen exposure when probing for a blood meal. (2) Thus, platelets aggregate, promoting clotting, and release of vasoconstrictor mediators promoting hemostasis. Blood feeders can inhibit platelet aggregation by preventing fibrinogen, thrombin or cAMP/cGMP stimulation (*Boophilus (Rhipicephalus) microplus*). (3) Platelet activation and degranulation also occur after thromboxane A2 that results in vasoconstrictor response and (4) the NO present within bug's saliva can prevent haemostatic effect.(5) They can also bind to ADP or (6) Prevent the action of ADP through salivary apyrase to prevent platelet aggregation. Adapted from Andrade *et al*, 2005.

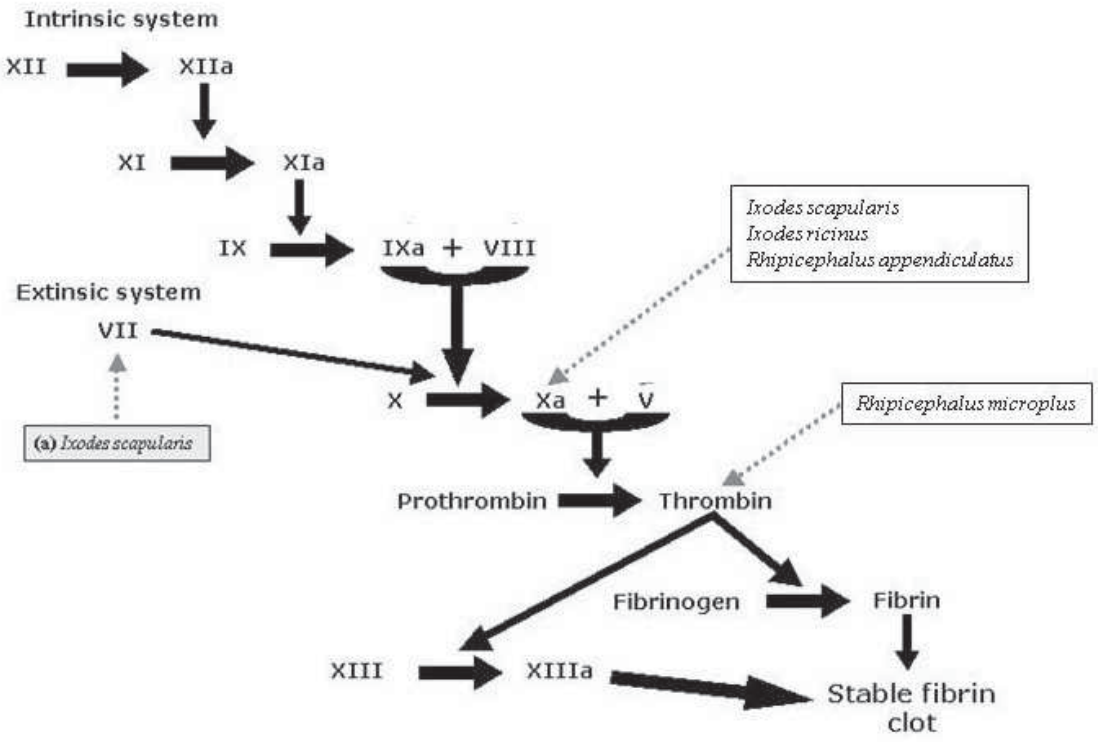

Figure 1.3: Blood-coagulation cascade (intrinsic and extrinsic system) activated in response to tissue injury. The blood-coagulation cascade is activated after blood vessels injury resulting in the production of active thrombin, which cleaves fibrinogen to fibrin that polymerizes forming a stable clot blocking blood loss. Salivary anticoagulants from blood feeding arthropods inhibit specific targets of the coagulation cascade. They target components such as factor IXa; VIII; Xa (*Ixodes scapularis, Ixodes ricinus* and *Rhipicephalus appendiculatus*); V, VII (*Ixodes scapularis*) and thrombin *(Boophilus microplus*) resulting in inhibition or delayed blood-thrombin (*Boophilus microplus*) and coagulation response. Adapted from Andrade *et al*, 2005.

One group of proteins that has attracted attention as a potential tick vaccine candidate are the serpins or serine proteinase inhibitors. In humans and other mammalian hosts, serpins have a central role in many physiological processes such as coagulation, fibrinolysis, fertilisation, development, malignancy, neuromuscular patterning and inflammation (Rubin, 1996). Parasite and arthropod produced serpins can therefore impair these same physiological functions, specifically the blood coagulation cascade, the inflammatory cascade and complement activation (Armstrong, 2001). Serpins inhibitory function occurs when the specific protease cleaves the peptide backbone of the serpin, but in the process the serpin becomes covalently attached. This results in a conformational change in the protease making it catalytically inoperative (Huntington et al., 2000). Serpins have been found in a range of blood feeding organisms such as on the surface of *Schistosoma* (Blanton *et al.*, 1994, Modha *et al.*, 1996) where it is speculated to prevent clot formation around the parasite and provide protection from activated neutrophils (Modha *et al*., 1996). In *Schistosoma* it has also been shown that serine proteases produced by the parasite are non-immunogenic due to the formation of protease-serpin complexes with host-derived serpins leading to speculation that recombinant inactive *Schistosoma* serine proteases could be an effective vaccine candidate (Modha *et al.*, 1996). Recent results (Yan *et al.*, 2005) have shown that a recombinant serpin from *S. japonicum* gives moderate protection in immunised mice against *S. japonicum* infection. Serpins have also been found in the mosquito *Aedes aegypti* (Stark and James, 1998) and arthropod serpins are involved in the regulation of anti-parasite defence systems (Levashina et al., 1999). Ticks are likely to use serpins to disrupt defensive host processes (Sugino *et al*., 2003).

1.2.3: Host Immunity.

The long length of feeding by ticks means that many tick antigens are exposed to the host immune system for a considerable time. Blood feeding arthropods have thus needed to develop immunomodulatory molecules to prevent the host becoming sensitized to salivary molecules facilitating the blood meal (Gillespie *et al*., 2000) (Figure 1.3). Innate immunity represents the first line of defence but lacks immunologic

memory and involves the complement cascade, acute phase proteins, inflammatory inductors, granulocytes, mast cells, dendritic cells, macrophages and NK cells (Andrade *et al*., 2005) and thus it is not surprising that tick saliva contains molecules that modulate many of these. Saliva of *I. dammini* antagonises anaphylatoxin and bradykinin (Ribeiro and Spielman, 1986) and inhibits C3a release and C3b deposition (Ribeiro, 1987) disrupting the complement cascade. Hard ticks, such as *I.scapularis*, have kininases to inhibit bradykinin (Ribeiro and Mather, 1998) and histamine-binding proteins to prevent itch and thus grooming by the host (Paesen *et al*., 1999). Tick saliva has also been shown to induce immune suppression of innate immune cells. *I. dammini* salivary homogenates inhibit neutrophil function (Ribeiro *et al*., 1990), *Dermacentor reticulatus* salivary gland extract decreases NK cell activity (Kubes *et al*., 2002) and *I.ricinus* saliva inhibits dendritic cell migration, maturation and function while promoting Th2 responses (Skallova *et al.*, 2008).

Ticks and other blood feeders have also developed the means to suppress the host's acquired immune response and the identification of these molecules is crucial in the development of any tick vaccine. Ticks have a systematic immunosuppressive effect on the host (Andrade *et al*., 2005). Salivary PGE_2 is thought to be responsible for suppressing exposed lymphocytes in vitro and lymphocytes isolated from tick infested experimental animals demonstrated a greatly reduced response to mitogens after culturing in vitro (Inokuma *et al*. 1994, Ramachandra and Wikel 1992, Ribeiro *et al*. 1985). The saliva of *I.scapularis* possesses a protein that binds to IL-2 (Gillespie *et al.*, 2001). This prevents activation of T-cells, B-cells, macrophages and NK cells, all of whom possess receptors for IL-2, demonstrating the importance of preventing its binding. Salivary gland extracts of *R. appendiculatus* have been shown to modulate cytokines, reducing the expression of IFN-α, TNF-α, IL-1α, IL-1β, IL-5, IL-6, IL-7 and IL-8 by LPS-stimulated human peripheral blood leukocytes (Fuchsberger *et al.*, 1995). This indicates that the tick is attempting to shift the host immune response to the Th2 pathway which would favour blood feeding and affect the expression of host acquired

resistance (Schoeler *et al.*, 1999). A number of studies have concluded that exposing lymphocytes to tick saliva favours the development of a Th2 response characterised by high levels of IL-4 and IL-10 and low levels of IFN-γ and IL-12 (Ferreira and Silva, 1999, Mejri *et al.*, 2002, Schoeler *et al.*, 1999). This is advantageous to the feeding tick as Th1-mediated inflammation of the host skin may cause the tick to be removed by grooming (Alexander, 1986). Parasites such as *Borrelia burgdorferi* also induce strong Th2 responses in susceptible mice while downregulating Th1 cytokines suggesting that immunisation targeted at stopping both tick and pathogen control of these responses would be effective in stopping pathogen transmission (Zeidner *et al.*, 1997) and possibly tick attachment. Saliva from *D. andersoni* contains molecules that suppress T-cell proliferation (Bergman et al., 1995) and reduce the expression of adhesion molecules on the lymphocyte surface (Macaluso and Wikel, 2001), retarding migration and modifying the host's immune cells at the feeding site, possibly through the same ERK signalling mechanism found to inhibit fibroblast migration.

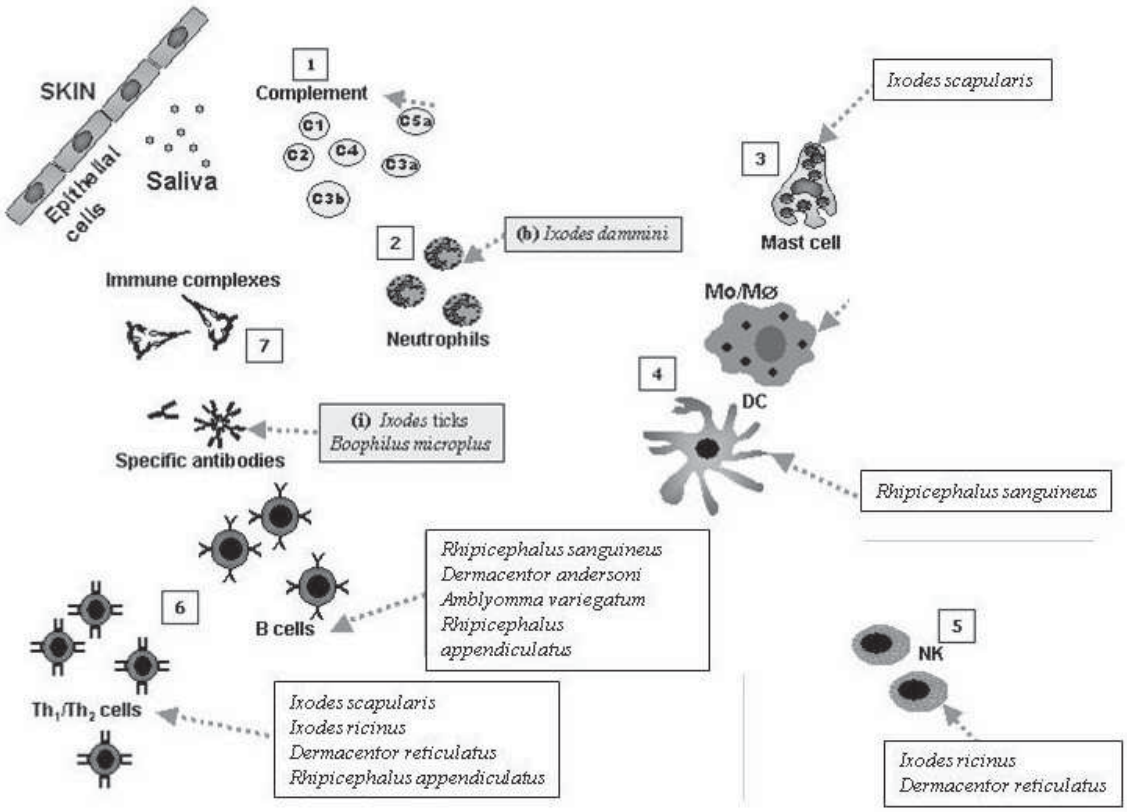

Figure 1.4: Host immune response is modified by arthropod's saliva. Salivary molecules can act on different effector cells and mediators of the immune system: (1) The complement system: inhibition of complement release of vasoactive mediators and cell activation in both classical and alternative pathways; (2) Neutrophils: inhibition of neutrophil function (*Ixodes dammini*); (3) Mast cells: reduction of mast cell degranulation and release of inflammatory mediators (*Ixodes scapularis*); (4) Antigen Presenting Cells: macrophages: inhibition of macrophage activation and dendritic cells: reduction of dendritic cell differentiation, maturation and cytokine production (*Rhipicephalus sanguineus*); (5) NK cells: reduction of NK cell cytotoxicity (*Dermacentor reticulatus* and *Ixodes ricinus*); (6) Lymphocytes: B cells: inhibition of cell proliferation and modulation of immunoglobulin production (*Rhipicephalus sanguineus, Rhipicephalus appendiculatus, Amblyomma variegatum* and *Dermacentor andersoni*) and T cells: modulation of cytokine production, reduced proliferative response and impaired leukocyte traffic (*Ixodes scapularis, Ixodes ricinus, Dermacentor reticulatus, and Rhipicephalus appendiculatus*) and (7) Antibodies and Immune Complexes: modification of immunoglobulin responses profile (*Boophilus microplus* and Ixodes ticks). Adapted from Andrade *et al*, 2005).

Even though ticks can modulate and suppress the acquired immune response of the host, natural and experimental hosts can still acquire resistance to tick feeding and tick salivary molecules. There is much anecdotal evidence for bandicoots and dogs repeatedly exposed to *I. holocyclus* being immune to holocyclotoxin and the only current treatment for paralysis caused by *I. holocyclus* is injection of hyperimmune sera from dogs continually fed on by *I. holocyclus*. In other, non-neurotoxin producing ticks, this resistance is seen by reduced tick engorgement, decreased numbers and viability of ova, impaired moulting, and death of feeding ticks (Wikel 1996, 1999). Ticks in response have evolved mechanisms to minimise the effect of host resistance. Ixodid ticks produce a unique family of immunoglobulin binding proteins (Wang and Nuttall 1995a,b) discovered after the observation that ticks excrete host immunoglobulins in their saliva during feeding and the immunoglobulins retain their binding capacity (Wang and Nuttall, 1994). *R. microplus* has been shown to modulate the isotype of host antibody response (Kashino *et al.*, 2005).

As mentioned previously, the modulation of the host immune system by tick salivary components provides an environment conducive for pathogen transmission. Ticks are vectors for Lyme disease, rickettsial diseases, babesiosis, ehrlichiosis and encephalitis viruses (Andrade *et al*., 2005). However, it is thought that relatively subtle differences in cytokine expression following exposure to tick saliva and pathogens may explain different host susceptibility to pathogen transmission (Andrade *et al*., 2005). *R. appendiculatus* salivary gland extracts have been shown to enhance the uptake of *Theileria* parva sporozoites into lymphocytes, macrophages and afferent lymph veiled cell (Shaw *et al.*, 1993). In *I.scapularis*, the immunosuppressive protein Salp15 has been identified as a saliva activated transmission factor that helps *Borrelia burgdorferi* establish in mammalian hosts (Ramamoorthi *et al*., 2005), but work has shown that it also suppresses asthma in mice (Paveglio *et al.*, 2007) and may have use in allogenic transplantation (Motameni *et al.*, 2004). Mice and guinea pigs infested with pathogen

free *I. scapularis* develop acquired resistance to *Borrelia burgdorferi* infection when challenged with infected ticks (Nazario *et al.*, 1998, Wikel *et al.*, 1997).

1.2.4: The problems of blood feeding.

One area of possible tick control that has received less attention is the digestion of the blood meal and the mechanisms employed by blood feeders to remove the toxic products such as heme which can reach concentrations of 150mg/mL in engorged female hard ticks (Graca-Souza *et al*., 2006). The tick blood meal is an extremely difficult sample to analyse due to the massive concentrations of host proteins, but the use of equalizer technology (Boschetti *et al.*, 2007) and fractionation techniques such as the MCE are making the blood meal a more easily analysed sample, as presented in latter chapters of this thesis. Anti-malarial drugs such as Chloroquine act on Plasmodium by preventing heme aggregation into hemozoin (Goldberg et al., 1990) and heme can promote the formation of free radicals leading to oxidation of lipids (Gutteridge and Smith, 1988), proteins (Aft and Mueller, 1984) and DNA (Aft and Mueller, 1983). Oxidative damage to phospholipid membranes occurs at a heme concentration of 100µM and blood heme is 10mM implying mid gut cells in blood feeders are exposed to toxic levels of heme in vivo (Graca-Souza *et al*., 2006). Hard ticks deal with heme in a different way to mosquitoes, accumulating the heme in a non-crystalline aggregate (Lara *et al*, 2003) after digesting hemoglobin intracellularly rather than in an extracellular environment as employed by other arthropods (Mendiola *et al*., 1996). After being taken up by receptor-mediated endocytosis, the hemoglobin is directed to specific acidic vesicles for digestion before the heme is moved to the cytosol and then the hemosome where it is packed into an organized aggregate (Lara et al., 2003 & 2005). The receptor mediating hemoglobin endocytosis maybe a vaccine target, preventing hemoglobin uptake for digestion although this may not release heme and cause toxicity.

Citelli *et al* (2007) have found that inhibiting catalase, part of the mechanism that deals with heme-produced free radicals, in female *R. microplus* increases levels of H_2O_2 in the gut impairing heme detoxification and results in diminished lifespan and lower egg-laying rates. This extended to ticks feeding on calves injected with catalase inhibitors having diminished lifespan, oviposition and engorgement. This tick catalase maybe an effective vaccine candidate. The catalase being inhibited could be from the tick or the host blood meal and evidence from *Plasmodium falciparum* has shown that the parasite using antioxidant defenses taken from the blood meal (Becker *et al*., 2004). Evidence also exists for the catalase being of tick origin as catalase activity has been found to increase before initiation of feeding in *R. microplus* (Freitas *et al*., 2007). An alternative to catalase-based removal of H_2O_2 through free heme is the thioredoxin/thioredoxin reductase system (TrxR/Trx) which has been shown to be upregulated in *Anopheles aegypti* after a blood meal (Sanders *et al*., 2003).

In addition to heme sequestration in the hemosome, hard ticks also employ heme-binding proteins such as the Hemelipoprotein (HeLp) found in *R. microplus* ticks (Maya-Monteiro *et al*., 2000) with homologs found in other ticks such as *O. parkeri* (Gudderra *et al.*, 2001) and *D. variabilis* (Donohue *et al.*, 2008). HeLp has been shown to have a preventative anti-oxidant role and reduce heme promoted lipid peroxidation and also functions as an inter organ heme transporter (Maya-Monteiro *et al*., 2000). This function is important as *R. microplus* does not possess a heme synthesis pathway and relies on the blood meal for the heme it requires (Braz *et al*., 1999). Vitellogenin (Vg) is the only haemolymph storage protein studied in any great detail (Thompson *et al.*, 2007). Following its production in the fat body, Vg is released into the haemolymph, transporting heme to the developing eggs and being used as a nutritional source for the developing embryos (Gudderra *et al.*, 2002). The disruption of the aspartic endopeptidases employed to degrade vitellogenin, vitellin, another hemeprotein found in tick egg yolk, and host haemoglobin has been suggested as a method of tick

control (Boldbaatar *et al.*, 2006, Nascimento-Silva *et al.*, 2008, Pohl *et al.*, 2008, Sojka *et al.*, 2008).

1.2.5: Tick bioinformatic resources.

In addition to the evolution that has occurred in both electrophoresis and mass spectrometry, both of which will be described in depth in chapter two, that allow the researcher to analyse smaller and smaller amounts of sample with increased dynamic range of protein concentration, and revolution has occurred in bioinformatics over the last decade. The complete sequencing of many genomes, such as human, mouse, model bacteria and yeast to name a few, and the determination of the open reading frames (ORF's) within these genomes is the essential component that allows a researcher to search data generated from proteins and peptides fragmented by tandem mass spectrometry and identify homologous proteins. The integration of technology and bioinformatics allows the tackling of large biological problems as never before, where previously one molecule was studied at a time (Williams, 1999). However, in bioinformatics not all organisms are created equally.

Unlike human, mouse and many commonly used experimental organisms, ticks have very little gene and protein sequence data available to search data generated from proteomic projects. At the time of writing, the NCBI databases contained nearly 4950 protein entries, 1 520 036 gene sequences and 298289 EST's for Acari and 2742 proteins, 1 156 020 genes and 281 564 EST's for Ixodida. The progression of the *I. scapularis* genome sequencing project (Pagel Van Zee *et al* 2007) initiated in 2004 (Hill and Wikel, 2005) has contributed 1 512 001 nucleotide sequences to the dataset available to be searched. However many of these sequences are listed as clones whose sequencing is still in progress or as genomic scaffold with no annotation of what the gene codes for. A search of these 1 512 001 sequences for those entries containing a CDS, or coding region that has been translated to protein sequence, shortens the list of entries to 693. This is significant because the NCBI non-redundant database (NCBInr) is

comprised of sequences compiled from GenBank CDS translations and the PIR, Swiss-Prot, PRF, and PDB protein databases. Thus only 693 of the 1 512 001 *I. scapularis* nucleotide sequences are included in NCBInr. The remaining sequences need to have their CDS determined to be included in NCBInr or a six frame translation to be performed, creating an extremely large dataset to be searched. It will be a considerable amount of time before the *I. scapularis* genome will be of use in tick proteomic projects.

Expressed Sequence Tag (EST) libraries are available for *R. microplus* (BmiGI), *A. variegatum* (AvGI), *R. appendiculatus* (RaGI) and *I. scapularis* (IsGI; Table 1.1) and these are able to be searched by the Peaks ProteinID algorithm (Bioinformatic solutions). ESTs or single pass sequencing of cDNA ends is an inexpensive way of obtaining genome and transcriptome data when compared to whole genome sequencing (Guerrero *et al.*, 2005). Guerrero *et al* (2005) estimated that BmiGI contains from 35% to 59% of the protein encoding portion of the *R. microplus* genome, assuming each of the BmiGI's 8270 unique sequences represents a unique gene. The result of this is at least 41% of the *R. microplus* genome is unrepresented and thus unsearchable. The latest release of BmiGI contains 9851 unique sequences thus representing 41-70% of the genome. This still leaves at least 30% unsearchable.

High throughput studies have been completed for the salivary glands of the American tick species *I. scapularis* (Ribeiro *et al*., 2006), *I. pacificus* (Francischetti *et al*., 2005) and *D. andersoni* (Alarcon-Chaidez et al., 2007), the tropical tick *Amblyomma variegatum* (Nene *et al*., 2002), the soft ticks. *O. parkeri* (Francischetti *et al.*, 2008) and *Argas monolakensis* (Mans *et al.*, 2008) and the castor bean tick *I. ricinus* (Chmelar *et al.*, 2008). These salivary transcriptomes or sialomes consist mainly of ESTs. As well as adding to the dataset able to be searched by MS/MS data, sialomes reveal much about the evolution of blood feeding. The datasets point to a large expansion of genes coding for a number of protein families as ticks diverged through their evolution (Francischetti

et al., 2008). For example, in *I. scapularis* and *I. pacificus,* several hundred putative secreted peptides belong to nearly a dozen gene families (Francischetti *et al.*, 2005, Ribeiro *et al.*, 2006). What all these salivary transcriptome studies have in common is that the vast majority of the identified proteins have no confirmed or known function (Francischetti *et al.*, 2008). The study of the *I.ricinus* sialome is unique as it uses four different time points in feeding to reveal the remarkable changes that occur in the expression of salivary proteins after attachment (Chmelar *et al*, 2008).

Gene indice name	Release version/date	Total ESTs	Tentative Consensus sequences	Singleton ESTs	Singleton ETs
A. variegatum	1.1 – 6th May, 2008	2152	490	1661	1
R. microplus (BmiGI)	2.1 – 3rd June, 2008	14586	9851	4696	39
R. appendiculatus	2.1 - 14th July, 2008	7583	2642	4917	24
I. scapularis	3.0 – 25th June, 2008	38392	20932	17437	23

Table 1.1 - Current composition and release dates of Expressed Sequence Tag (EST) databases available for ticks. Databases are hosted at compbio.dfci.harvard.edu/tgi/. The following information describing terms and assembly is from The Gene Indices website (compbio.dfci.harvard.edu/tgi/).

Terminology.

Expressed Sequence Tags (EST) are partial, single-pass sequences from either end of a cDNA clone. The EST strategy was developed to allow rapid identification of expressed genes by sequence analysis.

Tentative Consensus sequences are created by assembling ESTs into virtual transcripts. In some cases, TCs contain full or partial cDNA sequences (ETs) obtained by classical methods. TCs contain information on the source library and abundance of ESTs and in many cases represent full-length transcripts. Alternative splice forms are built into separate TCs. TCs are actual assemblies, with a consensus sequence, and not simply clusters of overlapping sequences.

Singleton ESTs are ESTs that are not contained in an assembly. These ESTs went through the assembly process but did not meet the match criteria to be assembled with any other EST in the collection of ESTs and other GenBank sequences used to create the consensus sequences for a particular Gene Index.
ETs represent mature transcripts. The ETs are curated for nomenclature and links have been made to related accessions. Sequences were downloaded from Genbank (mRNAs and sequences derived from genomic sequences). RefSeq divisions of GenBank (both known and model) were also downloaded and parsed. Where available, 5' and 3' non-coding regions were included. Alternative splice forms of genes are explicitly represented.

1.3.1: Vaccines against ticks.

Salivary molecules are attractive vaccine targets to stop both tick blood feeding and the transmission of parasitic diseases carried by ticks. However, the only successful tick vaccine to date is based on the midgut membrane antigen Bm86 from *R. microplus* which was commercialized in 1994 and has been used since in the field in Cuba, Australia, Mexico, Columbia and Brazil (de la Fuente et al., 2007). Bm86 has had a relatively minor effect on tick control efforts due to a number of factors, both commercial and technical. The most important technical problems are the inefficacy of Bm86-based vaccines against some ticks species (de la Fuente *et al.*, 2000, De Vos *et al.*, 2001) and incorrect use of the vaccine in an integrated control strategy (de la Fuente *et al.*, 2007). Since the introduction of Bm86-based vaccines, no new vaccines have appeared on the market (Nuttall *et al*., 2006) but many candidates have been evaluated (Table 1.2). Of these, only 64TRPs is being evaluated in cattle trials and as a transmission blocking vaccine against tick-borne encephalitis virus (Nuttall *et al*., 2006). Thus there is a need to find new potential candidates for broad spectrum vaccines controlling ticks and the disease they transmit.

1.3.2: Tick allergens.

Ticks, in common with other arthropods, can cause IgE-mediated allergic responses in susceptible individuals (Beaudouin *et al.*, 1997, Dorey, 1991, Stone *et al.*, 1989). Thus, attachment of *I. holocyclus* to humans can cause sensitisation and allergic reactions after repeated exposure ranging from localised skin irritation to life threatening anaphylaxis. *I. holocyclus* specific IgE was detected in human sera in the late 1980's (Dorey, 1991, Gauci *et al.*, 1988a, Gauci *et al.*, 1988b), but the actual proteins that the IgE reacts to has not been determined. This is mainly because the tools, such as Edman peptide sequencing and tandem mass spectrometry, were not available to the authors to determine the identity of the proteins when the previous studies were published. There have been no publications on the allergens from

Calreticulin	Jaworski *et al*, 1995; Ferreira *et al*, 2002
Immunoglobulin-binding protein	Wang and Nuttall, 1999
Histamine-binding protein	Paesen *et al*, 1999
Collagen-like protein	Mulenga *et al*, 1999b
Salivary protein	Tsuda *et al*, 2001
Cement protein	Bishop *et al*., 2002
64TRPs	Trimnell *et al*., 2005, Trimnell *et al*., 2002
Vitellin	Tellam *et al*., 2002
Vitelin-Degrading Cysteine Endopeptidase	Seixas *et al*, 2008
Yolk Pro-Cathepsin	Leal *et al*, 2006
Recombinant paramyosin	Ferreira *et al*., 2002a
Serine proteinase inhibitors	Imamura *et al.*, 2005, Sugino *et al.*, 2003
Male gonadal protein	Weiss and Kaufman, 2004
Tropinin I-like protein	Myung-Jo, 2005
Tick cDNA sequences	Almazan *et al*., 2005

Table 1.2 - Tick proteins evaluated as possible vaccine candidates.

I. holocyclus for almost 20 years. The human allergic sera used in the work of Dorey (1991) to characterise the allergens of *I. holocyclus* is still available for use in this work, as is a hyperimmune dog sera that is the current treatment for tick paralysis. Repeating Dorey's immunoblot work with 2-DGE and tandem mass spectrometry to identify the allergens is an aim of this project as those allergens have potential as vaccine candidates. In addition, identification of these allergens will aid in understanding the mechanism of allergic reactions and improve diagnosis and therapy.

1.4: Proteomics and ticks.

The use of proteomics to study parasitic arthropods has been limited for two main reasons. Firstly, the scarce amount of material available for proteomic methodologies. In this work, the resources were not available to establish a laboratory tick colony or artificial in vitro feeding assays as reported by Krober and Geurin (2007) and the twelve month life cycle from nymph to adult, along with the difficulties of receiving ethics approval for a tick colony on animals, meant colony establishment could not happen within the constraints of a PhD candidature. Adult ticks were collected in the field by professional tick collectors. This issue aside, conventional proteomics has generally been seen by the tick research community as not sensitive enough and prone to excessive sample losses (Valenzuela, 2002). Secondly, many researchers, both within the parasitology field and in the wider scientific community believe that the main proteomics separation tool, two dimensional gel electrophoresis, is too reliant on operator skill and often produces poor quality results. In reality, two-dimensional gels provide the highest resolution of a particular proteome and enable a frozen-in-time view of the proteome that includes the post-translationally modified (PTM) and processed versions of proteins (Herbert *et al*, 2007). The separation and analysis of post-translationally modified forms of proteins should be a priority, as PTMs play a major role in determining protein function (Packer *et al.*, 2008) - a crucial reason to include 2-D electrophoresis in the broad suite of tools that constitutes modern proteomics. The availability of high quality pre-cast immobilised pH gradients (IPGs)

and second dimension gels has made 2-D electrophoresis a simple technique; mainly reliant on a high quality sample. The key reason that two-dimensional gels provide poor results is poor sample preparation, however, this does not have to be the case. Since 1996 a number of publications have reported and reviewed the use of novel reagents such as thiourea and new sulfobetaine surfactants, which improve protein solubilisation prior to isoelectric focusing (Rabilloud *et al.*, 2007). Thiourea at 2M, in combination with urea at 7M produces a far more chaotropic sample solution than the conventional 8M urea. However, the increased chaotropic power required a new class of surfactants to cope with the highly denaturing environment. Rabilloud's group developed a range of new chaotrope tolerant surfactants, the best of which, amido-sulfobetaine 14 (ASB-14) and C7bZ0 in combination with urea and thiourea (Rabilloud *et al*., 1999), provide what is currently the highest level of solubilising power for IEF. The increased solubility with these new reagents and separation on narrow range commercial IPGs has significantly increased the total number of resolvable proteins. However, the increased numbers of proteins solubilised from a single sample can cause difficulties when attempting to separate whole extracts on a single 2-D gel and complexity reduction via fractionation is essential (Herbert *et al.*, 2007). Fractionation has assumed a central position in proteomics, as it is now clear that the dynamic range of protein expression and chemical diversity within even the simplest proteome cannot be resolved by single extraction and separation steps. The methods for fractionation can be broadly divided into biological, ie. membrane or organelle purification, and chemical, ie. solubility or isoelectric point. These methods will be further discussed in chapter two.

The published literature regarding ticks has focused more on the pathogenic diseases carried by ticks rather than the ticks themselves. This parallels the situation with mosquitoes and the pathogens they carry such as malaria. Although two-dimensional gel electrophoresis and nanoscale chromatography/mass spectrometry techniques are

commonly used for disease characterization in human and mammalian systems, literature detailing the use of these techniques in ticks is scarce.

Madden *et al* (2004) attempted to profile proteins in saliva from two related tick species, *Ammblyomma americanum* and *Amblyomma maculatum* by one and two-dimensional gel electrophoresis and MALDI mass fingerprinting. 1-D-SDS-PAGE profiles of saliva showed proteins between 5-100kDa as expected. Two intense bands with molecular weights of 92 and 98 kDa were reported to be identical to the major hemolymph protein in each tick species (Dillwith *et al*., unpublished), however this data remains unpublished. The N-terminal sequences for the isolated 92 kDa protein from *A. americanum* and *A. maculatum* are identical and are nearly identical to published sequences for hemolymph proteins from *Rhipicephalus (Boophilus) microplus* (Maya-Monteiro *et al*., 2000) and *Dermacentor variabilis* (Gudderra *et al*., 2001). However, these proteins were not able to be identified in two-dimensional gels and the authors (Madden *et al*, 2004) noted that in their experience, proteins above 90kDa do not focus well in IPG strips. A number of authors and reviews have reported proteins up to 250kDa being resolved by 2-DE (Gorg *et al.*, 2004). For those high molecular weight proteins that cause difficulty, cup loading has been shown to produce better results than passive rehydration (Gorg *et al*, 2004). Western blots of the 1-D-SDS-PAGE separated *A. americanum* and *A. maculatum* saliva suggested that majority of proteins were host proteins, in this case sheep. The authors observed that other work has shown that saliva from a number of ixodid tick species contains host serum proteins and the abundance of host proteins in the saliva of *A. americanum* and *A. maculatum* will make it difficult to use proteomic methods to characterize proteins of tick origin. This may vary among tick species, however with these two tick species the authors state it will be necessary to carry out a preliminary separation of host and tick proteins and such a group separation might best be accomplished using immunoaffinity chromatography with antibodies directed to host blood proteins.

Using two-dimensional gel electrophoresis and mass spectrometry, Untalan *et al* (2005) characterized the protein expression in unfed larvae of the cattle tick, *R. microplus*. The purpose of this was to assemble a database of proteins produced at the larval stage of development to use as a resource for the mining of molecules taking part in tick-host interactions. In this study, only one protein, tropomyosin, was unequivocally identified from its peptide mass map. Ten other proteins were assigned putative identities by subjecting Mascot-generated peptide sequences to BLAST searching of heterologous databases. These were a cytoskeletal protein (troponin I), multiple cuticular proteins, a glycine-rich salivary gland-associated protein and proteins with a presumed housekeeping role (arginine kinase, a high-mobility group protein and a small heat shock protein). Using sequential extraction, 550 Tris-soluble and 250 urea-soluble proteins were resolved in the isoelectric point (pI) range of 3–10, with the majority of Tris-soluble proteins present in the pI range of 5–8 and in the molecular mass (Mr) range of 10–100 kDa, whereas abundantly expressed urea-soluble proteins were predominantly acidic and of low molecular mass. The majority of the proteins identified had not previously been reported in *R. microplus*, and only a few examples of such proteins exist in the public database from the Subclass Acari. These include two troponin I-like sequences (*Haemaphysalis longicornis* and *Sarcoptes scabiei*), two ArgK sequences (*Psoroptes ovis* and *S. scabiei*) and one each of the HSP20-like (*D. variabilis*) and HMG-like sequences (*D. variabilis*).

Two-dimensional gel electrophoresis was used to examine host-vector-pathogen interactions in *Ixodes ricinus* by Vennestrom and Jensen (2007). A single freshly moulted nymph was crushed and the protein released solubilised in a single extraction step before focusing the sample on an 18cm pH 4-7 IPG strip. This was repeated for 20 nymphs frozen at time points from 2-30 weeks and the gels analysed to determine temporal expression of proteins. The authors were able to resolve approximately 4000 proteins, of which 9 were identified as host proteins (mouse or hen) that persisted in

the tick for 30 weeks, supporting the observations of Madden *et al* (2004). The authors were only able to identify one tick protein, actin, using MALDI MS/MS data.

Oleaga *et al* (2007) examined the salivary proteome of the soft argasid ticks, *Ornithodoros moubata* and *O. erraticus* using two-dimensional gel electrophoresis combined with Western blotting and *de novo* peptide sequencing to determine important allergenic proteins in the saliva. In *O.moubata* salivary gland extract (SGE), 70 protein spots were detected by silver staining. By probing Western blots with sera from pigs infested with each species, 18 antigenic spots were found which corresponded to some of the most highly abundant silver stained spots. The 18 antigenic spots were identified as 18 isoforms of a protein similar to *O. savignyi* TSGP1, however this identification was made using a single *de novo* sequenced peptide from only one of those 18 spots. TSGP1, belonging to the lipocalin family, is thought to play a role in salivary gland granule biogenesis (Mans and Neitz, 2004). Two non-antigenic spots were identified as the moubatin precursor by MS/MS ion search of NCBInr. In *O. erraticus* SGE, 180 spots were revealed by silver staining and 59 antigenic spots by Western blot. Six novel proteins were identified from *de novo* peptide sequence as similar to unknown secreted protein DS-1 precursor, NADPH dehydrogenase subunit 5, proteasome alpha subunit, ATP synthase F0 subunit 6, lipocalin and alpha tubulin. MS/MS ion searches did not identify any of the spots. *De novo* peptide sequences for both ticks were also searched against the four tick EST databases at The Gen Index databases (http://compbio.dfci.harvard.edu/tgi/), namely, *Amblyomma variegatum*, *R. microplus*, *Ixodes scapularis* and *Rhipicephalus appendiculatus*. Two proteins were identified, a fragment in the proteasome alpha subunit of *A. variegatum* and alpha tubulin from *I. scapularis* and *A. variegatum*. It is interesting to note that many spots analysed yielded one peptide for which the MS/MS data could be *de novo* sequenced and thus identify the protein. Most identifications were made with *de novo* sequence rather than MS/MS ion searches. The authors also note their surprise in finding none

of *O.moubata*'s reported pharmacologically active salivary proteins in the SGE protein spots analysed, except Moubatin.

A differential display experiment using two-dimensional gel electrophoresis was conducted by Rachinsky *et al*. (2007a & b) on the ovaries and midgut of uninfected and *Babesia bovis*-infected *R. microplus*. This study used a sequential extraction methodology to fractionate Tris-soluble proteins from proteins soluble in chaotrope/surfactant solutions (urea, thiourea and SB3-10). The authors also used liquid phase isoelectric focusing to fractionate these extracts and better resolve protein spots on narrow range IPGs. In the ovaries, 43 differentially expressed proteins were found and 19 were able to be identified by searching a translated *R. microplus* EST database, again illustrating the problem of MS/MS ion spectra searching of neglected genomes. Upregulated proteins identified included calreticulin, myosin chains, cytochrome c oxidase and serine protease inhibitors. Downregulated proteins included putative salivary proteins and tick lysozyme. In the midgut, 20 differentially expressed proteins were found and 16 identified. Proteins upregulated included six proteins implicated in signalling processes and five metabolic enzymes while downregulated proteins included a molecular chaperone, a cytoskeletal protein and a member of the prohibitin family. This work demonstrates that a well designed differential display can reveal many potential vaccine targets for both tick control and pathogen transmission that would not have been easily revealed with previously used separation technologies.

The salivary gland proteome of *Argas monolakensis* was examined using a variety of gel-based and chromatographic techniques, as well as cDNA analysis, in the work of Mans *et al* (2008). These included Edman sequencing of SDS-PAGE separated proteins blotted to PVDF, 2-DGE separated extracts with 78 spots subjected to trypsin digestion and MS/MS ion searching with Mascot, and sequential separation of the extract with anion and cation exchange chromatography followed by reversed phase

chromatography and peptide mass fingerprinting of 116 fractions. Of 14 proteins subjected to Edman sequencing, 13 had N-terminal sequences matching cDNA sequences. It was thought that these would correspond to the most abundant cDNA sequences, but the protein with the most transcripts (150), lipocalin, was not found by Edman sequencing and this was attributed to many cellular proteins being N-terminally blocked. Of the 78 spots cut from the 2-DGE, 18 matched cDNA transcripts and 14 proteins were identified, seven correlating with Edman products. The most prominent spot corresponded to lipocalin. The 116 ion exchange/reversed phase fractions yielded 52 identifications and 27 proteins, 11 of which had been identified by Edman sequencing and four by 2-DGE. Each method also identified unique proteins not found by the other methods, Edman sequencing giving four unique proteins, 2-DGE giving two and 2-DLC giving 14. This result has also been observed in yeast (Breci *et al*., 2005) where 2-D-PAGE and 2-DLC analysis gave complementary datasets with many hundreds of proteins uniquely identified in each approach. In *A. monolakensis* salivary glands, 29 proteins were identified which corresponded to ~60% of total salivary protein content.

Francischetti *et al* (2008) examined the salivary gland proteome of the soft tick *Ornithodorus parkeri* as part of a study of the salivary transcriptome, which was used to as a database to search the proteome data. A combination of different gel based techniques were used. 1-D-PAGE was used to separate proteins sequentially extracted with water and SDS sample buffer, then bands cut and trypsin digested for tandem mass spectrometry or the gel Western blotted and blotted proteins subjected to Edman sequencing. Edman sequencing revealed the presence of three lipocalins, three peptides from the basic tail superfamily, three Kunitz-containing peptides and a protein of unknown function. 2-D-PAGE on mini (7cm) gels resulted in 60 spots being cut for trypsin digestion and tandem mass spectrometry. Several of these were identified as lipocalins and a guanylate-binding protein and a ribosomal protein were also identified, but many spots were not identified. The greatest number of protein

identifications came from trypsin digests of 1-D gel separated proteins. Twenty-one lipocalins, six Kunitz-domain proteins, and five basic tail proteins were identified as well as cystatin, a fibrinogen-domain containing peptide and a phospholipase A2. As with all of the studies reviewed here, the majority of the proteins have no confirmed or known function.

This review of the literature demonstrates that two-dimensional gel electrophoresis is able to be used to study many aspects of the tick proteome, including host interactions and pathogen transmission. Very importantly, it also shows that host proteins are a significant issue when trying to characterise tick or pathogen proteins that are present at low abundance. With blood feeders, such as ticks, the endogenous proteins are substantially masked by the massive abundance of host proteins, especially albumin and α- and β-globins from haemoglobin. The necessity to study tick-host-pathogen interactions in their native environment means that mixed proteome samples are inevitable. Complex mixture fractionation and abundant protein depletion techniques are thus critical to detect the low abundance proteins produced by the tick-host-pathogen interaction. Another issue revealed by the tick literature is the lack of complete or extensively sequenced genomes. The lack of sequence databases often requires mass spectrometry-based *de-novo* peptide sequencing to enable identification of proteins by cross-species matching. This is considerably more time consuming than using direct database matching algorithms such as Mascot and Sequest. This will be further discussed in chapter two.

Chapter two – The proteomics of *Ixodes holocylcus*.

Preface.

In chapter one, the tick proteomic research reported to date was reviewed. The majority of these studies used two dimensional gel electrophoresis (2-DGE) as the main separation technique, although this is changing. Two recent papers from the Laboratory of Malaria and Vector Research at the National Institute of Health used 2-DGE in concert with 1-D SDS-PAGE and ion exchange chromatography to fractionate the proteins in excised salivary glands of *Ornithodorus parkeri* and *Argas monolakensis* (Francischetti *et al.*, 2008, Mans *et al.*, 2008). Rachinsky *et al* (2007 a & b) used liquid phase isoelectric focusing to fractionate proteins from excised ovaries and mid guts prior to focusing on narrow range IPG strips. While adding a proteome fractionation step, the dissection of ticks and the excision of individual organs from unengorged *Ixodes holocyclus* was not performed as much of the previous work in our laboratory dating back to the early 1990's has been performed on whole *I. holocyclus* ticks and low numbers of ticks were available. Thus the work in this chapter details the fractionation of the proteins from whole unengorged ticks using a variety of techniques and technologies that could subsequently be applied to dissected organs. The aim of this work was to find the fractionation workflow that would reduce or isolate high abundance proteins and allow the greatest number of tick proteins to be identified. The sample was first fractionated by protein solubility and then by isoelectric point, using IPG strips or liquid phase isoelectric focusing, or by size, using SDS-PAGE with vertical slab gels or continuous elution electrophoresis. The types of fractionation methodologies able to be used was constrained by the need to keep the proteins intact until they could be purified to a high level of homogeneity, either in a 2-D gel spot or discreet 1-D gel band. This was necessary as there is a lack of completely or extensively sequenced tick genome or proteome data available for searching with mass spectrometry data. This results in many proteins spots not being able to be identified even though good quality MS/MS data for the peptides was acquired. These MS/MS spectra were to be subjected to *de novo* peptide sequencing and, by keeping

the proteins intact until they were as homogenous as possible, individual peptides and their MS/MS data could be grouped together as being part of one protein. Once the sample preparation and fractionation work flows were determined they could then be used as the basis for determining the presence of specific protein types such as enzymes using the 2-DGE based fluorescent zymogram approach presented in chapter three. The work flows will also be used to prepare samples for equalisation experiments as presented in chapter four.

2.1: Sample Preparation of Ticks for Gel-based Proteomics.

In choosing a sample preparation methodology it must first be decided which part of the proteome the researcher wishes to isolate and the subsequent analysis techniques that will be used, such as 2-DGE or chromatographic methods. For a thorough review of sample preparation and fractionation techniques, the reader is directed to the review of Herbert *et al* (2007). The ongoing problem in proteomic sample preparation for both 2-DGE and multidimensional chromatography is the issue of high abundance proteins, including host proteins in the case of ticks, masking the existence of low abundance proteins. This issue of dynamic range has been addressed to some extent with improvements in fluorescent stains such as Sypro Ruby (Invitrogen) and Flamingo (Bio-Rad), which offer over five orders of magnitude in dynamic range. However, fractionation or sequential extraction (Herbert, 1999) of complex proteomes allows the enrichment of specific parts of the proteome and thus greater sensitivity. To analyse and characterise a protein it must first be able to be isolated in sufficient quantity to be analysed and thus to analyse low abundance proteins, the researcher must start with large amounts of material as no techniques exist to amplify the concentration of proteins as DNA concentration can be amplified using the polymerase chain reaction (Herbert and Harry, 2008). Using the example of yeast, 50% of the protein content of a cell is the product of 100 genes and the remaining proteins are the result of thousands of genes (Ghaemmaghami *et al*., 2003). More importantly, 75% of the yeast proteome has been found to be expressed at less than 5000 copies per cell. To put this into perspective, 1 fmol is generally considered an appropriate amount to determine a protein's identity using modern tandem mass spectrometry techniques. 1 fmol of a 50kDa protein is approximately equivalent to 6 x 10^8 molecules and if the 50kDa protein is present at 5000 copies per cell, 1.2 x 10^5 cells are required to obtain 1 fmol of protein. Of course, this does not take into account losses during sample handling and fractionation and thus the minimum number of cells required is more like 1 x 10^6.

All of the work presented in this chapter has used whole ticks as the starting material and grinding of ticks in liquid nitrogen with a mortar and pestle as the first step in sample preparation. The protein present in this powder can then be wholly extracted by sonication in ice-cold 10% trichloroacetic acid (TCA) in acetone, or sequentially extracted using firstly 40mM Tris-HCl pH 8.8 to obtain a soluble protein fraction and secondly 7M urea, 2M thiourea, 1% C7BzO surfactant and 40mM Tris-HCl pH 8.8 to obtain less soluble proteins. Following these sequential extractions, there is still insoluble protein present, but their solubilisation and analysis is beyond the scope of this work and it is likely that this protein is chitin and other cuticle proteins. C7BzO is a relatively new zwitterionic surfactant which has been shown to have greater solubilising power than CHAPS or SB 3-10 (Chevallet *et al*., 1998) and has not been used in any tick proteomic studies published to date. Following sequential extraction, all protein samples are buffer exchanged by precipitation, membrane filtration or gel filtration based centrifugal devices such as a BioSpin (Bio-Rad) into 7M urea, 2M thiourea, 1% C7BzO surfactant and 40mM Tris-HCl pH 8.8 for cysteine reduction and alkylation with tributylphosphine (5mM) and acrylamide monomers (20mM) in a single step. This negates the need for reducing agents during isoelectric focusing and SDS-PAGE and greatly improves spot resolution on two-dimensional gels (Herbert *et al*., 2001). The use of acrylamide instead of iodoacetamide for alkylation is due to iodoacetamide being scavenged by thiourea and acrylamide not reacting with reducing agents, as iodoacetamide can, allowing the reduction and alkylation to be done in a single step (Herbert *et al*, 2007). Samples are then buffer exchanged into 7M urea, 2M thiourea, 1% C7BzO surfactant and the conductivity of the sample checked with a Horiba conductivity meter. Samples reading above 200uS are desalted, checked again and then centrifuged at maximum speed in a microfuge before isoelectric focusing. These simple measures remove the majority of streaking problems that can occur on two-dimensional gels.

2.1.1: Subfractionation techniques.

Most 2-DGE based proteomic experiments use a 'one extract, one gel' approach. As this mainly identifies high abundance proteins, it is advantageous to subfractionate protein samples by physical or biological properties to reduce sample complexity (Herbert and Harry, 2008) which can have the side effect of increasing resolution. This is necessary because the diversity and dynamic range of proteins in an organism means that even the simplest proteome cannot be captured in its entirety using any single extraction and separation step (Herbert and Harry, 2008). By reducing sample complexity, each fraction can be studied in more detail and proteins which would have remained undetected in a total extract are present in sufficient quantities.

Fractionation by liquid phase isoelectric focusing is very useful as it allows the use of very narrow range IPG strips which can be loaded with a fraction containing proteins with isoelectric points (pI) only in the range of the strip. This allows the loading of a far greater amount of sample as capacity is not being wasted on proteins that focus to the extreme ends of narrow range IPGs. These proteins that focus to the ends of the IPG can precipitate due to the high protein concentration and are conductive as they are still charged due to not being at their isoelectric point, causing streaking and smearing problems (Herbert and Righetti, 2000). The two main liquid phase IEF techniques used in our lab are the Multi-Compartment Electrolyser (MCE) (Herbert & Righetti, 2000. Proteome Systems) and the RotoFor (Righetti *et al*, 2003, Bio-Rad). Using the MCE to subfractionate whole *E.coli* lysate, Herbert & Righetti (2000) were able to load five times the amount of protein onto a narrow range pH 4-5 IPG without introducing streaking or smearing seen in an unfractionated sample.

Liquid phase IEF techniques can also be coupled to preparative electrophoresis techniques such as the Bio-Rad Prep Cell which uses SDS-PAGE to separate proteins which are then eluted from the gel and collected using a peristaltic pump and fraction collector. Each fraction contains either a single protein, isoforms of the same protein

or a small number of proteins of similar pI and molecular weight depending on the initial pH fractionation method and gel percentages chosen for the second dimension separation steps. The advantage of this approach is that proteins are in liquid rather than gel and should be more easily enzymatically digested. Peptide recovery is higher than from gel pieces which ranges from 50-85% (Granvogl *et al.*, 2007b) and unlike gel pieces, these peptides can be used for both Edman degradation or MS/MS *de novo* sequencing experiments.

Another powerful subfractionation technique is the use of affinity chromatography (Righetti *et al*., 2003). Depletion of high abundance proteins in human serum, plasma and cerebrospinal fluid using specific antibodies bound to chromatography beads is a common approach, however in some cases proteins can be lost by non-specific binding. In ticks and other arthropods causing an allergic response, affinity chromatography using hyperimmune sera of the host can be used to isolate the specific proteins causing the response. Affinity fractionated samples are then separated further by two-dimensional gel electrophoresis after reduction and alkylation which is not carried out before the chromatography to ensure comformational epitopes are not removed. This targeted or hypothesis driven approach fractionates proteins that have caused an allergic response in an individual human host to salivary proteins that are injected into the host over a short period of time (hours). The work of Dorey (1991) has shown that sera from different individuals have antibodies specific for different tick proteins, although there are some commonly seen allergenic proteins on Western blots. In the case of hyperimmune dog sera, a commercially available anti-tick toxin treatment produced from dogs continually fed on by female *I. holocyclus,* antibodies in the sera are directed against salivary proteins excreted into the host over several days, some of which are involved in tick toxicoses. As mentioned in chapter one, these dogs show immunity to the tick toxin while continually fed on. Affinity chromatography techniques are also useful in

phosphoproteomics, glycoproteomics and thiol/disulphide proteomics (Azarkan *et al*., 2007).

2.1.2: Two dimensional gels vs multi-dimensional chromatography.

The lack of comprehensive databases of ticks and even related arachnids means that tick proteomic projects should use work flows that keep the proteins intact until they can be purified to homogeneity, essentially a top-down proteomics approach. This is necessary so that peptide sequences determined by *de novo* interpretation of tandem MS data come from the same protein, which cannot be done with global enzymatic digestion experiments where many digested proteins can be present in the same LC/MS/MS experiment or peptides from the same protein can be separated into different fractions in the first dimension. The main argument put forward for the use of multi-dimensional chromatography techniques is that it reveals poorly soluble proteins such as alkaline and membrane proteins that are reported to not be found with two-dimensional gels, mainly due to their low abundance and poor solubility (Santoni *et al.*, 2000). This argument has been weakened with improvements in surfactants, such as ASB-14 (Herbert, 1999), the use of organic solvents in a sequential extraction workflow (data not published, 2007; Molloy *et al*., 1999), the use of SDS as a solubilising reagent and SDS removal by TCA/acetone precipitation, and using salt to disrupt membranes in the presence of high pH or chaotropes that allows these proteins to be resolved on two-dimensional gels. Also, multi-dimensional chromatography suffers from the same problem as 2-DGE, namely the presence of high abundance proteins and the low abundance of membrane proteins.

Where chromatography can be very useful is the analysis of the tick salivary proteome where the amount of sample is extremely limiting. By being present in saliva, salivary proteins are by their nature soluble proteins and are able to be separated using denaturing and non-denaturing chromatography in the same way spider and scorpion venoms are separated (reviewed by Escoubas *et al*, 2008). Saliva can be fractionated

offline by reversed phase, ion exchange or gel filtration chromatography, homogeneous proteins tryptically digested, analysed by reversed phase LC/MS/MS and individual peptide sequences generated by *de novo* interpretation of MS/MS spectra. Using a sequence-similarity search approach, as outlined below (Waridel *et al.*, 2007), homologous proteins can be identified from other organisms. The emergence of nanoflow liquid chromatographs using >75 μm ID columns at flow rates >500 nl/min means that useful data can be generated from as little as 10 μl of collected saliva. The use of static nanospray tips at flow rates of >50 nl/min provides greater sensitivity for less complex samples. One recently introduced chromatographic technique that is of great use analysing poorly characterised organisms is LC-MALDI where the eluate from a nanoflow chromatograph is spotted onto a MALDI target plate rather than ionised by electrospray (Escoubas *et al.*, 2008). Peptides from a digested protein are still separated on a chromatographic timescale but individual peptides can be repeatedly sampled from the MALDI target on as long a timescale as necessary, unlike electrospray. Combining the LC-MALDI chromatographic approach with a MALDI TOF/TOF mass spectrometer allows excellent quality MS/MS data to be produced for *de novo* sequencing.

Multi-dimensional chromatography may also be useful for the quantitation of previously determined proteins at different stages of the tick lifecycle using isotopic labelling techniques such as iTRAQ (Applied Biosystems) although one could similarly use Difference Gel Electrophoresis (DIGE, GE Biosciences) on an uncharacterised sample. DIGE or iTRAQ could also be used to focus on identifying only those proteins that are up or down regulated between different samples, such as ticks fed on hosts infected with pathogens and hosts pathogen free as in the work of Rachinsky *et al* (2007). The iTRAQ reagents consist of 4-8 isobaric mass 'tags' that label primary amines of peptides in globally digested samples (Zieske, 2006). These tags add the same mass to all peptides and thus do not affect chromatographic retention or precursor mass. Fragmentation of the precursor, producing MS/MS data to identify the

peptide, also fragments the tag and each different tag produces a different reporter mass which can be quantitated against other reporters as they are generated at the same time and in the same experiment (Ross et al., 2004). Thus different samples can be individually digested and labelled before being combined and analysed to ascertain which proteins are up and down regulated in different 'treatments'. In the case of ticks, identification of the differentially displayed proteins may not be possible due to the protein not matching entries in current databases as discussed below. However if the difference in expression is significant, it would warrant further isolation and characterisation of the protein with a view to recombinant expression. Similarly, if ticks can be artificially fed (Krober and Guerin, 2007) SILAC (stable isotope labelling in cell culture) could be used to quantitate changes in protein expression on exposure to different pathogens in an appropriately designed experiment.

Selected or Multiple Reaction Monitoring (SRM/MRM) is a mass spectrometry based technique that can be likened to Western blotting in that it is able to detect very small amounts of a particular protein in a complex mixture of proteins. SRM/MRM experiments rely on prior knowledge of a proteotypic peptide from the protein of interest and the peptide's fragmentation products (Lange *et al.*, 2008). Using a triple quadrupole mass spectrometer (figure 2.1), Q1 can be set to only transmit the precursor mass of this proteotypic peptide to the collision cell (Q2) for fragmentation. Q3 is then set to only transmit a diagnostic fragment/s of the peptide. The result is much faster scan speeds and very low signal-to-noise ratios allowing detection of significantly lower concentrations of the peptide (Lange *et al.*, 2008). Using SRM/MRM experiments, attomole amounts of peptides are able to be identified and quantitative data can also be obtained. Prior knowledge of a proteotypic peptide's sequence can be derived from either a database entry or previous MS/MS data on protein spot from 2-D-PAGE. As described in section 1.2.5 and elaborated on in the next section, many tick proteins are not present in databases or have very low homology with other proteins from other organisms in the database. However if a protein spot is shown to

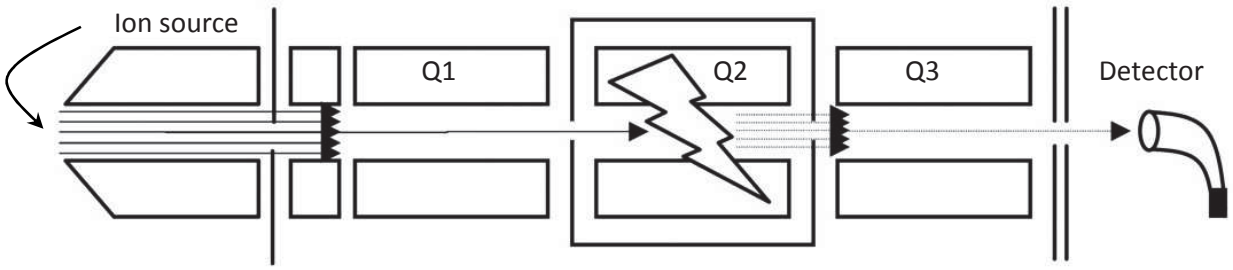

Figure 2.1 - Schematic diagram of triple quadrupole mass spectrometer operation during a multiple reaction monitoring experiment.
Quadrupole 1 (Q1) is set to only transmit the precursor mass of this proteotypic peptide to the collision cell (Q2) for fragmentation with either nitrogen or argon gas molecules. Quadrupole 3 (Q3) is then set to only transmit a diagnostic fragment/s of the peptide. All other precursors and fragments have unstable trajectories and are removed by the vacuum system. The extremely short scan times achieved allow sensitivity to be increased from femtomole levels to attomole levels.

be differentially displayed and good quality *de novo* peptide sequence for an unknown protein is obtained, this information would be sufficient to further characterise levels of this protein in samples such as serum from hosts by SRM/MRM experiments.

Another approach to increase the sensitivity of the mass spectrometer is a technique referred to as gas phase fractionation (Breci *et al.*, 2005). This method increases sensitivity by repeated analysis of the same sample over a discreet mass range of 100amu. The sample is analysed by 1-DLC/MS, but only the mass range between 400-500 m/z is scanned and only multiply charged ions in this range are selected for fragmentation. The experiment is then repeated for a 500-600 m/z range, 600-700 m/z range and so forth until 1400-1500 m/z. This increases sensitivity because the narrow mass range means shorter scan times and high abundance ions outside the range being scanned are not selected for fragmentation. Thus lower abundance ions are analysed that would not be previously as they would have eluted from the emitter before being selected while higher abundance ions are analysed. The disadvantage to this method is the necessity for 11 times the sample being consumed than used in a single standard full scan (400-1500 m/z) LC-MS/MS experiment. This amount of sample is often not available.

2.1.3: Tick databases and protein sequence determination.

As mentioned in section 1.5.2, ticks have very little gene (with coding region determined) and protein sequence data available to search data generated from proteomic projects. EST libraries provide the largest dataset available and searching of these EST databases with MS and MS/MS data with search algorithms such as Mascot, Sequest and Peaks ProteinID (Bioinformatic Solutions) can provide otherwise unknown identifications. It should be noted that entry descriptions are not often helpful and the EST often need to be BLAST searched to find other homologous proteins of known function (Choudhary *et al.*, 2001). Also, EST databases need to be translated into their correct open reading frame rather than increasing the size of the dataset by employing

a six frame translation. The program Prot4EST (Wasmuth and Blaxter, 2004) has been used in the work of Rachinsky *et al*. (2007b) to intuitively translate the available tick EST libraries and is currently being assessed in our laboratory to generate a dataset searchable with the SPIDER algorithm (Han *et al* 2005) which cannot use EST datasets. In many cases however, protein identification by BLAST searching either Edman degradation generated sequences (Rego *et al*, 2006; Francischetti *et al*, 2003) or *de novo* MS/MS sequences of enzymatically generated peptides may be only way to identify tick proteins of interest. The work of Oleaga *et al* (2007) reviewed in section 1.4.1 demonstrates how difficult this approach is.

Sequence-similarity search based tools, widely used in genomic projects, are slowly becoming more prevalent to identify proteins directly from MS/MS spectra without using current database search tools which use correlation of matched peptides. Current workflows, such as that outlined in Waridel *et al* (2007), use conventional database search tools, in this case Mascot, to filter the MS/MS data of known peptide sequences that are either from proteins present in the database or unknowns sharing peptides with known proteins. The remaining spectra are then filtered against a library of background spectra consisting common protein and chemical contaminants. These steps typically removed 75% of parent ion spectra in an unbiased fashion and without quality assessment with stringency of the filtering controlled by a user-defined p-value. The rectified dataset is then subjected to *de novo* sequence interpretation and sequence-similarity searching using MS-BLAST. This workflow completely automates identification of novel proteins even if there is only low sequence homology with the database entry and is a process we are eager to apply to our tick and other data from poorly characterised organisms.

The majority of MS/MS data is generated by subjecting tryptic peptides to collisional activated dissociation (CAD). This type fragmentation is affected if the tryptic peptides are too small or too large. In the case of large peptides generated by missed cleavages,

the presence of multiple basic residues prevents random protonation of the peptide backbone, limiting backbone dissociation to specific sites and inhibiting the generation of a diverse set of fragments (Mikesh *et al.*, 2006). This can result in no peptide sequence being able to be determined. In addition, MS/MS data quality can suffer when instruments are optimized for high throughput identification of samples using intelligent data acquisition experiments where collision energies are determined on-the-fly and fragmentation may not be occurring when the maximum concentration of peptide is entering the mass spectrometer.

A recent development in mass spectrometry, electron transfer dissociation (ETD), is a way of producing not only high quality spectra for peptide sequence determination, but also the post-translational modifications (PTM's) present on peptides are preserved rather than lost, which is the case when peptides undergo CAD fragmentation (Mikesh *et al*, 2006). ETD uses ion/ion chemistry to transfer an electron from a radical anion, fluoranthene, to a protonated peptide, fragmenting the peptide backbone by causing cleavage of the Cα-N bond on a millisecond timescale (Syka *et al.*, 2004). This generates complimentary c and z-ions while maintaining PTM's, regardless of peptide length, amino acid sequence or PTM's present (Swaney *et al* 2007), allowing the peptide sequence to be determined. ETD has been shown to be capable of producing sufficient ion series to determine 15-40 amino acids from the N and C-terminal of intact proteins and this was limited by the mass range of the instrument (Coon *et al.*, 2005). Approximately three ETD spectra/second vs. four CAD spectra/second can be obtained which is compatible with a chromatographic timescale (Mikesh *et al*, 2006). In addition, ETD uses relatively low cost ion trap instruments putting it within reach of more researchers. The only disadvantage is ETD is less sensitive than CAD because of the production of more product ions (McLafferty *et al.*, 2007) and higher charge states (3+ or more) are required for effective fragmentation (Swaney *et al* 2007; Pitteri *et al*, 2005) necessitating the use of enzymes other than trypsin such as Endo LysC.

In a recent paper by Zubarev *et al* (2008), the authors describe ETD and CAD as complementary techniques that should be used together, with high mass accuracy, to generate high quality, complementary *de novo* peptide sequence data. This data would allow much higher confidence identifications of proteins with a reduction in false positives by combining CAD generated b and y-ions with ETD generated c and z-ions and ETD's ability to differentiate between leucine (Leu) and isoleucine (Ile) by secondary fragmentation. Isobaric Leu/Ile (Xle) account for 1 in 6 amino acids in natural proteins and thus any peptide greater than 4 amino acids long will likely contain an Xle residue (Zubarev *et al*, 2008). This approach would be of great use with any organism with little genome and protein sequence data.

Proteomics technology and sample preparation is easily able to be applied to the problem of ticks as vectors of disease. The recent work of Rachinsky *et al*. (2007a & b) has demonstrated that a well designed differential display experiment can provide valuable insight into ticks as a vector. But due to the lack of bioinformatics resources available, researchers of ticks and other "neglected genomes" have to scour the literature to find tools to better identify the proteins that are important in the tick lifecycle and their suitability as vectors of disease.

2.1.4: Mapping the *Ixodes holocyclus* proteome with two-dimensional gel electrophoresis.

Very little is known about the proteins present in any life stage of *I.holocyclus* and only 20 protein sequences have been submitted to NCBI. The proteome could be homologous to other better characterized tick species such as *R. microplus* however the work of Mans and Neitz (2004) suggests that this may not be the case. The authors point out that tick species may have diverged at a very early point and that blood feeding in the main tick families may have evolved independently. If this is indeed the

case, proteins with a certain function in one family could have quite a different amino acid sequence to the protein with a similar function in another family. The homology of *I.holocyclus* proteins with those of other tick species may provide more evidence supporting this suggestion.

For the initial surveying of the *I.holocyclus* proteome, two-dimensional gel electrophoresis was chosen. 2-DGE keeps the proteins intact throughout the fractionation and separation process resulting in single protein isoforms being cut from the gel for trypsin digestion and identification by mass spectrometry. The resolution of proteins into single isoforms is dependent on sample complexity and the resolution of the IPG strip used with narrow range strips (pH 4-7 or 5-8 for example) being more highly resolving than broad range strips (pH 3-10). This approach is imperative when working on poorly characterized organisms with limited databases because the peptides fragmented by MS/MS are from one protein. Current search algorithms for MS/MS are unable to cope with peptide sequence variation and a single amino acid substitution in a peptide can result in the peptide not matching any entries in the chosen database. However, by ensuring the peptides are from a single homogenous protein isoform, unmatched MS/MS data can subsequently be *de novo* sequenced and BLAST searched, revealing the amino acid substitution and thus the protein's identity. In addition, 2-DGE allows characterization of the pI and MW of the protein spot, providing further confirmatory information on the protein's identity. Also, isoelectric focusing is very sensitive to contaminants such as salt, so a successful sample preparation method for 2-DGE will be easily transferable to other separations platforms.

This initial set of experiments was carried out on unengorged male and female ticks due to the fact that a reliable supply was found. The whole tick proteome was fractionated by solubility using first, a physiological buffer (Tris-HCl pH 8.0 with salt) and then chaotropes and surfactants (Urea, Thiourea, C7BzO). These were run on

broad range 11cm IPG's (3-10) in the first dimension and on 4-12% BisTris gels in the second dimension before staining with fluorescent stain (Flamingo, Bio-Rad) and over staining with colloidal Coomassie blue for spot cutting. Following initial experiments to optimize sample preparation and 2-DGE, the prepared sample was subjected to fractionation using a multi-compartment electrolyser (MCE) to provide significantly greater loading capacity and resolution of the proteins on narrow range IPG's. Also, 2-DGE separated proteins were blotted to nitrocellulose for probing with allergic sera from humans and dogs. Finally, 1-D SDS-PAGE and 1-DLC/MS was used to perform a shotgun proteomics experiment to ascertain whether the approach would yield complementary data to that obtained using 2-DGE and the MCE.

2.2: Methods.

2.2.1: Sample preparation and protein fractionation.

Figure 2.2 diagrammatically outlines the experimental steps used in this chapter. 100 ticks were snap frozen in liquid nitrogen and ground to a fine powder in a mortar and pestle pre cooled with liquid nitrogen. The powder was transferred to a centrifuge tube and 5 mL of 40mM Tris-HCl, pH 8.8 with or without 150mM LiCl added. Lithium was added in the later sample preparation for two reasons, the first being that other experiments in our lab have shown that the addition of LiCl resulted in more proteins being solubilised due to its ability to disrupt interactions between proteins and contaminants such as cell wall material. The second reason was to replicate solubilisation conditions used in the equalization experiments presented in chapter five, where the powder was solubilised in phosphate-buffered saline, which contains 150mM NaCl. It is well documented in the literature that salts can act to disrupt interactions between macromolecules. For example, in the membrane proteomics literature 100mM sodium carbonate stripping has been considered the benchmark in membrane stripping methods since its introduction in 1982. It is not clear why Fujiki *et al.* (1982) selected sodium as the salt of choice for membrane stripping when the literature on the boundary effects of other anions and cations is extensive. In 1888 Franz Hofmeister defined the series of anions and cations, the Hofmeister series, that have consistent chaotrope-like effects on the solubility of protein. Salt ions, either cations or anions, pair with ionised groups on proteins and can either enhance or reduce protein solubility depending on factors such as buffer pH and the isoelectric points of the affected proteins. Observations made from the Hofmeister series literature and by experiments conducted at the UTS Proteomics Centre suggested that lithium may be a more effective protein solubilising agent than sodium.

The solution was then sonicated for 6 x 30 seconds using a microtip probe at maximum power, typically 25-30W. The solution was chilled for 30 seconds on ice between

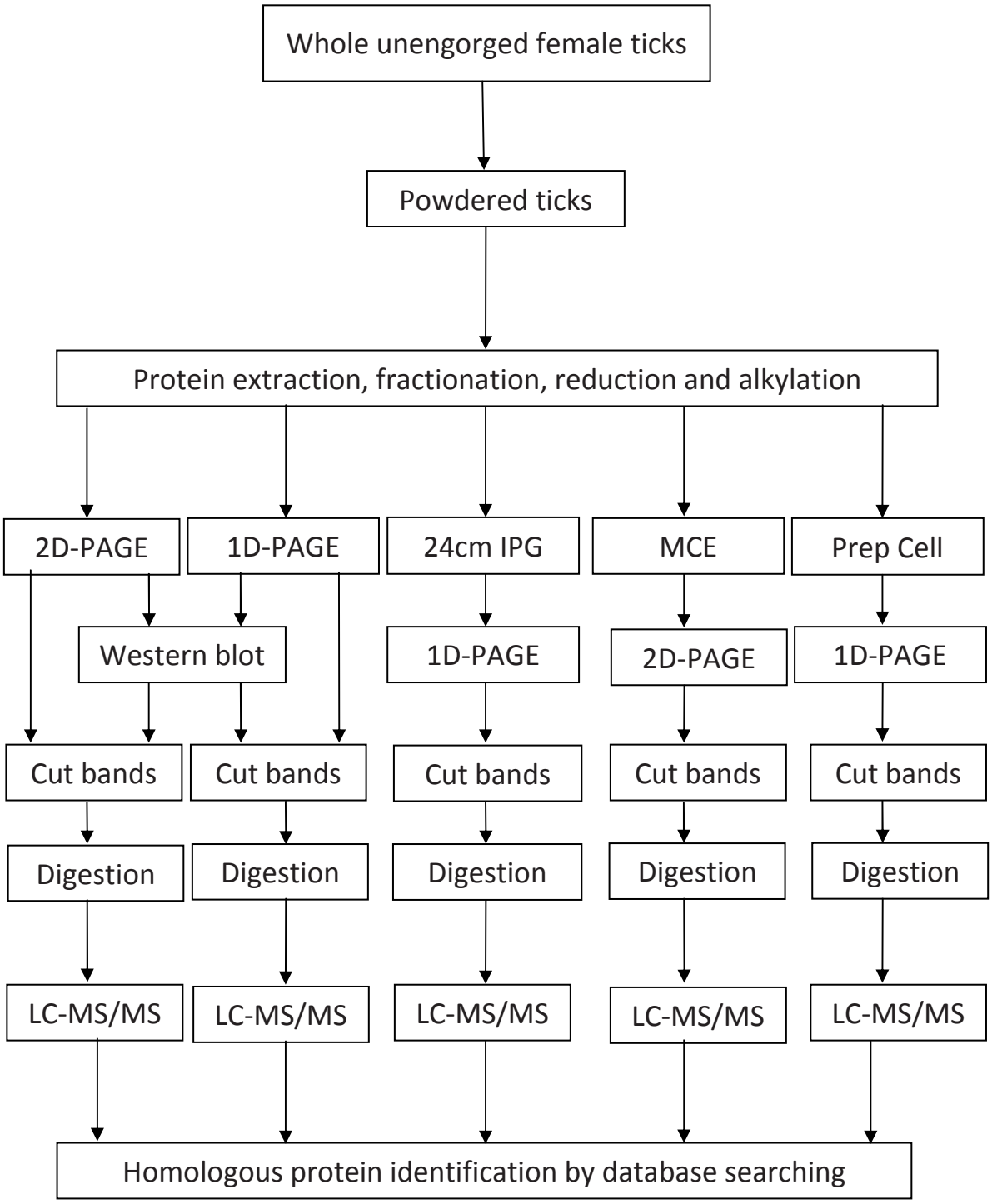

Figure 2.2: Diagramatic outline of the experimental steps carried out in this chapter. Exact methods are described in detail in the body text of the chapter.

sonicator bursts before being centrifuged at 38000g for 10 minutes at room temperature. The supernatant, containing the soluble proteins and designated fraction 1, was then removed and replaced with another 5mL of 40mM Tris-HCl, pH 8.8 and the sonication and centrifugation procedure repeated a total of three times. The extraction was repeated once more before the pellet was resuspended in 5mL of 7M urea, 2M thiourea, 1% C7BzO in 40mM Tris-HCl, pH 8.8 (UTC7Tris) with or without 150mM LiCl and the sonication and centrifugation procedure repeated a total of three times. Protein in remaining pellet was then extracted by adding 5mL of 150mM Tris-HCl pH 8.8 + 0.2% SDS and sonicating in waterbath for 30 minutes before centrifuging insoluble material at 38000g for 10 minutes.

As the samples were already at pH 8.8, no pH adjustment was necessary prior to reduction and alkylation of cysteine. Reduction and alkylation was carried out in a one step process using tributylphosphine (5mM final concentration) and acrylamide monomers (20mM final concentration) at room temperature for 90 minutes. The reaction was stopped and excess acrylamide quenched by adding DTT (10mM final concentration) and the proteins precipitated by the addition of five volumes of acetone and allowed to stand for 30 minutes before centrifugation at 3000g for 10 minutes. The acetone was poured off and any excess allowed to drain, by standing the tubes upside down on a tissue for 5 minutes. The protein pellet was then resuspended in 7M urea, 2M thiourea, 1% C7BzO (UTC7) and the sample's conductivity checked with a Horiba micro conductivity meter. If the conductivity was below 200µS, the sample was ready for isoelectric focusing, while conductivity readings above this value meant desalting of the sample with MicroBioSpin 6 columns (Bio-Rad) equilibrated in UTC7 (2 x 500µL) was necessary. Samples containing LiCl were always desalted, as acetone precipitation does not remove all of the LiCl.

2.2.2: One dimensional SDS-PAGE and protein concentration determination.

To ascertain the efficiency of protein extraction with each solution, 5µL of each fraction was combined with 2x sample buffer (250mM Tris-HCl pH 6.8, 10% glycerol, 2% SDS, bromophenol blue), heated to 95^{O}C for 5 minutes, briefly centrifuged and loaded into the wells of a 4-12% Criterion XT BisTris gel with MES running buffer (Bio-Rad). The gel was run at 160V until the bromophenol blue dye front had been run off the gel. Gels were fixed with 40% methanol, 10% acetic acid for 30 minutes before staining with Bio-Rad Flamingo fluorescent protein stain for 90 minutes. Images of the gels were taken using a Bio-Rad PharosFX Plus Molecular Imager utilising a 532nm excitation laser and 605nm band pass filter. The gel is shown in figure 2.4. This methodology was used for 1-D SDS-PAGE size fractionation of the sequential extracts for analysis of gel slices by 1-DLC/MS (section 2.3.3).

To determine the concentration of protein present, a dilution series of the extracts was prepared along with a dilution series of known concentrations of Bovine Serum Albumin (BSA). These samples were diluted with 2x sample buffer and loaded into the wells of a 4-12% Criterion XT BisTris gel with MES running buffer (Bio-Rad). The gel was run at 160V until the bromophenol blue dye front had been run approximately 0.5cm into the gel. The gel was then fixed, stained with Flamingo and imaged as described above. The gel is shown in figure 2.2. Using the Quantity One software (Bio-Rad), a standard graph of BSA concentrations was produced and the protein concentration of the Tris and UTC7Tris extracts determined and these are shown in figure 2.3.

2.2.3: Isoelectric focusing.

Bio-Rad 11cm pH 3-10 IPG Readystrips were partially rehydrated with 100µL of UTC7 for 30 minutes at room temperature. 150µL of sample containing approximately 450-600µg of extracted protein was actively loaded by pipetting the sample under the IPG strip, which had been laid facedown and covered in paraffin oil in a Proteome Systems IsoelectrIQ focusing tray. The sample is then drawn into the strip by applying a slowly

increasing voltage to the strip. The IsoelectrIQ instrument was programmed in the following way to focus the proteins in the sample. Step 1: 100-3000V in a convex ramp over 5 hours with current limited to 75uA/strip. Step 2: 3000-10000V in a linear ramp over three hours with current limited to 50uA/strip. Step 3: 10000V with a current limit of 35uA/strip until focusing had reached at least 100000 volt hours. Due to failure of the IsoelectrIQ instrument, some focusing experiments were carried out using a Bio-Rad Protean IEF cell with the only difference being the current being limited to 50uA/strip for the entire program. Isoelectric focusing on 24cm pH 3-10 IPG's was also carried out using the same program. Once focused, the strips were then electrophoresed and the gel was fixed and stained with Flamingo as described in section 2.2.4. Focused 24cm IPG strips were cut into 1 cm sections, which were placed in a 0.5mL tube with 50µL of 2x sample buffer. The pieces were then boiled for 10 minutes and the solution loaded into the wells of a 12 well, 4-12% Criterion XT gel and electrophoresed as described in section 2.2.2.

2.2.4: SDS-PAGE.

Focused IPG strips were equilibrated for 25 minutes in 7M urea, 250mM Tris-HCl pH 8.5, 2% SDS, with bromophenol blue. Equilibrated strips were then placed on top of a midi sized (11cm) 4-12% Criterion XT BisTris gel with MES running buffer (Bio-Rad). The gel was run at 160V until the bromophenol blue dye front had been run of the gel (approximately 70 minutes). Gels were fixed with 40% methanol, 10% acetic acid for 30 minutes, stained with Bio-Rad Flamingo fluorescent protein stain for 90 minutes and imaged using a Bio-Rad PharosFX Plus Molecular Imager utilising a 532nm excitation laser and 605nm band pass filter. Following this, the gels were overstained with colloidal Coomassie blue stain overnight and spots either excised manually or with a Bio-Rad EXQuest spot cutter. The gels are shown in figure 2.4. Spot numbers were determined for the Flamingo stained gels using Proteomeweaver (Bio-Rad) gel analysis software.

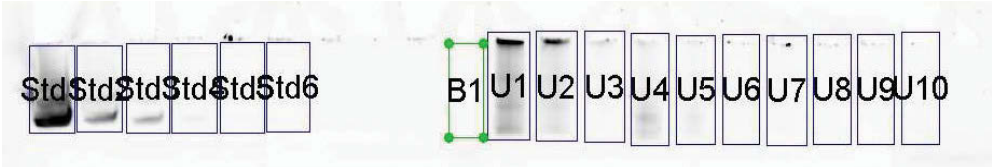

Figure 2.3 - Protein assay of extracted protein from whole, unengorged female *I. holocyclus* ticks.
Samples loaded are as follows.

- BSA standards (loaded in 10µL).
  1) 5ug.
  2) 1ug.
  3) 0.5ug.
  4) 0.1ug.
  5) 0.05ug.
  6) 0.01ug.
  7) 0.005ug.
  8) 0.001ug.
  9) 0.0005ug.
- Samples (loaded in 10µL).
  1) Tris Ext 1. 1:50 dilution.
  2) Tris Ext 1. 1:100 dilution.
  3) Tris Ext 1. 1:500 dilution.
  4) UTC7Tris Ext 1. 1:50 dilution.
  5) UTC7Tris Ext 1. 1:100 dilution.
  6) UTC7Tris Ext 1. 1:500 dilution.
  7) SDS. 1:50 dilution.
  8) SDS. 1:100 dilution.
  9) SDS. 1:500 dilution.
- Background labelled B1.

R-squared for BSA standard curve = 0.9988
Tris extracted protein concentration: 16.6µg/µL
UTC7Tris extracted protein concentration: 3.25µg/µL

2.2.5: Protein Digestion.

Excised spots were washed in 50mM ammonium hydrogen carbonate, 50% acetonitrile buffer twice for 10 minutes and then dehydrated with 100% acetonitrile. The acetonitrile was removed and the dehydrated gel piece dried by rotary evaporation for 1 minute before being rehydrated in 50mM ammonium hydrogen carbonate buffer containing 125ng of trypsin (Trypsin Gold, Promega) and incubated overnight at 37^{O}C. The gel pieces and digest solution were sonicated in a water bath for 10 minutes at room temperature before the digestion solution was removed to a clean tube and 50% acetonitrile, 2% formic acid added to extract more peptides with the aid of another 10 minutes in the sonicating water bath. The pooled extracts were concentrated by rotary evaporation to approximately 15µL and frozen at -20^{O}C until analysis by mass spectrometry.

2.2.6: Mass Spectrometry.

All mass spectrometry and data analysis was carried out in-house by the author. Using an Eksigent AS-1 autosampler and Tempo nanoLC system, 10µL of the peptide digest was loaded at 20µL/min onto a Michrom reversed phase trapping cartridge to wash away any contaminants. The bound peptides were then eluted at 500nL/min from the trap column onto a 75µm x 100mm monolithic silica C18 column (New Objective) with buffer A (2% acetonitrile, 0.2% formic acid). Peptides were then eluted at 500nl/min by a gradient from 2% buffer B (98% acetonitrile, 0.2% formic acid) to 30% buffer B over 8 minutes and to 80% buffer B after a further 2 minutes. In later experiments, due to greater experience with the mass spectrometer, the flow rate was reduced to 300nL/min (to increase sensitivity) and the initial gradient was extended to 2%-30% buffer B over 15 minutes. The gradient was further extended to 2%-30% buffer B over 30 minutes for the 1-D-SDS-PAGE separated samples.

Peptides were eluted from the column and ionised by a MicroIonspray II head holding a nanoelectrospray emitter (75µM tapering to 20µM, New Objective) at 2300V into the source of a QSTAR Elite Quadrupole TOF mass spectrometer set up for Intelligent Data Acquisition using Analyst 2.0 software (Applied Biosystems/MDS Sciex). Briefly, the data system analyses ions transmitted from Q1 to the TOF analyser for 2+ to 5+ ions and when such an ion is detected and measured above 30 counts per scan, it is selected by Q1 and transmitted to the Q2 collision cell for fragmentation by collision-induced dissociation (CID) with N_2 gas molecules. Fragment ions are then measured by the TOF analyser. The data system also determines the best collision energy for the fragmentation using the dynamic CE algorithm (part of Analyst software). Selected ions are then excluded for 15 seconds (for 2-DGE spots) or 45 seconds (for all other samples) to allow other lower abundance peptides to be sampled. A 15 second exclusion time means that a peptide can be fragmented again when the ion is at its highest intensity as it elutes from the analytical column which provides better MS/MS data for identification and *de novo* sequence determination. This methodology is practical when analysing spots from 2-DGE as a relatively low number of peptides (1-20) are present when compared to samples generated by 2-DLC and other shotgun methodologies that don't purify proteins to homogeneity.

2.2.7: Data Analysis.

The process for searching the data generated by from mass spectrometry is shown diagrammatically in figure 2.4. Raw MS/MS data was searched against the LudwigNR (comprised of the UniProt, plasmoDB and Ensembl databases), MSDB and SwissProt databases using the Mascot search algorithm (v2.1.0) provided by the Australian Proteomics Computing Facility. PEAKS ProteinID algorithm (Peaks Studio v4.5, Bioinformatics Solutions Inc) was used to search NCBI non-redundant (compiled from GenBank CDS translations and the PIR, SWISS-PROT, PRF, and PDB protein databases) and a self-assembled database of all protein sequences in the Acari taxon contained in NCBI's protein database along with EST libraries for *Amblyomma variegatum, Rhipicephalus (Boophilus) microplus, Ixodes scapularis* and *Rhipicephalus*

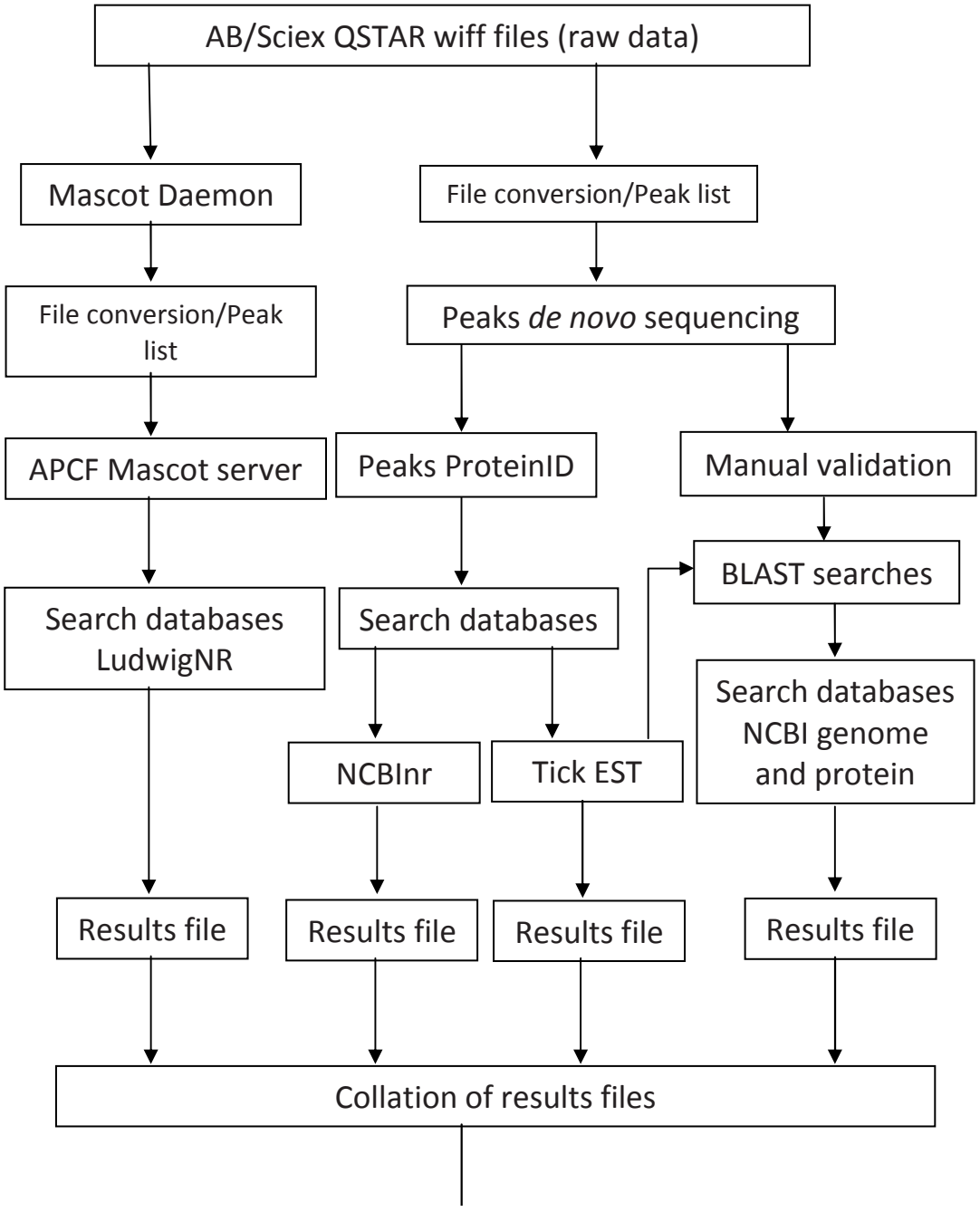

Figure 2.4: Diagramatic outline of the steps involved in searching mass spectrometry data generated. Exact methods are described in detail in the body text of the chapter.

appendiculatus (compbio.dfci.harvard.edu/tgi/). Matching nucleotide sequences in the EST databases were then compared to the NCBInr database with BLASTX (http://blast.ncbi.nlm.nih.gov/Blast.cgi). As Peaks Studio is not as prevalent in its use as Mascot or Sequest, an explanation of its results scoring is necessary. Peaks uses sequence tag information produced by *de novo* sequencing of the MS/MS spectra to filter the database being searched before fragment ion fingerprinting. The protein score, expressed as a percentage, is then calculated from the confidence of the ten best peptide hits for the protein and normalised against the other identified proteins. In contrast, Mascot uses a probability based algorithm to determine the best peptide match to an MS/MS spectra. Spots that were not identified by these ion searching methods were subjected to automated *de novo* sequence analysis using the PEAKS Studio 4.5 software package. Automatically *de novo* sequenced peptides with a score greater than 90% confidence were examined manually using PEAKS and BioAnalyst (Applied Biosystems/MDS Sciex) to ensure the determined amino acid sequence was correct and not due to noise in the mass spectrum. These peptide sequences were then used in MS-BLAST searches (Shevchenko, 2001. dove.embl-heidelberg.de/Blast2/msblast.html). For the shotgun experiments (sections 2.3.3 and 2.3.5), false discovery rates were determined by Mascot performing a decoy database search. During this search, for every protein or peptide sequence tested, a random sequence of the same length is automatically generated and tested. The false discovery rate is then reported as the percentage of the total number of peptide matches at a defined significance threshold, typically 0.05. In this work, the significance threshold was reduced to 0.01 to ensure the false discovery rate was as low as possible and the rates are reported in table 2.8.

2.2.8: Sequential extract separation by multi-compartment electrolyser.

The multi-compartment electrolyser (MCE) was fitted with pH 3, pH 5, pH 8 and pH 11 membranes creating five chambers, pH 1-3, pH 3-5, pH 5-8, pH 8-11 and pH 11-14. Extracted tick protein was diluted to a final volume of 5mL in UTC7 and loaded into the 5-8 chamber of the MCE while the other chambers were filled with 7M urea, 2M

thiourea. Separation of the proteins by pI was effected by applying the following program. Step 1, 6 hours at 100V, step 2, 6 hours 100-1500V using a linear ramp, step 3, at least 6 hours at 1500V. Following this program, the fractionated sample was removed as quickly as possible to 15 mL tubes to prevent the proteins diffusing back across the membranes. The pH membranes were also removed and placed in a 15mL tube and washed with 1mL of UTC7 to recover any protein bound to the membrane. The fractions were then concentrated using Vivaspin centrifugal concentrators with a 3 kDa cut off (Sartorious) to approximately 500µL and the protein content and concentration assessed by 1-D-SDS-PAGE as described in section 2.2.2. The fractions were then subjected to 2-DGE as described above using an appropriate narrow range 11cm IPG. These strips were pH 3-6 (Bio-Rad) for the pH 3-5 fraction, pH 5-8 (Bio-Rad) for the pH 5-8 fraction, pH 6-11 (Amersham) for the pH 8-11 fraction, and pH 4-7 (Bio-Rad) for the pH 5 membrane fraction. Following overnight focusing, the strips were equilibrated and second dimension SDS-PAGE performed as described above. After gel staining, spots of interest were then excised from the gel and analysed by mass spectrometry and database searching as described above. The gel images are shown in figure 2.6, the proteins identified are listed in table 2.3.

2.2.9: Size fractionation using preparative, continuous elution electrophoresis.

The Bio-Rad Model 491 Prep-Cell was used to perform SDS-PAGE as described in the manufacturer's instructions with the following modifications. A 10cm long, 28mm diameter, 10% acrylamide gel tube was cast and allowed to set overnight. After assembly of the Prep Cell and filling with Tris/Glycine running buffer without SDS, the sample, prepared by boiling in 2x sample buffer with SDS, was carefully layered onto the top of the gel tube and electrophoresis performed at 200V. When the bromophenol blue dye front began to elute from the gel tube into the collection manifold, 154 500µL fractions were collected at 50µL/min with a Bio-Rad FP-1 fraction collector. The buffer in the elution chamber was 50mM Tris-HCl pH 8.8. SDS was present only in the sample to minimize the amount of SDS present that could

potentially affect downstream processing of the protein fractions such as trypsin digestion.

10µL of every fifth fraction was analysed by SDS-PAGE as described in section 2.2.2 (figure 2.8). This was done to determine quality of the separation and the fraction number of highly abundant proteins. Fractions were combined as described in section 2.3.3 to isolate the high abundance proteins from lower abundance proteins. The proteins in the combined fractions were concentrated with centrifugal concentrators with a 3000Da cutoff and analysed by SDS-PAGE again to further assess the fractionation and estimate the protein concentration (figure 2.8).

2.2.10: Western blots of unengorged tick extracts with human allergic sera.

Sequentially extracted protein from unengorged *I.holocyclus* was prepared as described in section 2.2.1. 1-D and 2-D SDS-PAGE was conducted as described in sections 2.2.3 and 2.2.4. Proteins were blotted to nitrocellulose by the Khyse-Anderson (1984) method and a HEP-1 semi-dry blotter (Owl Separation Systems, Portsmouth, NH). The membranes were then blocked with 0.1% Tween-20 in TBS and probed with 10% allergic sera (human or dog) in TBS/0.01% Tween-20 for two hours as described by Dorey (1991). The human sera was sourced from a hyperallergic individual and obtained by plasmophoresis. The dog sera used is a commercially available treatment for *I. holocyclus* induced paralysis, produced from sera obtained from dogs continuously fed on by female *I. holocyclus.* Each sera will therefore contain antibodies to proteins from different stages of tick feeding, the human sera to tick proteins present at tick attachment and the dog sera to proteins present after many days of tick feeding. Following 3 x 10 minute washes with TBS/0.05% Tween-20, the blots were incubated for one hour with either anti-human IgE or anti-dog IgG, both conjugated to alkaline phosphatase. After another 3 x 10 minute washes with TBS/0.05% Tween-20, sera bound proteins were detected by incubation of the blot with BCIP/NBT (SigmaFAST) until bands or spots appeared. The reaction was stopped by washing the

gel in water and drying the blot on paper towel. The resulting blots are shown in figures 2.12 and 2.13. Detected spots in the Western blot (figure 2.13A) were matched to a Flamingo stained gel of the same sample separated in the same way. The matched spots were excised and analysed as described in sections 2.2.5, 2.2.6 and 2.2.7. The homologous proteins identified are listed in table 2.6.

2.3: Results.

2.3.1: Fractionation by protein solubility.

Figure 2.5 shows the SDS-PAGE of the sequentially extracted proteins of female unengorged *I.holocyclus*. It is evident that three 5mL extractions with 50mM Tris-HCl pH 8.8 + 150mM LiCl has been sufficient to extract all of the readily extractable soluble proteins in the sample as extract three contains very little protein. The subsequent denaturing extraction, UTC7Tris-HCl +150mM LiCl, which is extract 4, has a high protein concentration and the pattern is very different compared to the Tris-HCL extract. An extra extraction with UTC7Tris may have been useful as there is a significant amount of protein present in extract six and the pattern in extract seven, where proteins have been extracted with SDS, is almost the same as fraction six.

The gel images presented in figure 2.6 show the proteins present in unengorged whole female *I. holocyclus*. What is immediately apparent is the stark difference in the spot patterns produced when fractionating this proteome by protein solubility. Very few protein spots are seen above 100kDa and a number of spot "chains" are present. These chains are usually isoforms of the same protein that have been post-translationally modified. The determination of post-translational modifications was beyond the scope of this work. The Tris-HCl extracted, or soluble, proteins have isoelectric points uniformly distributed across the pH gradient. In contrast, the UTC7Tris-HCl extracted, or insoluble, proteins have a very concentrated collection of small acidic proteins and a distinct pattern of highly alkaline proteins.

Approximately 380 spots were detected in the Tris-extracted sample gel stained with Flamingo (figure 2.6A), but only 90 spots were excised from the Coomassie overstained gel for trypsin digestion. This was because only these 90 spots were deemed to have an optical density above a threshold that would allow homologous proteins to be

identified after analysis of the spot's tryptic peptides by mass spectrometry. Searching the LudwigNR database with Mascot identified homologous proteins in 40 of these spots with at least one high scoring peptide (Table 2.1). The spots returning no match with Mascot were searched against a tick EST database compiled in-house from the EST databases of *Amblyomma variegatum, Rhipicephalus (Boophilus) microplus, Ixodes scapularis* and *Rhipicephalus appendiculatus* (compbio.dfci.harvard.edu/tgi/) using the Peaks ProteinID search algorithm. This identified the proteins in a further 15 spots although only 10 of these were annotated to indicate what the ESTs were homologous to. The remaining five ESTs were BLAST searched to determine their homology to other proteins. The remaining 35 spots were not able to be identified by these MS/MS ion searching methodologies. Manual inspection of the 35 data files showed very few multiply charged ions were present and these were of low intensity. These ions were intense enough (over 30 counts) to be selected for fragmentation, but the resulting MS/MS spectra was of too poor quality for successful MS/MS ion searching or *de novo* peptide sequencing.

The successful identification of only 20% of the 380 Flamingo-stained gel spots is not surprising. The latest fluorescent stains, such as Flamingo and Sypro Ruby (Invitrogen) can routinely detect sub-nanogram amounts of protein in a single spot on a 2-D gel. This amount is below the practical detection limit of most mass spectrometers which can be as low as 10fmol, but routinely is 50-100fmol in this operator's experience. The detection limit of the QSTAR Elite Q-TOF mass spectrometer used in these experiments is approximately 1ng (33 fmol of a 30kDa protein, 20 fmol of a 50kDa protein), or a spot visible with colloidal Coomassie blue stain and many spots were faint after Coomassie staining. Although the basic procedure for in-gel trypsin digestion has not changed since Wilm and co-workers published the procedure in 1996, there are multiple variations of the procedure and due to their complexity, successful preparation is highly dependent on the skill of the operator (Granvogl *et al.*, 2007a). Different methodologies vary temperature (Turapov *et al.*, 2008), ultrasonication

(Lopez-Ferrer *et al.*, 2005), microwave irradiation (Juan *et al.*, 2005, Sun *et al.*, 2006), organic solvents (Granvogl *et al.*, 2007a) and pressure (Lopez-Ferrer *et al.*, 2008) to optimally digest proteins in gel spots and time was not available to evaluate which method was best for our lab. Maintaining an LC/MS system at the peak of its possible sensitivity is only part of successful protein identification using mass spectrometry.

Tris-extracted protein (600ug) was loaded and focused in the IPG strip, before being separated by SDS-PAGE (figure 2.6A). The majority of this 600ug is in only 30-40% of the spots as shown by high spot intensity. The other spots are not visible after Coomassie staining and thus insufficient protein (or tryptic peptides) is present for successful identification. One solution to this problem is to load more protein onto the isoelectric focusing strip and thus have spots of greater intensity. Due to the presence of high abundance proteins, especially actin in the Tris extracted sample shown in figure 2.6A, loading more protein will result in poor isoelectric focusing with high abundance proteins precipitating in the strip and not allowing other proteins to focus. Another solution is to further fractionate the sample by isoelectric point, in this case using a multi-compartment electrolyser (MCE), and loading this sample onto narrow pH range strips of the same physical length. By pre-fractionating with the MCE, 3-4 times the protein can be loaded. If the sample loaded onto the gel in figure 2.6A was loaded onto a narrow range strip without MCE prefractionation, at least half of the protein would focus to the ends of the strip. A greater number of spots were detected in the Flamingo-stained gel of the UTC7Tris extract than the Tris extract (figure 2.6B). Approximately 450 spots were detected and 96 were excised for identification, or approximately 20% of the total spots detected by Proteomeweaver. Searching the NCBI non-redundant database with Mascot identified homologous proteins in 37 of these spots with at least one high scoring peptide (Table 2.2). The spots returning no match with Mascot were again searched against the compiled tick EST database using the Peaks ProteinID algorithm.

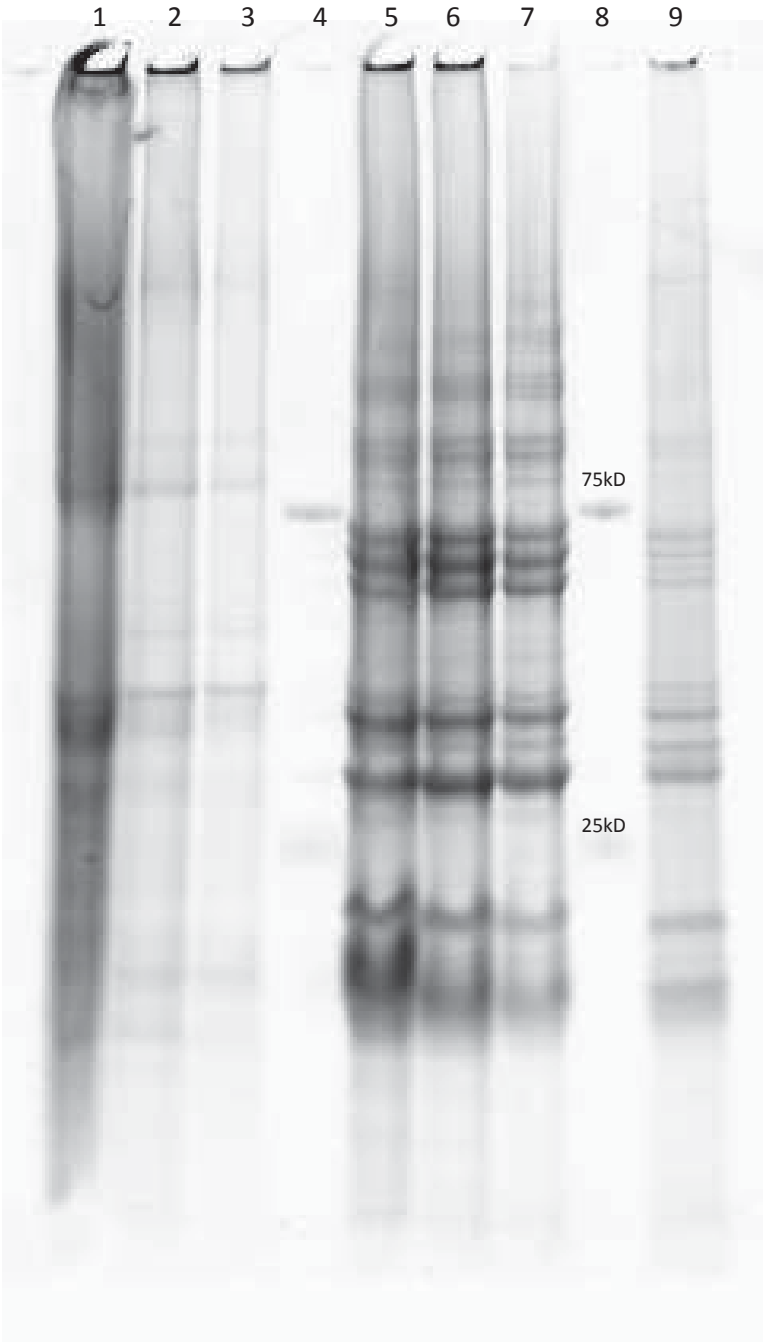

Figure 2.5: 1-D SDS-PAGE of sequentially extracted proteins from female unengorged *I. holocyclus*.

1) Tris-HCl pH 8.8 + 150mM LiCl extract 1.
2) Tris-HCl pH 8.8 + 150mM LiCl extract 2.
3) Tris-HCl pH 8.8 + 150mM LiCl extract 3.
4) Molecular weight markers.
5) UTC7Tris-HCl pH 8.8 +150mM LiCl extract 4.
6) UTC7Tris-HCl pH 8.8 +150mM LiCl extract 5.
7) UTC7Tris-HCl pH 8.8 +150mM LiCl extract 6.
8) Molecular weight markers.
9) 150mM Tris-HCl pH 8.8 + 0.2% SDS extract 7.

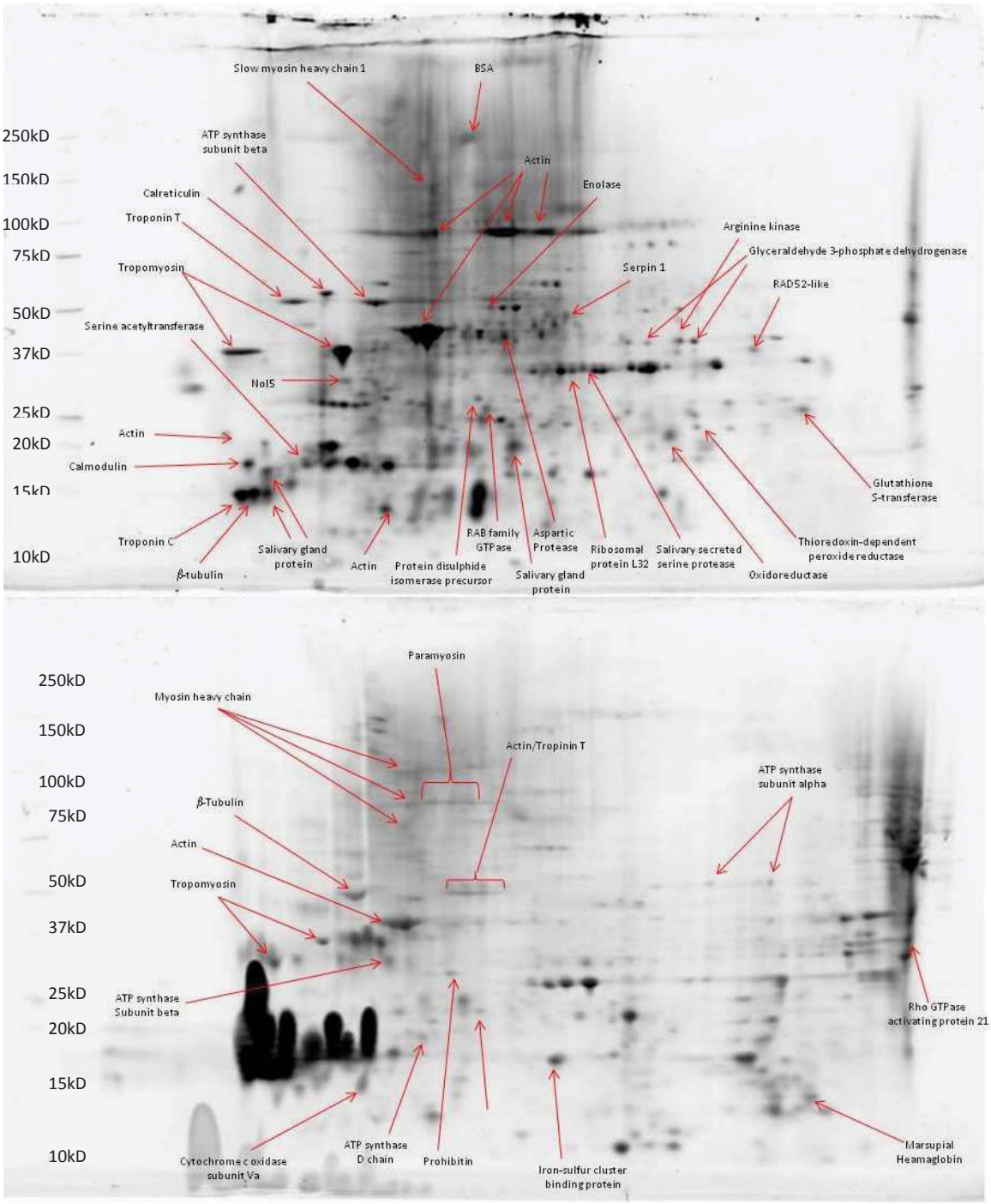

Figure 2.6: Sequentially extracted protein from whole, unengorged female *I. holocyclus* ticks separated by 2-DGE. The upper gel (2.4A) contains 600ug of protein extracted with 50mM Tris-HCl pH 8.8 + 150mM LiCl while the lower gel (2.4B) contains 490ug of protein extracted with 7M urea, 2M thiourea, 1% C7BzO, 50mM Tris-HCl pH 8.8 + 150mM LiCl. Isoelectric focusing was carried out on 11cm pH 3-10 strips and second dimension SDS-PAGE carried out on 4-12% BisTris gels.

Spot #	Description	pI	MW	Score	Matches	Coverage (%)	Database	Engine
1	tr|B0JYQ0|ALB protein.[Bos taurus]	5.95	69248	1096	36	29	trEMBL	Mascot
	sp|O97162|Tropomyosin.[Boophilus microplus]	4.7	32982	336	9	23	SwissProt	Mascot
2	sp|O97162|Tropomyosin.[Boophilus microplus]	4.7	32982	839	48	42	SwissProt	Mascot
3	tr|Q4PKE5|Actin 5.[Aedes aegypti]	5.3	41795	1122	114	68	trEMBL	Mascot
4	sp|P02769|Serum albumin precursor (BSA) (Allergen Bos d 6).[Bos taurus]	5.82	69248	346	7	11	SwissProt	Mascot
5	sp|P02769|Serum albumin precursor (BSA) (Allergen Bos d 6).[Bos taurus]	5.82	69248	542	15	19	SwissProt	Mascot
	tr|Q4PKE5|Actin 5.[Aedes aegypti]	5.3	41795	516	15	49	trEMBL	Mascot
A1	tr|Q4PKE5|Actin 5.[Aedes aegypti]	5.3	41795	511	17	45	trEMBL	Mascot
	tr|Q64K84|Calreticulin.[Ixodes jellisoni]	4.57	47612	125	4	8	trEMBL	Mascot
A2	tr|Q4PKE5|Actin 5.[Aedes aegypti]	5.3	41795	317	8	38	trEMBL	Mascot
A3	tr|Q5CAR2|Actin.[Ixodes ricinus]	5.3	41797	115	2	9	trEMBL	Mascot
A4	ISGI-G893P536FK9 similar to UP|Q6P6X6_BRARE (Q6P6X6) Nol5 protein, partial (29%)		76624	99%	3	6	ISGI	Peaks
A5	tr|A0A106|Enolase (EC 4.2.1.11) (Fragment).[Semiotellus sp. CD038]	5.82	41139	239	6	16	trEMBL	Mascot
A6	tr|Q5CAR2|Actin.[Ixodes ricinus]	5.3	41797	53	1	6	trEMBL	Mascot
A7	ISGI-TC8810 similar to UP|Q692U7_IXOSC (Q692U7) Salivary gland protein, partial (44%)		65972	90%	2	13	ISGI	Peaks
A8	tr|Q2VCI9|Aspartic protease-like.[Solanum tuberosum]	5.85	54859	85	1	3	SwissProt	Mascot
A9	AVGI-TC149		97165	90%	1	3	AVGI	Peaks
A10	tr|Q0MR22|RAD52-like protein.[Penicillium marneffei]	9	61801	56	1	2	trEMBL	Mascot
A11	ISGI-TC15028 homologue to UP|Q6P7N2_XENTR (Q6P7N2) MGC75629 protein (Eukaryotic translation initiation factor 2), partial (6%)		78825	90%	1	3	ISGI	Peaks
B1	tr|A8E4K0|Troponin T.[Haemaphysalis qinghaiensis]	4.93	46442	98	3	7	trEMBL	Mascot
B2	ISGI-TC627 similar to UP|Q692U7_IXOSC (Q692U7) Salivary gland protein, partial (51%)		80107	90%	1	4	ISGI	Peaks
C1	tr|Q6PTP3|ATP synthase subunit beta (EC 3.6.3.14) (Fragment).[Encope michelini]	4.96	45811	501	11	29	trEMBL	Mascot
C2	tr|Q0PHP0|Beta-1 tubulin (Fragment).[Aedes aegypti]	4.67	48053	121	3	10	trEMBL	Mascot
C3	ISGI-TC2548		78110	90%	1	4	ISGI	Peaks
	ISGI-TC7696 similar to UP|Q5C5Z8_SCHJA (Q5C5Z8) SJCHGC09463 protein (Fragment), partial (14%)		73496	90%	1	3	BMGI	Peaks
C4	ISGI-G894P545FB3		70171	90%	1	3	ISGI	Peaks
	AVGI-TC352		63608	90%	1	6	AVGI	Peaks
C7	BMGI-TC12871		46956	90%	1	5	BMGI	Peaks
C8	tr|O97117|Glutathione S-transferase.[Boophilus microplus]	8.22	25588	65	1	4	trEMBL	Mascot

C9	ISGI-TC15028 homologue to UP|Q6P7N2_XENTR (Q6P7N2) MGC75629 protein (Eukaryotic translation initiation factor 2), partial (6%)		78825	92%	2	3	ISGI	Peaks
C11	ISGI-DN973124 UP|Q8FZR3_BRUSU (Q8FZR3) Oxidoreductase, FAD-binding, partial (2%)		25553	90%	1	15	ISGI	Peaks
	BMGI-BEAF903TF		37753	90%	1	6	BMGI	Peaks
D1	ISGI-G894P53RG3 similar to UP|Q3L6K7_9ACAR (Q3L6K7) Myosin alkali light chain protein, partial (92%)		77193	99%	4	12	ISGI	Peaks
D2	ISGI-TC4720 similar to UP|Q692U7_IXOSC (Q692U7) Salivary gland protein, partial (48%)		79756	91%	2	4	ISGI	Peaks
D3	tr|B0FL77|Myosin light chain 2.[Bombyx mandarina]	4.67	22029	57	2	3	trEMBL	Mascot
D4	RAGI-CD796888		67973	90%	1	4	RAGI	Peaks
D5	ISGI-TC1051 similar to UP|Q7Q3Z3_ANOGA (Q7Q3Z3) ENSANGP00000010457 (Fragment), partial (64%)		79028	90%	1	3	ISGI	Peaks
D7	tr|A8DX87|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (Fragment).[Phascolion strombus]	8.78	34366	87	1	5	trEMBL	Mascot
	tr|B0LJ59|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12).[Dictyocaulus viviparus]	8.16	36408	54	1	4	trEMBL	Mascot
D8	ISGI-G894P30FH22		26023	99%	3	9	ISGI	Peaks
D10	RAGI-CD796888		67973	90%	1	4	RAGI	Peaks
D11	tr|Q4PN07 Thioredoxin-dependent peroxide reductase.[Ixodes scapularis]	9.1	24394	58	1	8	trEMBL	Mascot
D12	BMGI-TC6444		128681	90%	1	3	BMGI	Peaks
E1	sp|P15159 Troponin C.[Tachypleus tridentatus]	4.11	17422	127	5	11	SwissProt	Mascot
E3	tr|Q4PM87|Nonmuscle myosin essential light chain.[Ixodes scapularis]	4.49	16507	152	2	15	trEMBL	Mascot
E4	tr|Q7SZL6|Beta-actin.[Monopterus albus]	5.31	41768	153	3	10	trEMBL	Mascot
E5	tr|A9QUS4|Beta-actin.[Rachycentron canadum]	5.31	41754	514	14	41	trEMBL	Mascot
E6	ISGI-TC12609 similar to UP|Q5UAQ6_3OMMO (Q5UAQ6) Ribosomal protein L32, complete		38783	99%	2	6	ISGI	Peaks
E7	ISGI-TC11435 similar to UP|Q38EU9_9TRYP (Q38EU9) Arginine kinase, partial (52%)		85067	90%	3	10	ISGI	Peaks
E8	ISGI-TC4720 similar to UP|Q692U7_IXOSC (Q692U7) Salivary gland protein, partial (48%)		79756	90%	3	13	ISGI	Peaks
E9	tr|Q1HQR4|RAB family GTPase (Rab11).[Aedes aegypti]	5.73	24319	72	1	6	trEMBL	Mascot
E10	tr|A0MQ61|Slow myosin heavy chain 1.[Danio rerio]	5.51	222828	59	1	0	trEMBL	Mascot
E11	BMGI-BEADC84TF		64704	90%	1	4	BMGI	Peaks
E12	tr|Q4PKE5|Actin 5.[Aedes aegypti]	5.3	41795	414	11	41	trEMBL	Mascot
F1	sp|P02595 Calmodulin (CaM).[Patinopecten sp.]	4.04	16802	400	11	54	SwissProt	Mascot
F2	ISGI-G893P364FK12 UP|Q8DQS9_STRR6 (Q8DQS9) Serine acetyltransferase, partial (5%)		74240	90%	1	6	ISGI	Peaks
F4	tr|Q96HG5|Actin, beta (Fragment).[Homo sapiens]	5.56	40978	792	37	52	trEMBL	Mascot
	tr|Q4PKE5|Actin 5.[Aedes aegypti]	5.3	41795	785	39	60	trEMBL	Mascot
F5	tr|Q5CAR2|Actin.[Ixodes ricinus]	5.3	41797	259	8	30	trEMBL	Mascot
F6	sp|P60524 Haemoglobin subunit beta (Haemoglobin beta chain) (Beta-globin).[Canis familiaris]	7.96	15986	68	2	13	SwissProt	Mascot

	ISGI-DN970843 homologue to UP|Q4PMM2_IXOSC (Q4PMM2) Salivary secreted serine protease, partial		89976	92%	2	4	ISGI	Peaks
F7	tr|Q6QWP0|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (Fragment).[Phormictopus sp. SBH266263]	6.42	31683	118	3	9	trEMBL	Mascot
F9	ISGI-TC3271 weakly similar to UP|Q70JV3_BRUMA (Q70JV3) Protein disulphide isomerase precursor		91033	94%	4	9	ISGI	Peaks
G2	AVGI-TC149		97165	99%	3	3	AVGI	Peaks
	RAGI-TC1455 similar to SP|Q96B10|ABE1_HUMAN ATP-binding cassette sub-family E member 1 (RNase L inhibitor)		157902	90%	2	2	RAGI	Peaks
G3	sp|Q1HR36|14-3-3 protein zeta.[Aedes aegypti]	4.78	28210	146	3	14	SwissProt	Mascot
G4	tr|Q963H2|Enolase (EC 4.2.1.11) (Fragment).[Taphrorychus bicolor]	5.46	31183	151	2	9	trEMBL	Mascot
G5	ISGI-G894P559FA16 homologue to UP|Q3R2U8_XYLFA (Q3R2U8) Glycoside hydrolase, family 24, partial (9%)		73049	90%	1	6	ISGI	Peaks
G6	RAGI-CD791968 similar to GP|19528647|gb|AAL90438.1||AY089700 SD10213p [Drosophila melanogaster]		71055	99%	2	3	RAGI	Peaks
G7	tr|Q291K6|GA10074-PA (Fragment).[Drosophila pseudoobscura]	8.66	42742	62	1	3	trEMBL	Mascot
	ISGI-G893P555RG6 similar to UP|Q692U7_IXOSC (Q692U7) Salivary gland protein, partial (28%)		82391	99%	2	4	ISGI	Peaks
G8	RAGI-CD787506		42875	90%	1	12	RAGI	Peaks
G8	RAGI-CD796888		67973	90%	1	4	RAGI	Peaks
G9	tr|Q06B75|Serpin-1 precursor.[Ixodes ricinus]	5.97	48495	70	1	2	trEMBL	Mascot
	AVGI-BM291400		51142	90%	1	9	AVGI	Peaks
	ISGI-G893P515RB5 similar to UP|Q9BVZ6_HUMAN (Q9BVZ6) CNOT1 protein (Fragment), partial (45%)		75241	90%	1	5	ISGI	Peaks
	ISGI-G893P52RG23 similar to UP|Q6IIG2_DROME (Q6IIG2) HDC18328, partial (12%)		75093	90%	1	3	ISGI	Peaks
G10	sp|P02769|Serum albumin precursor (BSA) (Allergen Bos d 6).[Bos taurus]	5.82	69248	324	8	13	SwissProt	Mascot
	BMGI-TC13064		62209	90%	2	7	BMGI	Peaks
	BMGI-TC12978		127860	90%	1	4	BMGI	Peaks
G11	sp|P02769|Serum albumin precursor (BSA) (Allergen Bos d 6).[Bos taurus]	5.82	69248	151	2	4	SwissProt	Mascot
	ISGI-G894P560RK1		74670	99%	3	3	ISGI	Peaks

Table 2.1: Protein identifications of spots cut from 2D gel separated whole unengorged female I.holocyclus extracted with 50mM Tris-HCl pH 8.8 + 150mM LiCl. Data files from the mass spectrometer were searched against the NCBI non-redundant database with Mascot. Data files that returned no results with Mascot were researched against a tick EST database compiled from the *Amblyomma variegatum, Rhipicephalus (Boophilus) microplus, Ixodes Scapularis* and *Rhipicephalus appendiculatus* gene indices (compbio.dfci.harvard.edu/tgi/). Spot numbers not listed did not return any results when searched by either search engine. An image of the gel annotated with the listed spot numbers can be found in the appendix.

Spot #	Accession number/Description/Species	pI	MW	Score	Matches	Coverage (%)	Database	Engine
A1	BMGI-TC6096 (BLASTX: cuticular protein [Tachypleus tridentatus - horseshoe crab])		141312	94%	4	6	BMGI	Peaks
A2	BMGI-TC6096 (BLASTX: cuticular protein [Tachypleus tridentatus - horseshoe crab])		141312	99%	18	7	BMGI	Peaks
	ISGI-TC7535 weakly similar to UP|Q3V6R9_TACTR (Q3V6R9) Cuticular protein (Fragment), partial (50%)		79631	99%	17	10	ISGI	Peaks
A3	50S ribosomal protein L7/L12 [Robiginitalea biformata HTCC2501]		13021	99%	4	14	NCBInr	Peaks
	TC7535 weakly similar to UP|Q3V6R9_TACTR (Q3V6R9) Cuticular protein (Fragment), partial (50%)		79631	99%	19	20	ISGI	Peaks
A4	tr|A0B6U3|Cell division protein pelota.[Methanosaeta thermophila]	5.91	38766	48	1	3	trEMBL	Mascot
A5	TC6096 (BLASTX: cuticular protein [Tachypleus tridentatus - horseshoe crab])		141312	99%	13	7	BMGI	Peaks
A6	sp|P02662|Alpha-S1-casein precursor (Contains: Antioxidant peptide).[Bos taurus]	4.96	24513	137	3	11	SwissProt	Mascot
A7	ribonuclease E [Parvularcula bermudensis HTCC2503]		112276	99%	3	8	NCBInr	Peaks
	TC8932 (BLASTX: putative secreted salivary gland peptide [Ixodes scapularis])		105773	99%	13	25	BMGI	Peaks
A8	transcriptional repressor protein MetJ [Photorhabdus luminescens subsp. laumondii TTO1]		12235	99%	17	4	NCBInr	Peaks
	TC8932 (BLASTX: putative secreted salivary gland peptide [Ixodes scapularis])		105773	99%	16	37	BMGI	Peaks
A9	TC8932 (BLASTX: putative secreted salivary gland peptide [Ixodes scapularis])		105773	99%	11	3	BMGI	Mascot
A11	tr|Q9DD57|Beta tubulin.[Chionodraco rastrospinosus]	4.74	49718	985	33	43	trEMBL	Mascot
A12	transcript:ENSGALT00000002103[Gallus gallus]	5.3	41607	1269	92	71	NCBI	Mascot
	tr|Q4PKE5|Actin 5.[Aedes aegypti]	5.3	41795	1231	95	70	trEMBL	Mascot
B1	sp|O97162|Tropomyosin.[Boophilus microplus]	4.7	32982	798	31	44	SwissProt	Mascot
B2	sp|O97162|Tropomyosin.[Boophilus microplus]	4.7	32982	81	1	5	SwissProt	Mascot
B3	sp|O97162|Tropomyosin.[Boophilus microplus]	4.7	32982	69	1	5	SwissProt	Mascot
B4	TC6096 (BLASTX: cuticular protein [Tachypleus tridentatus - horseshoe crab])		141312	99%	19	7	BMGI	Peaks
B5	tr|Q2U835|Ubiquitin carrier protein (EC 6.3.2.-).[Aspergillus oryzae]	9.05	20997	54	4	4	trEMBL	Mascot
B6	tr|Q09JF2|Cytochrome c oxidase subunit Va.[Argas monolakensis]	6.32	17753	200	5	11	trEMBL	Mascot
B7	tr|A1V933|Iron-sulfur cluster-binding protein, Rieske family.[Burkholderia mallei]	5.95	40195	55	1	2	trEMBL	Mascot
B8	G893P549RC21 similar to UP|Q692U7_IXOSC (Q692U7) Salivary gland protein, partial (14%)		76554	99%	3	6	ISGI	Peaks
B9	BM289816 CG11907 gene product {Drosophila melanogaster}		60372	99%	3	4	AVGI	Peaks
B10	ens transcript:ENSEEUT00000009411[Erinaceus europaeus]	6.16	56569	93	2	4	NCBInr	Mascot
B11	TC8489 similar to UP|Q08195_TOBAC (Q08195) Cysteine-rich extensin-like protein-2, partial (15%)		96285	99%	2	4	ISGI	Peaks
B12	TC12634		38481	99%	9	14	ISGI	Peaks
C1	tr|Q5CAR2|Actin.[Ixodes ricinus]	5.3	41797	163	3	11	trEMBL	Mascot
C2	tr|Q17C86|Actin.[Aedes aegypti]	5.3	41659	134	3	11	trEMBL	Mascot

C3	tr|A0F027|Beta cytoplasmic actin (Fragment).[Scophthalmus maximus]	5.08	17475	78	2	17	trEMBL	Mascot
C4	tr|Q4PM92|ATP synthase D chain.[Ixodes scapularis]	5.39	20082	104	3	11	trEMBL	Mascot
C6	tr|A0T2V0|Beta-actin [Penaeus monodon]	5	41973	143	3	9	trEMBL	Mascot
C7	ens transcript:ENSAPMT00000014142 [Apis mellifera]	8.83	31119	204	7	13	NCBInr	Mascot
	tr|Q1HR13|Prohibitin.[Aedes aegypti]	5.36	29885	147	6	10	trEMBL	Mascot
C8	tr|Q16EV9|Heat shock protein.[Aedes aegypti]	5.79	16611	57	1	9	trEMBL	Mascot
C10	CD785531 similar to SP|Q43822|PLSB_PHAVU Glycerol-3-phosphate acyltransferase, chloroplast precursor (GPAT). {Phaseolus vulgaris}		60558	99%	4	12	RAGI	Peaks
C11	>ISGI-G894P526RF12 homologue to predicted protein {Neurospora crassa}		75062	99%	4	3	ISGI	Peaks
C12	sp|Q6H1U7|Hemoglobin subunit beta (Hemoglobin beta chain) (Beta-globin). [Macropus eugenii]	7.1	16122	200	3	22	SwissProt	Mascot
D1	TC1525 weakly similar to UP|Q6PFB2_MOUSE (Q6PFB2) Rcc1 protein, partial (38%)		74067	99%	6	8	ISGI	Peaks
D2	G894P535FO5 weakly similar to UP|CRYM_MACFL (Q28488) Mu-crystallin, partial (32%)		69945	98%	6	16	ISGI	Peaks
D3	G893P516RI1		63899	93%	3	8	ISGI	Peaks
D4	G893P539FO19 weakly similar to UP|Q4J544_AZOVI (Q4J544) Amylo-alpha-1,6-glucosidase, partial (3%)		67405	99%	4	9	ISGI	Peaks
D5	TC10154		141526	95%	8	5	ISGI	Peaks
D10	tr|Q5T5J2|Rho GTPase activating protein 21.[Homo sapiens]	7.63	21273	67	2	4	trEMBL	Mascot
D12	tr|Q4PKE5|Actin 5.[Aedes aegypti]	5.3	41795	506	15	42	trEMBL	Mascot
E1	tr|Q5CAR2|Actin.[Ixodes ricinus]	5.3	41797	299	7	20	trEMBL	Mascot
	tr|A8E4K0|Troponin T.[Haemaphysalis qinghaiensis]	4.93	46442	286	11	10	trEMBL	Mascot
E3	G893P564RB16 weakly similar to UP|Q291D6_DROPS (Q291D6) GA10086-PA (Fragment), partial (12%)		66683	99%	5	11	ISGI	Peaks
E6	tr|Q4U3C7|Actin.[Orientobilharzia turkestanicum]	5.28	41720	214	5	19	trEMBL	Mascot
E7	sp|Q6H1U7|Hemoglobin subunit beta (Hemoglobin beta chain) (Beta-globin). [Macropus eugenii]	7.1	16122	67	1	8	SwissProt	Mascot
E8	tr|Q4PKE5|Actin 5.[Aedes aegypti]	5.3	41795	337	7	21	trEMBL	Mascot
	tr|Q17F.3|ATP synthase subunit beta (EC 3.6.3.14).[Aedes aegypti]	5.02	53940	331	4	11	trEMBL	Mascot
E9	tr|B0UTS7|AMP-dependent synthetase and ligase.[Haemophilus somnus]	8.85	63014	54	1	1	trEMBL	Mascot
F1	sp|P05661|Myosin heavy chain, muscle.[Drosophila melanogaster]	5.91	224328	346	5	3	SwissProt	Mascot
F2	sp|Q86RN8|Paramyosin.[Boophilus microplus]	5.53	101930	1616	28	32	SwissProt	Mascot
F3	sp|P05661|Myosin heavy chain, muscle.[Drosophila melanogaster]	5.91	224328	337	7	3	SwissProt	Mascot
F5	sp|P05661|Myosin heavy chain, muscle.[Drosophila melanogaster]	5.91	224328	551	9	4	SwissProt	Mascot
F6	tr|Q9N4N8|Putative uncharacterized protein.[Caenorhabditis elegans]	5.3	50252	56	2	1	trEMBL	Mascot
F7	tr|Q3LGW2|Actin.[Ornithodoros moubata]	5.3	41795	377	12	22	trEMBL	Mascot
F8	tr|Q4PKE5|Actin 5.[Aedes aegypti]	5.3	41795	320	8	23	trEMBL	Mascot

F9	tr|Q3LGW2|Actin.[Ornithodoros moubata]	5.3	41811	455	12	28	trEMBL	Mascot
F10	tr|Q4RG96|Chromosome 12 SCAF15104, whole genome shotgun sequence. (Fragment).[Tetraodon nigroviridis]	5.15	41159	356	8	25	trEMBL	Mascot
	tr|B0LF74|Beta actin.[Xestia c-nigrum]	5.46	41829	350	9	25	trEMBL	Mascot
G1	ens|ENSLAFP00000001043|ENSLAFG00000001234 transcript:ENSLAFT00000001234[Loxodonta africana]	8.07	58751	149	3	5	NCBInr	Mascot
G7	tr|Q5VKZ3|ATP synthase subunit alpha (EC 3.6.3.14) (Fragment).[Neoastelia spectabilis]	8.33	44817	228	4	8	trEMBL	Mascot
G9	sp|Q6H1U7|Hemoglobin subunit beta (Hemoglobin beta chain) (Beta-globin).[Macropus eugenii]	7.1	16122	76	2	15	SwissProt	Mascot
H6	tr|B0IW54|ATP synthase F1, alpha subunit (EC 3.4.21.68).[Rhizobium leguminosarum bv. trifoli WSM1325]	6.22	54616	54	1	3	trEMBL	Mascot

Table 2.2: Protein identifications of spots cut from 2D gel separated whole unengorged female I.holocyclus extracted with 7M urea, 2M thiourea, 1%C7BzO, 50mM Tris-HCl pH 8.8 + 150mM LiCl. Data files from the mass spectrometer were searched against the NCBI non-redundant database with Mascot. Data files that returned no results with Mascot were researched against a tick EST database compiled from the *Amblyomma variegatum, Rhipicephalus (Boophilus) microplus, Ixodes Scapularis* and *Rhipicephalus appendiculatus* gene indices (compbio.dfci.harvard.edu/tgi/). Spot numbers not listed did not return any results when searched by either search engine. An image of the gel annotated with the listed spot numbers can be found in the appendix.

This identified the proteins in a further 19 spots although only 10 of these were annotated to indicate what the ESTs were homologous to. The remaining 9 ESTs were BLAST searched to determine their homology to other proteins. Forty spots were not able to be identified by the employed MS/MS ion searching methodologies.

On visual inspection of the gel images in figure 2.6, it seems that the Tris extract has significantly more spots than the UTC7Tris extract. The Proteomeweaver software is able to detect very subtle differences in the Flamingo-stained spot density in the gel, which are difficult to determine with the naked eye. On manual inspection, more than 50% of the spots detected by Proteomeweaver in the UTC7Tris extract are very faint and well below the limit of detection by colloidal Coomassie blue and mass spectrometry. An approximately 10-fold increase in total protein would need to be loaded onto the IPG strip to identify these protein spots while also removing or depleting the very intense collection of spots at the acidic end of the gel. As stated above, this can be achieved by MCE fractionation and narrow range IPG's.

2.3.2: Fractionation by isoelectric point using a MultiCompartment Electrolyser.

Figure 2.7 shows 2-D gels of the UTC7Tris extract fractionated with a Multi-Compartment Electrolyser (MCE) and isoelectrically focused on narrow range IPG strips. Gel 2.6A is the protein present in the pH 3-5 chamber focused on a pH 3-6 IPG strip. The group of acidic proteins present on the gel of the entire sample separated on a pH 3-10 IPG strip (figure 2.6B) comprises approximately 13 very large spots. After fractionation of the sample with the MCE and loading the result on a narrow range strip, this group of acidic proteins is spread across the narrow pH range of the strip and resolved in more than 35 spots. However, the proteins in these spots did not match to any homologous proteins following trypsin digestion and analysis of the peptides by mass spectrometry followed by the previously employed MS/MS ion searching

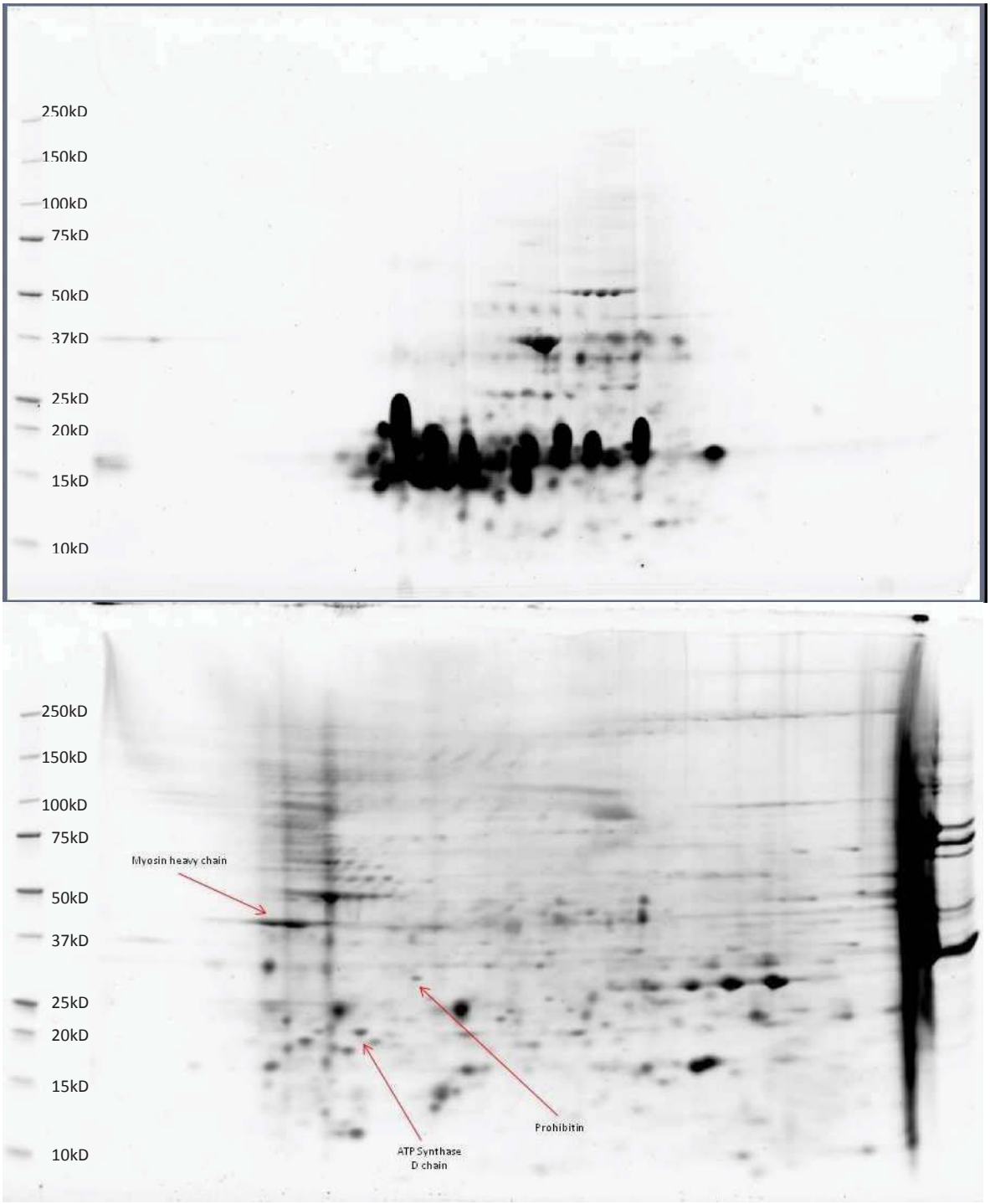

Figure 2.7: MCE separated, UTC7Tris extracted female unengorged *I.holocyclus*. Gel A (upper) is protein present in the pH 3-5 chamber of the MCE, isoelectrically focused on a pH 3-6 IPG strip. Approximately 158 individual protein spots were detected by Proteomeweaver. Gel B (lower) is protein present in the pH 5-8 chamber of the MCE, isoelectrically focused on a pH 5-8 IPG strip. Approximately 459 individual protein spots were detected by Proteomeweaver. SDS-PAGE separation of proteins in both IPG's was carried out on 4-12% Criterion XT BisTris gels with MES running buffer.

methodology with Mascot and Peaks ProteinID. *De novo* sequence determination was not attempted. A great number of spots not seen previously are now visible and this is repeated in all of the narrow pH range gels (figure 2.7B-D). The approximately 450 spots found by Proteomweaver in the UTC7Tris extracted samples separated on a pH 3-10 IPG has been resolved into approximately 700 spots using the MCE and narrow range IPG's. This number does not include the 350 protein spots found precipitated on the pH 5 membrane of the MCE that has been recovered and focused on a pH 4-7 IPG (figure 2.7D). A number of these spots are likely to appear in the gels of the pH 3-5 and pH 5-8 chambers. For example, Spot A3 from the pH 5-8 fraction and spot A9 from the pH 5 membrane (Table 2.3) identify to the same transcript from *Apis mellifera* with almost identical Mascot scores, number of matched peptides and protein coverage.

Without identification of all of the spots in these overlapping fractions (pH 3-5 chamber, pH 5 membrane, pH 5-8 chamber), it is difficult to ascertain how many unique proteins are present but it is estimated that by fractionating the sample with the MCE, the number of visible protein spots has approximately doubled from 450 to 900. This graphically illustrates the necessity to fractionate proteomes and not rely on the "one extract, one gel" approach previously employed in many proteomics projects. Analysis of protein spots following trypsin digestion and tandem mass spectrometry did not provide any extra identifications to those found on the wide, pH 3-10 IPG (figure 2.6). This is most likely due to more protein being needed to be loaded as reflected by the spot intensities on the gels. Protein concentration was not ascertained before loading the MCE or following concentration of the fractionated samples prior to isoelectric focusing.

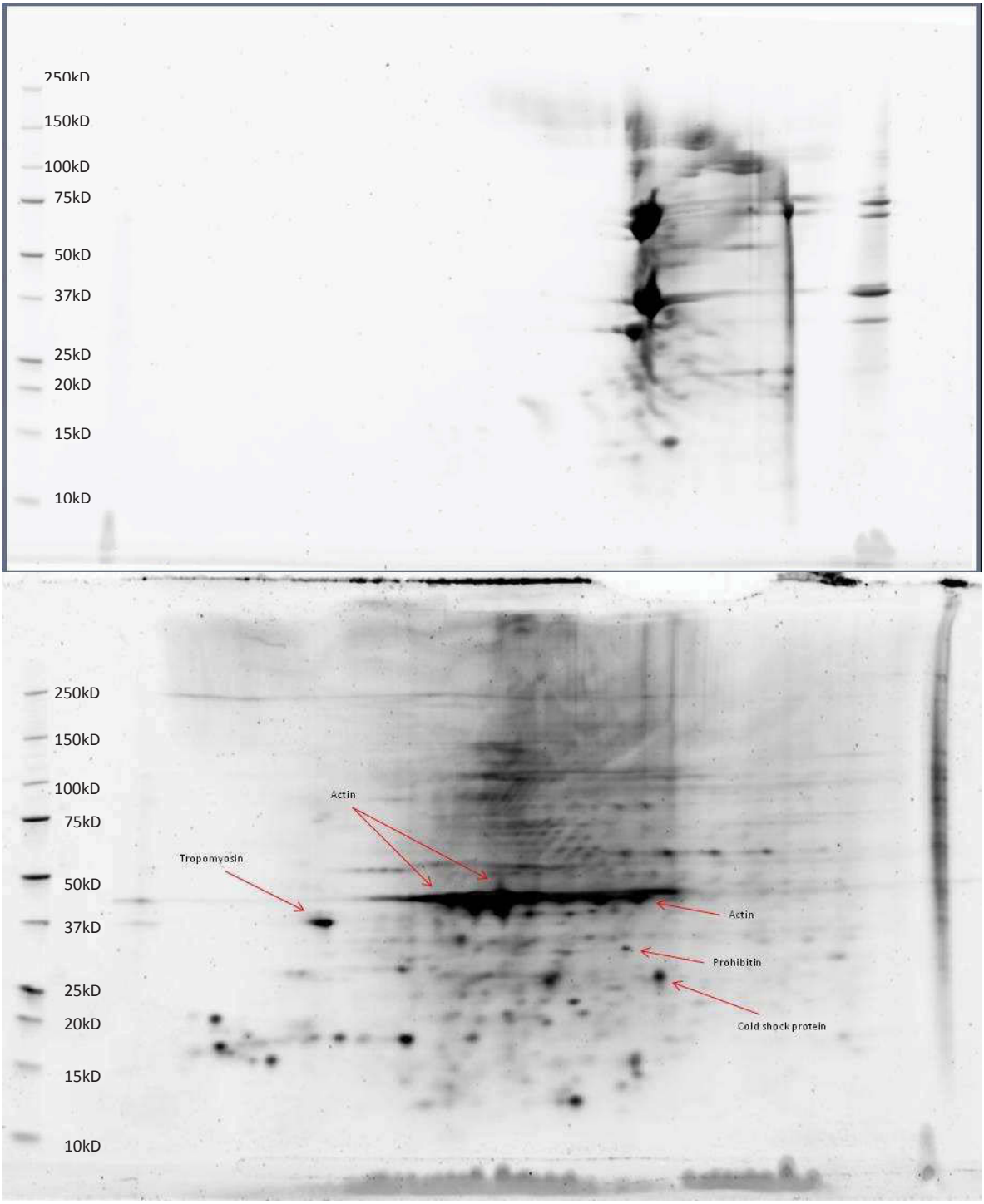

Figure 2.7: MultiCompartment Electrolyser (MCE) separated, UTC7Tris extracted female unengorged *I.holocyclus*. Gel C (upper) is protein present in the pH 8-11 chamber of the MCE, isoelectrically focused on a pH 6-11 IPG strip. Approximately 79 individual protein spots were detected by Proteomeweaver. Gel D (lower) is protein present on the pH 5 membrane of the MCE, isoelectrically focused on a pH 4-7 IPG strip. Approximately 356 individual protein spots were detected by Proteomeweaver. SDS-PAGE separation of proteins in both IPG's was carried out on 4-12% Criterion XT BisTris gels with MES running buffer.

Spot #	Description	pI	MW	Score	Matches	Coverage (%)	Database	Engine
	5-8 fraction							
A3	transcript:ENSAPMT00000014142 [Apis mellifera] Prohibitin. A subgroup of the band 7 domain of flotillin (reggie) like proteins.	8.83	31199	197	4	17	NCBInr	Mascot BLAST
A4	sp|P05661|Myosin heavy chain, muscle. [Drosophila melanogaster]	5.91	224328	302	4	2	NCBInr	Mascot
A5	tr|B0JYQ0|ALB protein. [Bos taurus]	5.95	69248	238	5	8	NCBInr	Mascot
B2	tr|Q2H0U0|Putative uncharacterized protein [Chaetomium globosum]	8.18	72255	67	2	1	NCBInr	Mascot
B3	tr|Q4PM92|ATP synthase D chain. [Ixodes scapularis]	5.39	20082	91	2	11	NCBInr	Mascot
B4	sp|P05661|Myosin heavy chain, muscle. [Drosophila melanogaster]	5.91	224328	308	4	2	NCBInr	Mascot
B5	sp|P02769|Serum albumin precursor (Allergen Bos d 6) (BSA). [Bos taurus]	5.82	69248	387	5	9	NCBInr	Mascot
C1	tr|Q17C86|Actin. [Aedes aegypti]	5.3	41659	74	2	7	NCBInr	Mascot
C4	sp|P05661|Myosin heavy chain, muscle. [Drosophila melanogaster]	5.91	224328	391	5	3	NCBInr	Mascot
	5 membrane fraction							
A7	sp|Q97162|Tropomyosin. [Boophilus microplus]	4.7	32982	552	9	21	NCBInr	Mascot
A8	tr|Q2FDF5|Cold shock protein. [Staphylococcus aureus]	4.77	8417	111	3	26	NCBInr	Mascot
A9	ENSAPMG00000008077 transcript:ENSAPMT00000014142 [Apis mellifera] (Prohibitin)	8.83	31199	200	4	17	NCBInr	Mascot
A10	tr|Q4PKE5|Actin 5. [Aedes aegypti]	5.3	41795	284	6	21	NCBInr	Mascot
B7	tr|Q4PKE5|Actin 5. [Aedes aegypti]	5.3	41795	735	20	48	NCBInr	Mascot
B8	tr|A0A4H8|Actin (Fragment). [Thaumatomonas sp. TMT002]	5.18	38618	75	2	7	NCBInr	Mascot
B9/10	sp|P02769|Serum albumin precursor (Allergen Bos d 6) (BSA). [Bos taurus]	5.82	69248	1095	26	25	NCBInr	Mascot
C7	tr|Q4PKE5|Actin 5. [Aedes aegypti]	5.3	41795	871	21	57	NCBInr	Mascot
C8	tr|A0ER78|Actin (Fragment). [Bicosoecida sp. MBIC11051]	5.22	35384	130	3	11	NCBInr	Mascot
C9	tr|A0A4H8|Actin (Fragment). [Thaumatomonas sp. TMT002]	5.18	38618	158	4	13	NCBInr	Mascot
D7	tr|A9QUS4|Beta-actin. [Rachycentron canadum]	5.31	41754	878	26	44	NCBInr	Mascot
G7	ens ENSXETP00000039264|ENSXETG00000018102 transcript:ENSXETT00000039264 [Xenopus tropicalis]	9.11	54293	68	2	1	NCBInr	Mascot
H7/8/9	sp|P02769|Serum albumin precursor (Allergen Bos d 6) (BSA). [Bos taurus]	5.82	69248	852	18	22	NCBInr	Mascot

Table 2.3: Protein identifications of spots cut from 2D gel separated whole unengorged female I.holocyclus extracted with 7M urea, 2M thiourea, 1% C7BzO, 50mM Tris-HCl pH 8.8 + 150mM LiCl, then separated with a Multi Comparment Electrolyser (MCE). MCE fraction is indicated in the table. Data files from the mass spectrometer were searched against the NCBI non-redundant database with Mascot. Data files that returned no results with Mascot were researched against a tick EST database compiled from the *Amblyomma variegatum, Rhipicephalus (Bcophilus) microplus, Ixodes Scapularis* and *Rhipicephalus appendiculatus* gene indices (compbio.dfci.harvard.edu/tgi/). Spot numbers not listed did not return any results when searched by either search engine. An image of the gel annotated with the listed spot numbers can be found in the appendix.

2.3.3: Comparison of two-dimensional gel electrophoresis with a 1-D SDS-PAGE/mass spectrometry approach to protein identification.

The previous experiments have shown that although fractionating by isoelectric point and focusing the fractionated proteins on narrow range IPG strips approximately doubles the number of protein spots, many spots are still too faint for identification. It is possible to excise all Flamingo-stained spots, but in practical terms and without the aid of robotics, the trypsin digestion of 400 spots in a gel is not realistic and the chances of keratin contamination increases with the number of samples and handling steps (Granvogl *et al.*, 2007a). In addition, 400 hours of mass spectrometry instrument time would be required to analyse the 400 tryptic digested spots from a single gel. In our lab, with single mass spectrometer and multiple projects, this is not possible. The use of 1-D-SDS-PAGE and 1-DLC/MS in a shotgun methodology is a complimentary technique to the 2-D-SDS-PAGE/1-DLC/MS approach used up to this point. By describing the method as shotgun it is necessary to point out that the sample is not as highly resolved as in 2-DGE, but all of the peptides of the same protein will be in the same sample analysed by 1-DLC/MS rather than spread across different fractions as in shotgun 2-DLC/MS methodologies. Although not reducing the amount of high abundance proteins, the digestion of all proteins present in the 1-D gel lane (and thus the sample), in combination with dynamic exclusion of previously selected peptides by the MS for 60 seconds, may reveal the identity of interesting low abundance proteins. The analysis of 15-20 slices of the 1-D gel lane is also much faster than analysing 400 individual spots (20 hours vs 400 hours of MS time). However, this method relies on the protein of interest's sequence being present in a database as unmatched individual peptides will not be able to be assigned to a certain protein present in the slice without prior knowledge of the complete sequence. *De novo* sequencing could be carried out on these peptides, which could then be clustered together for searching EST databases or BLAST searching of NCBI databases.

This experiment evaluates the usefulness of the 1-D-SDS-PAGE/1-DLC/MS methodology with unengorged tick samples. In a trial experiment, 100µg of the Tris-

HCl pH 8.8/150mM LiCl extract and the UTC7Tris-HCl pH 8.8 + 150mM LiCl extracted protein sample focused for figure 2.6A and B was loaded into separate single wells of a 4-12% Criterion XT gel with MES running buffer (as described in section 2.2.2). Figure 2.8 shows the flamingo-stained gel lanes. Pre-stained markers (10µL) were loaded in the adjacent lane and these were used as a guide to slice up the gel lane into 11 pieces using each marker as a cut point. The gel slices were then cubed into 1mm x 1mm pieces, washed, dehydrated and trypsin digested as described in section 2.2.5 although the volumes were increased due to the larger amount of gel pieces. After overnight digestion, the peptides were extracted as described in section 2.2.5 with an additional extraction step with 50% acetonitrile, 2% formic acid performed to ensure maximum recovery of the peptides. The samples were then analysed using the MS method described in section 2.2.6 at a flow rate of 300nL/min. A longer initial gradient was used (30% buffer B in 30 minutes) due to the greater number of peptides expected. Lengthening the gradient increases the time between the elution of peptides allowing the MS to select and fragment more peptides.

Table 2.4A and B shows the proteins identified from 100µg of sample. This amount was chosen as it was thought wise not to overload the 1-D gel lane causing "smearing" of high abundance proteins and other studies using a similar methodology report using this level of loading or less. However, as can be seen by the results table, only the high abundance proteins have been identified, especially actin which is present in many gel slices due to its high abundance. After removing the redundancy from the list of proteins, 35 proteins were found in the Tris extract and 15 proteins were found in the UTC7 extract. Being a complementary technique to 2-DGE, one would expect to see proteins in the list not identified from the 2-D gels due to the spot not being cut and analysed. This is the case with approximately 20 proteins being identified in this experiment, when both extracts results are combined and redundancy removed, that were not found in the 2-DGE experiments. However, the identification of such a small number of proteins was surprising when compared to other publications such as the

work of Breci *et al* (2005) who compared a number of different gel and chromatographic methodologies by analysing a yeast lysate. Using a similar 1-D SDS-PAGE/1-DLC/MS approach to that used in this work, nearly 900 individual proteins were found in 65µg of lysate. The difference in that experiment was that the gel lane was cut into 38 slices for trypsin digestion and 1-DLC/MS.

The experiment was repeated by separating 200ug of Tris and UTC7 extracted protein by 1-D SDS-PAGE and the gel lanes cut into 35 approximately 2mm slices which were further diced into 1 x 1mm pieces. The results of this experiment are shown in table 2.4C and D. This improved the number of proteins identified to 52 in the Tris extract and 95 in the UTC7 extract. But once again, high abundance cytoskeletal proteins dominate the list as their high abundance peptides were more frequently selected and fragmented by the mass spectrometer, even though fragmented peptides were excluded for 60 seconds before being selected again for fragmentation. Repeating the experiment with a much longer exclusion time (2-5 minutes) is an option for future consideration. An alternative to a longer exclusion time is the generation of an exclusion list of the ions/peptide masses of the peptides of high abundance proteins. This requires prior knowledge of the masses of the ion/peptide to be excluded and the ion/peptide's retention time. Both of these could be ascertained from these experiments.

This experiment identified a number of previously unseen proteins in any methodology employed thus far. Many of these were identified with 1 or 2 high scoring peptides. By combining the results of the two extracts (Tris and UTC7Tris) and removing redundancies, 80 proteins were identified using the 1-D SDS-PAGE/1-DLC/MS methodology, that were not identified by 2-DGE. In contrast, 60 proteins were identified by 2-DGE that were not found with 1-D SDS-PAGE/1-DLC/MS. Ribosomal proteins were overwhelmingly found by 1-D SDS-PAGE/1-DLC/MS. A single 2-DGE spot

was found to have homology to a ribosomal protein, while 24 ribosomal protein homologues were found by 1-D SDS-PAGE/1-DLC/MS, 21 of these from *I. scapularis*, two from *O. parkeri* and one from *R. microplus*. A significant number of homologous proteins with enzymatic activity were also identified using 1-D SDS-PAGE/1-DLC/MS, such as malate dehydrogenase, succinate dehydrogenase (both from *I. scapularis*), phospholipid-hydroperoxide glutathione peroxidise (*R. microplus*) phosphoglycerate kinase, fructose-bisphosphate aldolase and a mitochondrial peptidase (all from mosquitoes).

To complete the above data analysis, the mass spectrometer data files from the individual gel slices of a single gel lane were merged and searched with Mascot. This was done to reduce analysis time as by merging the data files, the higher abundance cytoskeletal proteins are reported as single hits in one output file. This kind of analysis could not be carried out with Peaks ProteinID algorithm, as the software does not allow the merging of separate data files into a single search. However, the data files were manually inspected using Peaks to determine if peptides were present that were not explained by the ProteinID results. These unmatched peptides would be from proteins with no sequence data in the NCBInr database. This analysis of the data provided very few unmatched peptides, which was unexpected.

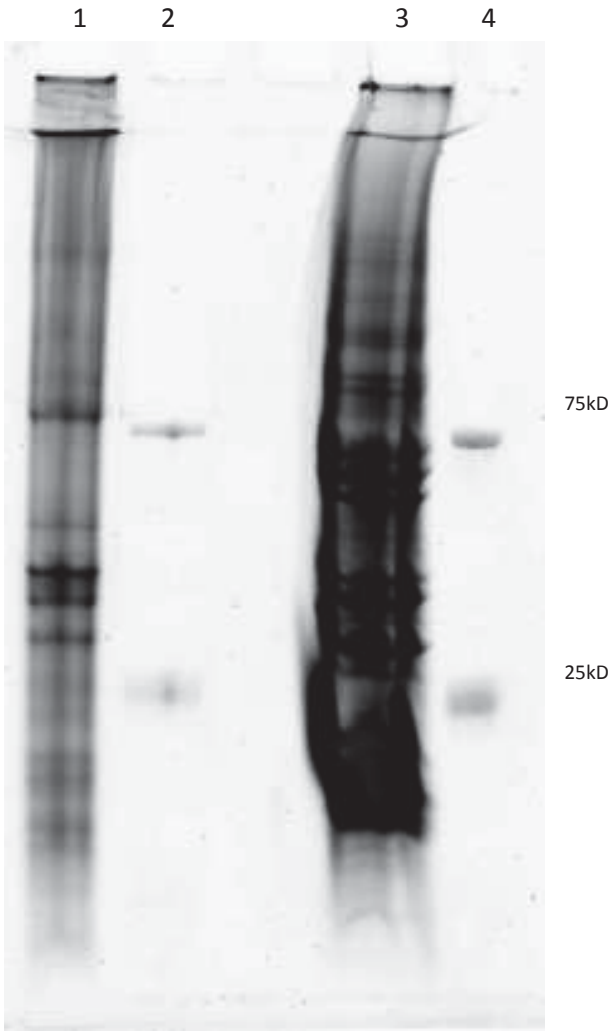

Figure 2.8: 1-D-SDS-PAGE separated unengorged female *I. holocyclus* protein extracts stained with Flamingo. Using concentrations determined in figure 2.2, 100µg of each sample was loaded into each well. The molecular weight (MW) markers loaded on the gel are pre-stained markers. Two of these markers have a pink dye bound to them, while the others have a blue dye bound. The protein markers with blue dye bound are not visualized with Flamingo fluorescent stain and hence are not seen on the image.

1) Tris-HCl pH 8.8/150mM LiCl extract.
2) MW markers
3) UTC7Tris-HCl pH 8.8/150mM LiCl extract.
4) MW markers.

Accession number/Description/Species	MW	Score	Matches	Unique peptides	Coverage (%)	Database	Engine	List position
sp|Q86RN8|Paramyosin.[Boophilus microplus]	101930	1726	68	28	30	SwissProt	Mascot	1
tr|Q4PKE5|Actin 5.[Aedes aegypti]	41795	1376	233	28	72	trEMBL	Mascot	2
sp|O97162|Tropomyosin.[Boophilus microplus]	32853	935	38	19	49	SwissProt	Mascot	32
sp|P05661|Myosin heavy chain, muscle.[Drosophila melanogaster]	224328	931	39	16	6	SwissProt	Mascot	33
ens transcript:ENSGACT00000018465[Gasterosteus aculeatus] (ATP synthase subunit beta (EC 3.6.3.14))	57348	509	16	7	17	NCBInr	Mascot	119
tr|Q4SKJ3|Chromosome undetermined SCAF14565, whole genome shotgun sequence.[Tetraodon nigroviridis]	25479	438	28	9	24	trEMBL	Mascot	143
tr|Q6W975|Sodium/potassium ATPase alpha subunit (Fragment).[Garypus californicus]	43358	392	8	8	22	trEMBL	Mascot	167
tr|A1C213|Beta-tubulin (Fragment).[Oxymonadida environmental sample]	41548	340	13	7	20	trEMBL	Mascot	197
sp|P0AA25|Thioredoxin-1 (Trx-1) (Trx).[Escherichia coli]	11799	267	27	5	47	SwissProt	Mascot	226
tr|Q8MV74|Myosin heavy chain type II (Fragment).[Cyrtophora citricola]	24985	236	7	3	6	trEMBL	Mascot	247
tr|Q16JS3|Tubulin alpha chain.[Aedes aegypti]	49890	224	17	4	11	trEMBL	Mascot	257
tr|Q17ER8|Histone h2a.[Aedes aegypti]	13386	214	15	4	37	trEMBL	Mascot	268
tr|Q171T4|Alpha-actinin.[Aedes aegypti]	103580	213	4	3	4	SwissProt	Mascot	269
tr|Q4GX79|Ribosomal protein Ubq/L40e.[Julodis onopordi]	14693	171	8	2	25	trEMBL	Mascot	295
ens transcript:ENSDART00000043743[Danio rerio]	71049	169	6	3	5	NCBInr	Mascot	298
tr|Q6QWP2|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (Fragment).[Lithobius sp. SBH266126]	32084	164	2	2	13	trEMBL	Mascot	304
tr|Q5AYB3|Malate dehydrogenase (EC 1.1.1.37).[Emericella nidulans]	37725	162	8	2	8	trEMBL	Mascot	311
tr|A6Y9G3|Elongation factor-1 alpha (Fragment).[Triops australiensis]	30112	149	9	3	16	trEMBL	Mascot	331
tr|Q5XXS7|Putative arginine kinase.[Oncometopia nigricans]	39958	129	7	3	10	trEMBL	Mascot	363
tr|O97117|Glutathione S-transferase.[Boophilus microplus]	25574	126	4	2	13	trEMBL	Mascot	366
tr|A0A9Q3|Class II myosin heavy chain (Fragment).[Molgula tectiformis]	150537	114	2	2	1	trEMBL	Mascot	388
tr|Q1HRQ7|Mitochondrial ATP synthase alpha subunit (ATP synthase alpha subunit mitochondrial).[Aedes aegypti]	59355	114	2	2	6	trEMBL	Mascot	389
tr|Q963H2|Enolase (EC 4.2.1.11) (Fragment).[Taphrorychus bicolor]	31183	114	3	3	9	trEMBL	Mascot	392
tr|A8P3E5|Fructose-bisphosphate aldolase (EC 4.1.2.13).[Brugia malayi]	39488	102	5	3	7	trEMBL	Mascot	421
sp|P02595|Calmodulin (CaM).[Patinopecten sp.]	16802	101	2	2	22	SwissProt	Mascot	422
tr|A7ANI1|Cytochrome c, putative.[Babesia bovis]	12509	74	1	1	10	trEMBL	Mascot	495
tr|B2X122|Serpin-2.[Spodoptera exigua]	41063	71	1	1	2	trEMBL	Mascot	509
tr|Q4PM87|Nonmuscle myosin essential light chain.[Ixodes scapularis]	16507	69	1	1	7	trEMBL	Mascot	515
tr|A3SA77|Putative uncharacterized protein.[Sulfitobacter sp. EE-36]	7137	68	6	1	12	trEMBL	Mascot	518

tr|A8X135|CBR-ARD-1 protein.[Caenorhabditis briggsae]	27215	67	1	1	5	trEMBL	Mascot	525
tr|Q291K6|GA10074-PA (Fragment).[Drosophila pseudoobscura]	42742	66	8	1	3	trEMBL	Mascot	527
tr|Q0V9W8|Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide.[Xenopus tropicalis]	28226	64	1	1	4	trEMBL	Mascot	541
tr|B0KZJ3|Allergen Aca s 8.[Acarus siro]	24017	59	1	1	4	trEMBL	Mascot	570
sp|Q47576|Acylphosphatase (EC 3.6.1.7) (Acylphosphate phosphohydrolase).[Thermobifida fusca]	10274	54	3	1	12	SwissProt	Mascot	617
tr|A0LLV9|A/G-specific DNA-adenine glycosylase (EC 3.2.2.-).[Syntrophobacter fumaroxidans]	44665	53	1	1	2	trEMBL	Mascot	623

Table 2.4A: Protein identifications from 1D SDS-PAGE/1DLC/MS analysis of unengorged female I. holocyclus protein extracted with 50mM Tris-HCl pH 8.8 + 150mM LiCl. 100ug of protein was loaded onto the 1D SDS-PAGE.

Accession number/Description/Species	MW	Score	Matches	Unique peptides	Coverage (%)	Database	Engine	List number
tr|Q4PKE5|Actin 5.[Aedes aegypti]		777	70	22	62	trEMBL	Mascot	1
sp|P05661|Myosin heavy chain, muscle.[Drosophila melanogaster]	224328	287	16	6	3	SwissProt	Mascot	56
tr|Q4SKJ3|Chromosome undetermined SCAF14565, whole genome shotgun sequence.[Tetraodon nigroviridis] Histone H2A	25479	269	14	5	24	trEMBL	Mascot	62
tr|Q6W975|Sodium/potassium ATPase alpha subunit (Fragment).[Garypus californicus]	43358	182	6	5	15	trEMBL	Mascot	92
tr|Q9U7Q1|Beta tubulin.[Trichuris trichiura]	49790	171	6	5	10	trEMBL	Mascot	97
sp|Q92GS8|ATP synthase subunit beta (EC 3.6.3.14) (ATPase subunit beta) (ATP synthase F1 sector subunit beta).[Rickettsia conorii]	51076	155	5	2	15	SwissProt	Mascot	103
tr|Q1HRQ7|Mitochondrial ATP synthase alpha subunit (ATP synthase alpha subunit mitochondrial).[Aedes aegypti]	59355	136	4	4	10	trEMBL	Mascot	110
ens transcript:ENSORLT00000015220[Oryzias latipes]	109295	136	3	2	3	NCBInr	Mascot	111
tr|Q95UM0|Elongation factor-1 alpha (Fragment).[Baeopterus philpotti]	35949	89	2	1	6	trEMBL	Mascot	140
sp|Q86RN8|Paramycsin.[Boophilus microplus]	101930	81	2	2	2	SwissProt	Mascot	147
tr|Q9GPM1|9.8 kDa basic protein.[Amblyomma hebraeum]	9852	67	1	1	16	trEMBL	Mascot	175
tr|A0BD42|Chromosome undetermined scaffold_10, whole genome shotgun sequence.[Paramecium tetraurelia]	199608	66	1	1	0	trEMBL	Mascot	179
ens|ENSCPOP00000006728|ENSCPOG00000007460 transcript:ENSCPOT00000007532[Cavia porcellus]	111196	64	4	1	1	NCBInr	Mascot	184
tr|Q11AW8|2-octaprenylphenol hydroxylase (EC 1.14.13.-).[Mesorhizobium sp.]	59152	56	26	1	2	trEMBL	Mascot	198

Table 2.4B: Protein identifications from 1D SDS-PAGE/1DLC/MS analysis of unengorged female I. holocyclus protein extracted with 7M urea, 2M thiourea, 50mM Tris-HCl pH 8.8 + 150mM LiCl. 100ug of protein was loaded onto the 1D SDS-PAGE.

Accession number/Description/Species	MW	Score	Matches	Unique peptides	Coverage (%)	Database	Engine	List position
sp|O86RN8|Paramyosin.[Boophilus microplus]	101930	1854	198	31	36	SwissProt	Mascot	1
tr|Q4PXE5|Actin 5.[Aedes aegypti]	41795	1277	415	23	68	trEMBL	Mascot	2
tr|Q6X4W2|Actin.[Rhipicephalus appendiculatus]	41781	1204	402	20	64	trEMBL	Mascot	3
sp_vs|P05661-16|(Mhc)Isoform C of P05661.[Drosophila melanogaster]	224278	1030	153	14	7	NCBInr	Mascot	20
tr|Q8T8B2|Beta-tubulin.[Bombyx mori]	50182	999	84	18	39	trEMBL	Mascot	29
tr|A9UCI2|ATP synthase subunit beta (EC 3.6.3.14) (Fragment).[Novocrania anomala]	46032	800	35	12	40	trEMBL	Mascot	81
tr|Q3TH56| heat shock protein 8.[Mus musculus] (HSP 70)	70828	714	76	10	18	trEMBL	Mascot	111
sp|O97162|Tropomyosin.[Boophilus microplus]	32982	690	68	13	29	SwissProt	Mascot	121
tr|Q4SKJ3|Chromosome undetermined SCAF14565 [Tetraodon nigroviridis] (Histone H2A)	25479	617	58	12	27	trEMBL	Mascot	160
tr|Q16JS3|Tubulin alpha chain.[Aedes aegypti]	49890	448	33	8	24	trEMBL	Mascot	270
sp|Q7PKQ5|Alpha-actinin, sarcomeric (F-actin cross-linking protein).[Anopheles gambiae]	106460	349	31	5	6	SwissProt	Mascot	365
tr|Q9BNV9|Elongation factor 1-alpha (Fragment).[Amblyomma sp. 'Amb2']	40075	321	41	5	15	trEMBL	Mascot	392
tr|Q2F5T3|ATP synthase subunit alpha (EC 3.6.3.14).[Bombyx mori]	59621	316	11	6	13	trEMBL	Mascot	397
tr|Q5XXS7|Putative arginine kinase.[Oncometopia nigricans]	39958	308	27	6	13	trEMBL	Mascot	406
tr|Q5DJT2|Histone H3 (Fragment).[Limnadopsis birchii]	9698	241	38	6	37	trEMBL	Mascot	496
tr|Q4PLZ0|Mitochondrial malate dehydrogenase (Fragment).[Ixodes scapularis]	17192	185	5	2	16	trEMBL	Mascot	602
tr|Q8T5G9|Triosephosphate isomerase (EC 5.3.1.1) (Fragment).[Archaeopotamobius sibiriensis]	24355	182	3	2	13	trEMBL	Mascot	610
tr|Q967G6|Myosin heavy chain (Fragment).[Trichinella spiralis]	64798	162	15	2	4	trEMBL	Mascot	659
tr|A8E4K0|Troponin T.[Haemaphysalis qinghaiensis]	46442	160	16	3	7	trEMBL	Mascot	665
tr|Q16PM9|Chaperonin-60kD, ch60.[Aedes aegypti]	60394	159	2	2	5	trEMBL	Mascot	668
tr|O35079|Polyubiquitin [Cricetulus griseus]	99531	155	9	3	4	trEMBL	Mascot	681
tr|Q17ER3|Histone h2a.[Aedes aegypti]	16708	153	8	3	31	trEMBL	Mascot	686
tr|Q6W975|Sodium/potassium ATPase alpha subunit (Fragment).[Garypus californicus]	43358	151	8	3	10	trEMBL	Mascot	692
tr|Q6QWP0|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (Fragment).[Phormictopus sp. SBH266263]	31683	144	21	2	8	trEMBL	Mascot	708
tr|Q7YZX3|Enolase (EC 4.2.1.11).[Onchocerca volvulus]	47123	141	4	2	9	trEMBL	Mascot	723
tr|B0FYV2|Eukaryotic initiation factor 4A.[Plutella xylostella]	47932	133	2	2	5	trEMBL	Mascot	746
tr|Q4PMB4|60S acidic ribosomal protein P0.[Ixodes scapularis]	34713	127	2	2	6	trEMBL	Mascot	770
tr|Q170J7|Moesin/ezrin/radixin.[Aedes aegypti]	69024	122	3	2	3	trEMBL	Mascot	787
tr|A0JMA1|Heat shock protein 90kDa alpha (Cytosolic), class A member 1, gene 2.[Xenopus tropicalis]	82389	109	2	2	3	trEMBL	Mascot	827

tr|B0L4I9|Glutathione S-transferase mu class (EC 2.5.1.18).[Rhipicephalus annulatus]	25574	96	16	2	13	trEMBL	Mascot	880
tr|Q29KW9|Phosphoglycerate kinase (EC 2.7.2.3) (Fragment).[Drosophila pseudoobscura]	43964	96	3	2	5	trEMBL	Mascot	881
sp|P02595|Calmodulin (CaM).[Patinopecten sp.]	16802	95	1	1	10	SwissProt	Mascot	887
tr|A9UYA2|Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8).[Monosiga brevicollis]	17827	89	3	1	15	trEMBL	Mascot	911
tr|Q64K84|Calreticulin.[Ixodes jellisoni]	47612	86	3	2	6	trEMBL	Mascot	925
tr|A6NA00|40S ribosomal protein SA.[Ornithodoros parkeri]	33045	81	5	1	10	trEMBL	Mascot	938
tr|Q17SU8|Fructose-bisphosphate aldolase (EC 4.1.2.13).[Aedes aegypti]	51671	78	6	1	4	trEMBL	Mascot	949
tr|Q1HDV2|Fructose-bisphosphate aldolase (EC 4.1.2.13).[Schistosoma japonicum]	39540	78	2	1	3	trEMBL	Mascot	952
tr|A5K6T9|Phosphoglycerate kinase (EC 2.7.2.3).[Plasmodium vivax]	45154	78	2	1	3	trEMBL	Mascot	953
tr|Q291K6|GA10074-PA (Fragment).[Drosophila pseudoobscura]	42742	77	22	1	3	trEMBL	Mascot	957
tr|Q7ZL32|Envelope glycoprotein (Fragment).[Human immunodeficiency virus 1]	18504	74	12	1	6	trEMBL	Mascot	963
sp|P00022|Cytochrome c.[Chelydra serpentina]	11600	74	1	1	10	SwissProt	Mascot	966
tr|Q8MWP3|Calreticulin.[Boophilus microplus]	47707	72	2	1	2	trEMBL	Mascot	977
tr|B2X122|Serpin-2.[Spodoptera exigua]	41063	71	5	1	2	trEMBL	Mascot	987
tr|Q0TR25|Putative alpha-N-acetylgalactosaminidase.[Clostridium perfringens]	178081	69	2	1	0	trEMBL	Mascot	1000
tr|B0WRR2|Vacuolar protein sorting 130.[Culex quinquefasciatus]	471954	69	16	1	0	trEMBL	Mascot	1006
tr|Q4PMC0|Ribosomal protein LP2.[Ixodes scapularis]	11448	64	1	1	12	trEMBL	Mascot	1046
tr|Q4PM16|60S ribosomal protein L23.[Ixodes scapularis]	14768	64	1	1	10	trEMBL	Mascot	1048
tr|Q4PMD5|Succinate dehydrogenase Ip subunit.[Ixodesscapularis]	31654	64	2	1	4	trEMBL	Mascot	1050
tr|Q4PM11|40S ribosomal protein S13.[Ixodes scapularis]	17107	64	2	1	7	trEMBL	Mascot	1053
tr|Q4PN07|Thioredoxin-dependent peroxide reductase.[Ixodes scapularis]	25674	63	1	1	4	trEMBL	Mascot	1063
tr|B0W1R7|Prohibitin-2.[Culex quinquefasciatus]	33146	63	1	1	3	trEMBL	Mascot	1064
tr|A8MLG2|Methionine aminopeptidase (EC 3.4.11.18).[Alkaliphilus oremlandii]	27323	58	16	1	3	trEMBL	Mascot	1104
tr|B1B544|Vitellogenin-C.[Haemaphysalis longicornis]	176180	53	10	1	0	trEMBL	Mascot	1203

Table 2.4C: Protein identifications from 1D SDS-PAGE/1DLC/MS analysis of unengorged female I. holocyclus protein extracted with 50mM Tris-HCl pH 8.8 + 150mM LiCl. 200ug of protein was loaded onto the 1D SDS-PAGE and the resulting lane cut into 35 slices.

Accession number/Description/Species	MW	Score	Matches	Unique peptides	Coverage (%)	Database	Engine	List position
sp|Q86RN8|Paramyosin.[Boophilus microplus]	101930	2446	256	40	38	SwissProt	Mascot	1
tr|Q4PKE5|Actin 5.[Aedes aegypti]	41795	1573	989	30	69	trEMBL	Mascot	2
sp|Q8ITB9|Tropomyosin.[Haemaphysalis longicornis]	32893	1226	139	25	55	trEMBL	Mascot	33
tr|Q9DD57|Beta tubulin.[Chionodraco rastrospinosus]	49718	999	66	25	47	trEMBL	Mascot	55
tr|Q179E8|Myosin heavy chain, nonmusde or smooth muscle.[Aedes aegypti]	221327	971	217	15	6	trEMBL	Mascot	59
tr|A7SHX4|Predicted protein.[Nematostella vectensis]	24545	855	243	16	35	trEMBL	Mascot	92
ens transcript:ENSDART00000043743[Danio rerio] (Heat shock 70 kDa protein 8)	71049	817	39	13	17	NCBInr	Mascot	100
tr|A9UCI2|ATP synthase subunit beta (EC 3.6.3.14) (Fragment).[Novocrania anomala]	46032	771	40	14	40	trEMBL	Mascot	122
sp|Q7PKQ5|Alpha-actinin, sarcomeric (F-actin cross-linking protein).[Anopheles gambiae]	106460	603	42	12	10	SwissProt	Mascot	210
tr|Q16JS3|Tubulin alpha chain.[Aedes aegypti]	49890	597	44	11	31	trEMBL	Mascot	213
tr|Q5XXS7|Putative arginine kinase.[Oncometopia nigricans]	39958	541	41	8	14	trEMBL	Mascot	350
tr|A5Y4F5|Histone H3a (Fragment).[Ostracoberyx dorygenys]	12368	396	67	8	59	trEMBL	Mascot	398
tr|Q2F5T3|ATP synthase subunit alpha (EC 3.6.3.14).[Bombyx mori]	59621	385	13	7	15	trEMBL	Mascot	410
tr|A0A9Q3|Class II myosin heavy chain (Fragment).[Molgula tectiformis]	150537	360	22	3	1	trEMBL	Mascot	435
tr|Q9BNV9|Elongation factor 1-alpha (Fragment).[Amblyomma sp. 'Amb2']	40075	343	43	6	15	trEMBL	Mascot	453
tr|A8E4K0|Troponin T.[Haemaphysalis qinghaiensis]	46442	338	53	8	12	trEMBL	Mascot	459
sp|Q7PPA5|Calcium-transporting ATPase sarcoplasmic/endoplasmic reticulum (EC 3.6.3.8).[Anopheles gambiae]	111788	330	15	8	9	SwissProt	Mascot	468
tr|Q16PM9|Chaperonin-60kD, ch60.[Aedes aegypti]	60755	328	8	5	14	trEMBL	Mascot	472
tr|Q6W975|Sodium/potassium ATPase alpha subunit (Fragment).[Garypus californicus]	43358	327	32	5	20	trEMBL	Mascot	475
tr|Q171J5|Histone h2a.[Aedes aegypti]	13215	293	46	5	31	trEMBL	Mascot	556
sp|P02595|Calmodulin (CaM).[Patinopecten sp.]	16802	277	7	4	38	SwissProt	Mascot	584
tr|Q4PM03|40S ribosomal protein S9.[Ixodes scapularis]	22573	266	8	6	19	trEMBL	Mascot	613
tr|Q4PM11|40S ribosomal protein S13.[Ixodes scapularis]	17199	260	6	5	27	trEMBL	Mascot	624
ens transcript:ENSFCAT00000005848[Felis catus]	222877	255	8	5	2	NCBInr	Mascot	633
sp|P13395|Spectrin alpha chain.[Drosophila melanogaster]	278132	250	6	4	2	SwissProt	Mascot	640
sp_vs|P54385-2||Gdh)Isoform C of P54385.[Drosophila melanogaster]	61043	247	4	4	10	NCBInr	Mascot	650
tr|A0A106|Enolase (EC 4.2.1.11) (Fragment).[Semiotellus sp. CD038]	41139	246	7	3	14	trEMBL	Mascot	652
tr|Q4PM67|Ribosomal protein S16.[Ixodes scapularis]	16615	244	6	4	23	trEMBL	Mascot	656
tr|A0PJ13|14-3-3-like protein.[Penaeus monodon]	27834	239	6	4	20	trEMBL	Mascot	665

tr|Q6WI29|Malate dehydrogenase (EC 1.1.1.37) [Branchiostoma belcheri tsingtauense]	35289	239	5	2	13	trEMBL	Mascot	668
tr|Q967G6|Myosin heavy chain (Fragment).[Trichinella spiralis]	64798	232	30	3	3	trEMBL	Mascot	688
tr|Q4PMD6|Ribosomal protein L30.[Ixodes scapularis]	12461	231	5	4	21	trEMBL	Mascot	690
tr|Q5NM89|ADP ribosylation factor 79F.[Aedes aegypti]	20675	223	7	6	33	trEMBL	Mascot	709
tr|O97117|Glutathione S-transferase.[Boophilus microplus]	25574	223	15	4	22	trEMBL	Mascot	712
tr|Q4PM30|Ribosomal protein S19.[Ixodes scapularis]	16509	212	6	3	14	trEMBL	Mascot	743
tr|A0JMA1|Heat shock protein 90kDa alpha (Cytosolic), class A member1, gene 2.[Xenopus tropicalis]	82389	203	4	4	7	trEMBL	Mascot	762
sp|Q4PM54|60S ribosomal protein L17.[Ixodes scapularis]	21329	202	3	3	11	SwissProt	Mascot	769
tr|A6N9R2|Ribosomal protein S18.[Ornithodoros parkeri]	17815	201	8	5	24	trEMBL	Mascot	770
tr|Q4PMB4|60S acidic ribosomal protein P0.[Ixodes scapularis]	34713	198	5	3	10	trEMBL	Mascot	790
sp|Q4PMB3|40S ribosomal protein S4.[Ixodes scapularis]	29590	180	6	5	17	SwissProt	Mascot	850
tr|A6N9Z4|40S ribosomal protein S3.[Ornithodoros parkeri]	26499	180	3	3	14	trEMBL	Mascot	852
tr|Q4PM87|Nonmuscle myosin essential light chain.[Ixodes scapularis]	16507	176	3	2	16	trEMBL	Mascot	878
tr|Q170J7|Moesin/ezrin/radixin.[Aedes aegypti]	69024	173	4	3	3	trEMBL	Mascot	883
ens transcript:ENSCINT00000018729[Ciona intestinalis]	34389	167	15	3	11	NCBInr	Mascot	904
tr|A3E3L5|Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8).[Pfiesteria piscicida]	17946	163	7	3	14	trEMBL	Mascot	915
tr|Q5DE07|SJCHGC09326 protein.[Schistosoma japonicum]	71818	162	2	2	5	trEMBL	Mascot	924
tr|Q64K84|Calreticulin.[Ixodes jellisoni]	47612	162	7	3	16	trEMBL	Mascot	925
tr|A9QQ34|Cysteine and glycine-rich protein.[Lycosa singoriensis]	11562	150	10	4	35	trEMBL	Mascot	980
tr|B0D3U0|Predicted protein.[Laccaria bicolor]	22596	148	4	2	18	trEMBL	Mascot	984
tr|Q4PM16|60S ribosomal protein L23.[Ixodes scapularis]	14768	148	7	3	25	trEMBL	Mascot	987
tr|Q7XY84|Histone H2B protein.[Griffithsia japonica]	13523	145	51	4	19	trEMBL	Mascot	998
tr|Q1MTI4|Triosephosphate isomerase (EC 5.3.1.1).[Danio rerio]	26836	141	6	3	11	trEMBL	Mascot	1016
sp|P15159|Troponin C.[Tachypleus tridentatus]	17422	140	5	4	7	SwissProt	Mascot	1025
tr|A6N9L6|40S ribosomal protein S28.[Ornithodoros parkeri]	7319	139	4	2	32	trEMBL	Mascot	1028
tr|Q0PZI5|Eukaryotic initiation factor 4A.[Callinectes sapidus]	48694	138	2	2	4	trEMBL	Mascot	1031
sp|Q4PM04|60S ribosomal protein L18.[Ixodes scapularis]	21603	134	2	2	14	SwissProt	Mascot	1064
tr|Q4PM10|Ribosomal protein S14.[Ixodes scapularis]	16110	133	3	2	15	trEMBL	Mascot	1065
tr|Q8II76|Alpha-2-macroglobulin precursor splice variant 1.[Ornithodoros moubata]	164969	133	8	2	1	trEMBL	Mascot	1067
tr|Q4PM82|Ribosomal protein S25.[Ixodes scapularis]	12828	128	3	2	18	trEMBL	Mascot	1096
tr|Q4PMD3|40S ribosomal protein S10.[Ixodes scapularis]	18073	127	7	3	24	trEMBL	Mascot	1099
tr|A0SHR2|Protein disulfide isomerase (EC 5.3.4.1).[Amblyomma variegatum]	54890	126	5	2	4	trEMBL	Mascot	1109

tr|A7RJJ7|Predicted protein.[Nematostella vectensis]	30761	125	3	2	10	trEMBL	Mascot	1112
tr|Q60WP2|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (Fragment) [Lithobius sp. SRH066126]	32084	122	5	2	15	trEMBL	Mascot	1122
tr|Q967M9|Enolase (EC 4.2.1.11) (Fragment).[Dryocoetoides cristatus]	40352	116	2	2	2	trEMBL	Mascot	1152
tr|A9QQC2|Cofilin[Lycosa singuriensis]	17106	115	7	3	14	trEMBL	Mascot	1158
tr|Q4PM37|Ribosomal protein L7-like.[Ixodes scapularis]	29529	115	3	3	10	trEMBL	Mascot	1159
tr|A2CEW4|5-aminoimidazole-4-carboxamideribonucleotide formyltransferase/IMP cyclohydrolase.[Danio rerio]	64017	106	2	2	4	trEMBL	Mascot	1200
tr|Q7PXW9|AGAP001633-PA.[Anopheles gambiae]	271590	106	12	3	0	trEMBL	Mascot	1202
tr|A6N9Z6|60s ribosomal protein L10.[Ornithodoros parkeri]	25493	100	2	2	10	trEMBL	Mascot	1223
tr|A6NA00|40S ribosomal protein SA.[Ornithodoros parkeri]	33045	99	2	2	10	trEMBL	Mascot	1228
tr|Q4PLY3|Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8).[Ixodes scapularis]	18072	99	2	2	9	trEMBL	Mascot	1230
ens|ENSAPMP00000028674|ENSAPMG00000013962 transcript:ENSAPMT00000028673 [Apis mellifera]	45278	97	2	2	4	NCBInr	Mascot	1242
tr|Q6B864|Cytochrome b5.[Ixodes pacificus]	15136	95	1	1	11	trEMBL	Mascot	1252
tr|A8P0E9|Isocitrate dehydrogenase (NAD) subunit alpha, mitochondrial, putative (EC 1.1.1.41).[Brugia malayi]	38754	93	1	1	4	trEMBL	Mascot	1260
tr|Q4PM66|Ribosomal protein L18a.[Ixodes scapularis]	20830	91	2	2	11	trEMBL	Mascot	1279
sp|Q4PMD1|60S ribosomal protein L36.[Ixodes scapularis]	8297	90	3	2	28	SwissProt	Mascot	1284
tr|Q6WNX0|Ferritn.[Ixodes scapularis]	19656	89	4	2	6	trEMBL	Mascot	1291
tr|Q16PZ7|Cytoplasmic dynein light chain.[Aedes aegypti]	14713	86	1	1	9	trEMBL	Mascot	1308
tr Q2XW13 Phospholipid-hydroperoxide glutathione peroxidase [Boophilus microplus]	19121	86	1	1	13	trEMBL	Mascot	1311
tr|Q5TMX9|AGAP012407-PA (Fragment).[Anopheles gambiae]	53099	86	1	1	2	trEMBL	Mascot	1314
tr|Q4PM07|Ribosomal protein L35.[Ixodes scapularis]	14328	85	2	1	10	trEMBL	Mascot	1316
tr|Q6J1M0|Rap 1A2 GTPase (MGC80662 protein).[Xenopus laevis]	21020	85	2	1	6	trEMBL	Mascot	1317
tr|Q5C3V3|SJCHG005011 protein (Fragment).[Schistosoma japonicum]	24043	84	1	1	4	trEMBL	Mascot	1323
tr|Q06875|Serpin-1 precursor.[Ixodes ricinus]	43906	80	5	1	7	trEMBL	Mascot	1359
tr|Q1HDV2|Fructose-bisphosphate aldolase (EC 4.1.2.13).[Schistosoma japonicum]	39540	80	3	1	3	trEMBL	Mascot	1360
tr|A5LHV9|Proteindisulfide isomerase-2.[Haemaphysalis longicornis]	55871	79	3	2	6	trEMBL	Mascot	1367
tr|Q6B8D1|Putative salivary secreted peptide [Ixodes pacificus]	10939	79	2	1	13	trEMBL	Mascot	1371
tr|Q17811|Vinculin.[Aedes aegypti]	107099	78	2	1	1	trEMBL	Mascot	1374
tr|B2D2C9|Ferritin.[Ornithodoros coriaceus]	20013	74	1	1	16	trEMBL	Mascot	1423
tr|A5EQ81|Hydrogenobyrinic acid a,c-diamide cobaltochelatase (EC 6.6.1.2).[Bradyrhizobium sp.]	128682	59	1	1	2	trEMBL	Mascot	1425
tr|A9P3C3|Ribosomal protein L23 (Fragment).[Boophilus microplus]	15450	73	1	1	9	trEMBL	Mascot	1431
tr|Q17A09|Mitochondrial processing peptidase beta subunit.[Aedes aegypti]	52216	68	1	1	2	trEMBL	Mascot	1498

tr|Q2HI18|Putative uncharacterized protein.[Chaetomium globosum]	76557	68	1	1	1	trEMBL	Mascot	1499
tr|Q4PLZ7|Probable microsomal signal peptidase 22 kDa subunit.[Ixodes scapularis]	20301	68	1	1	7	trEMBL	Mascot	1504
tr|A8X135|CBR-ARD-1 protein.[Caenorhabditis briggsae]	27215	68	2	1	5	trEMBL	Mascot	1508

Table 2.4D: Protein identifications from 1D SDS-PAGE/1DLC/MS analysis of unengorged female I. holocyclus protein extracted with 7M urea, 2M thiourea, 50mM Tris-HCl pH 8.8 + 150mM LiCl. 200ug of protein was loaded onto the 1D SDS-PAGE and the resulting lane cut into 35 slices.

2.3.4: Continuous elution preparative electrophoresis.

With such large amounts of protein being needed to load to have sufficient low abundance protein to be identified, the use of a Prep-Cell to fractionate the sample by molecular weight would seem to be a useful methodology. The Prep-Cell would allow high protein loads in the milligram range if necessary and due to its greater surface area for the proteins to enter the gel tube, smearing effects should be greatly reduced. This may prevent the cytoskeletal proteins being present in as many fractions as occurs in the 1-D SDS-PAGE/1-DLC/MS experiments. It would also provide a greater amount of material for other types of analysis such as 2-DLC or gas phase fractionation in the mass spectrometer (Breci *et al.*, 2005). The UTC7Tris extract was fractionated using a 10% acrylamide gel tube of 10cm in length. 154 fractions were collected over a 36-hour period and 10µL of every fifth fraction analysed by 1-D SDS-PAGE to determine how well the sample was fractionated. These gels are shown in figure 2.9 and show that fractionation has occurred as expected. The low intensity of the bands indicates that the fractions need to be combined and concentrated. This is necessary to firstly maximise the amount of protein available for trypsin digestion and mass spectrometry and secondly reduce the number of samples analysed by mass spectrometry to a manageable number. After combining and concentrating the samples to approximately 200µL, 10µL was removed and analysed by 1-D SDS-PAGE as shown in figure 2.9. The intensity of some of the bands present in this gel are not sufficient for further analysis and indicates that the starting amount of protein loaded into the Prep-Cell was not high enough or that some loss of protein has occurred through the process of fractionating and concentrating the sample. The collection of larger fraction volumes in fewer tubes and the precipitation of the fractionated proteins rather than the use of centrifugal concentrators may improve this.

It was envisaged that the high cross sectional surface area of the prep cell would be advantageous by preventing “smearing” of the high abundance cytoskeletal proteins. However, the Model 491 is in fact too large for the amount of sample available in this

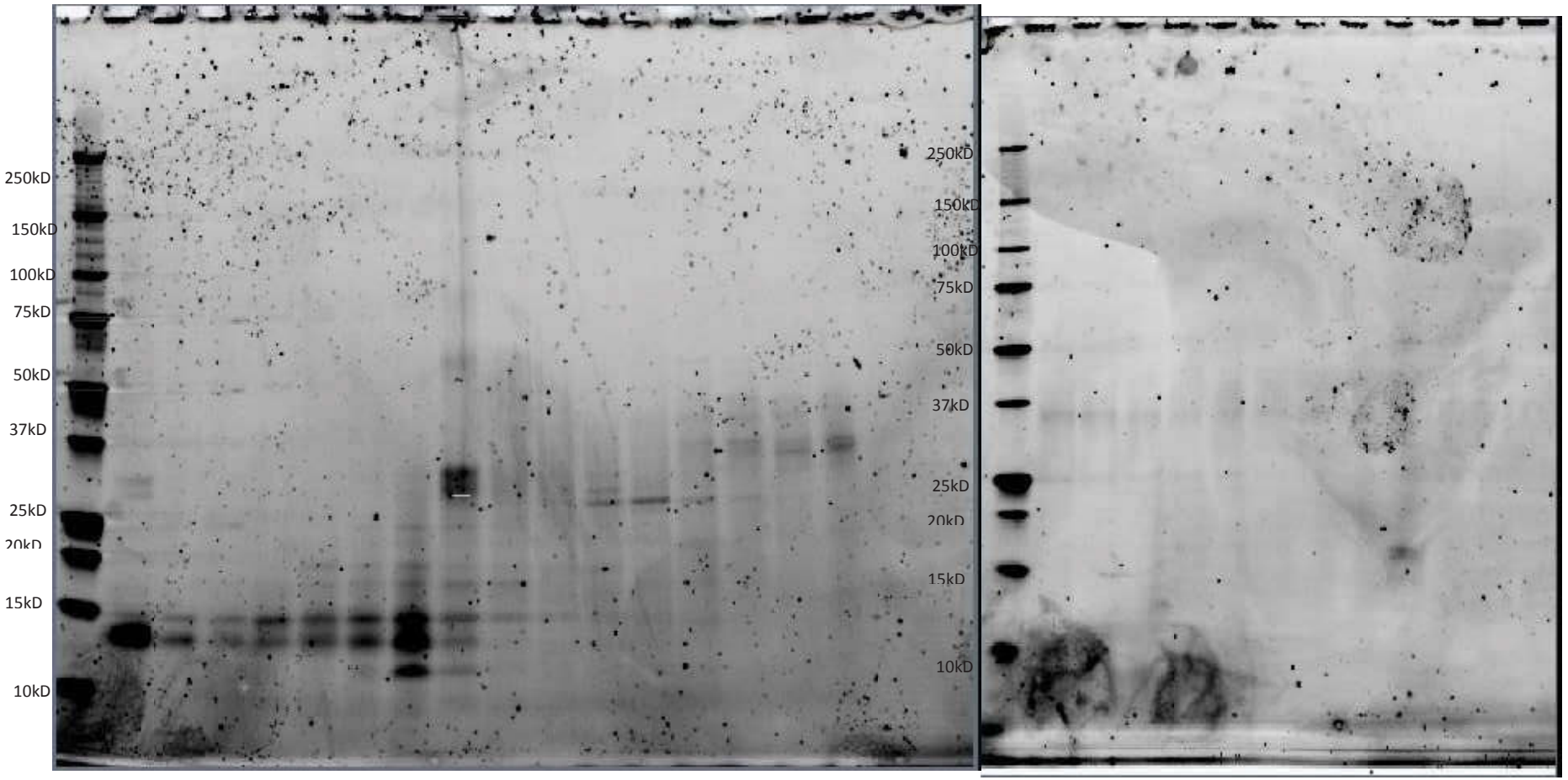

Figure 2.9: 1-D SDS-PAGE analysis of UTC7Tris-HCl pH 8.8 + 150mM LiCl extracted protein from female *I.holocyclus* fractionated with the Model 491 Prep-Cell. Speckling is due to colloidal Flamingo stain deposits on gel due to very low amount of protein present in gel which has been scanned at higher than usual PhotoMultiplier (PMT) voltage to enable visualization of proteins. The upper gel shows fractions from collection rack A (90 fractions) and the lower gel shows fractions from collection rack B (64 fractions). Fraction numbers are indicated on the gel image.

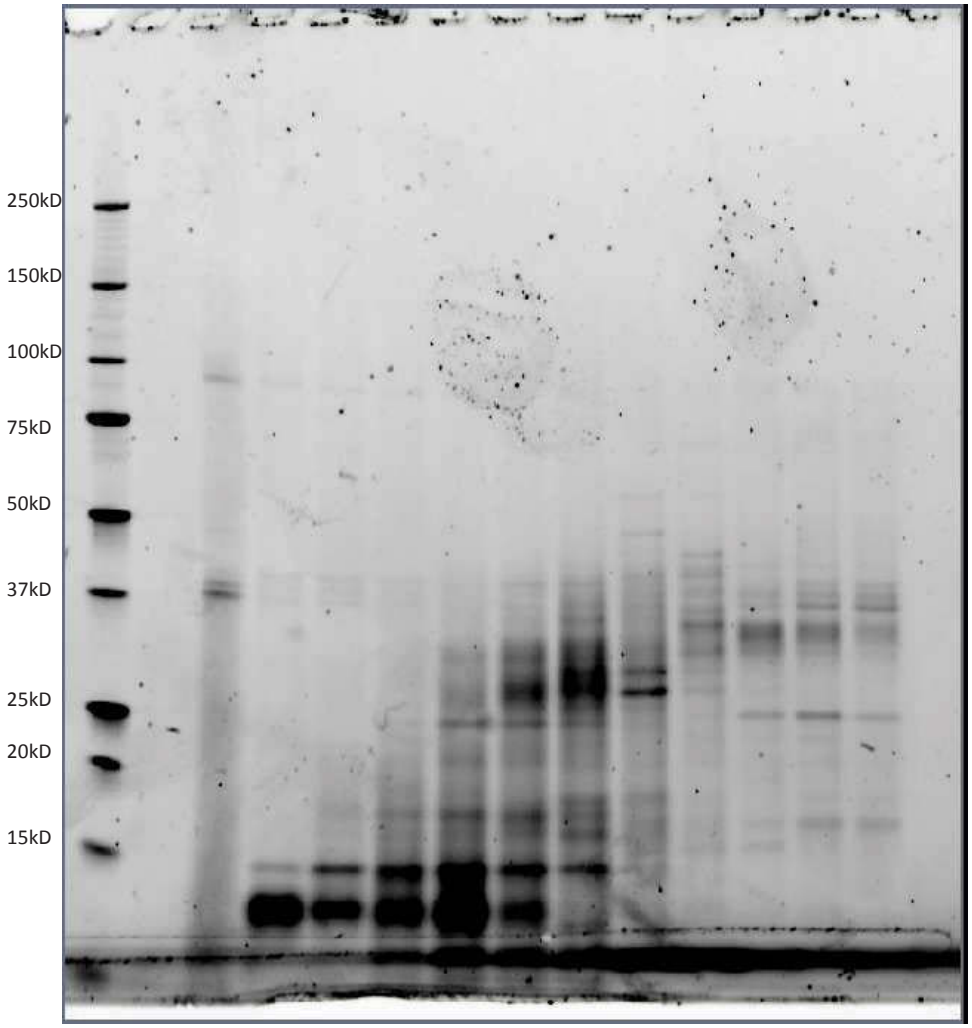

Figure 2.10: 1-D SDS-PAGE of combined and concentrated Prep-Cell fractions.

1) MW markers
2) Rack A, fraction 2-10
3) Rack A, fraction 11-17
4) Rack A, fraction 18-27
5) Rack A, fraction 28-41
6) Rack A, fraction 42-47
7) Rack A, fraction 48-53
8) Rack A, fraction 54-62
9) Rack A, fraction 63-76
10) Rack A, fraction 77-90
11) Rack B, fraction 8-19
12) Rack B, fraction 20-29
13) Rack B, fraction 30-39

study. An intermediate solution between the 1mm wide vertical slab gel and the Model 491 Prep Cell is required, but no such device was available. A lack of sample prevented this experiment from being repeated with this particular sample as it was reserved for other experiments. These experiments could be repeated with MCE fractionation of the sample prior to 1-D-SDS-PAGE or Prep Cell as described in section 2.1.1. A lack of sample also prevented this experiment from being performed.

2.3.5: Isoelectric fractionation using immobilised pH gradient strips and 1-D SDS-PAGE.

As described above, liquid phase isoelectric focusing using the MCE or Rotofor is a useful fractionation technique to be used prior to a number of other techniques including narrow range IPGs, 1-D SDS-PAGE and 1-DLC/MS. However it does have one major drawback and that is the large liquid volumes and relatively large sample sizes required in the currently available devices. The MCE has chamber volumes of 5mL and the fractionation of small amounts of sample can result in losses due to the need to concentrate the fractions by precipitation or membrane concentrators. Also, protein digestion in liquid has proven to have poor reproducibility in our laboratory compared to in gel digestion. This is likely due to subtle variations in the removal of solubilisation reagents, which cause proteins to precipitate or trypsin to lose activity if not properly removed. By having the protein “locked” in the gel piece, multiple washes and buffer changes can be performed without loss of protein. A variation of this, called Tube-Gel digestion (Lu and Zhu, 2005), has been employed to trypsin digest bacteriorhodopsin, a seven-transmembrane protein. When digested in solution, no peptides are found by mass spectrometry, but when the protein is polymerised in acrylamide, washed and in gel digested, five peptides are found provided 21% coverage of the protein.

One alternative for sample concentration ranges between midi sized (11cm) 2-DGE and the MCE is to use 18 or 24cm IPG strips and large format gels. The main problem with this approach lies in the second dimension gels. Large format precast gels are

expensive and not available in the long life BisTris buffer system. The casting and running of large format gels in-house is not a trivial exercise and requires a level of expertise and experience. To avoid wasting time, sample and reagents with large format gels, it was decided to use the loading capacity and high resolution of a 24cm IPG combined with the ease and robustness of 1-D SDS-PAGE in the midi format. One way of combining these is to cut the focused strip in half, trim the end of the strip where proteins rarely focus and perform the second dimension on two midi sized Criterion XT gels as a normal 2-D gel would be performed. Although this method has worked with success in our lab, it was decided it would not be used for reasons discussed in section 2.3.1, where it was noted that only a subset of visible spots are able to be excised and analysed.

The method employed first focuses the sample on a 24cm IPG strip allowing higher sample loads than 11cm IPG strips or 1-D SDS-PAGE without sacrificing resolution. The cutting of 1cm lengths means that fractions of approximately 0.3 of a pH unit can be isolated, offering far higher resolution that an MCE or Rotofor. After focusing, the strip was cut into 1cm long pieces, boiled in SDS buffer to elute the proteins from the gel, and the solution containing the proteins loaded into the wells of two 12 well 1-D gel. After electrophoresis and staining, the gel lanes can then be cut into pieces depending on band intensity rather than the regular 2mm slices employed in section 2.3.3, thus keeping the number of samples to be analysed by mass spectrometry to a reasonable number.

Tris extracted unengorged female tick protein (450µL) was used to passively rehydrate a 24cm pH 3-10 IPG strip for five hours prior to isoelectric focusing. Following cutting of the strip and SDS-PAGE as described above, the resulting gel is shown in figure 2.11 and dissected as indicated. The gels show distinct differences in band patterns in different gel lanes although there are three bands that are present in all gel lanes. Following isoelectric focusing, it was noticed that the IPG strip was swollen about one–

third of the strip length from the alkaline end. This is likely to be Tris-HCl in the sample that interferes with the focusing and causes the swelling as the protein and buffer in the strip build up as they can't pass the Tris molecules that concentrate around pH 8 due to the alkaline pK value of Tris (Cameron Hill, personal communication). The swollen part of the strip was loaded into lane 15 and there is a greater amount of protein present in this lane to the adjacent lanes, supporting this conclusion.

The sample was desalted using MicroBioSpins and the experiment was repeated. The IPG strip did not swell during isoelectric focusing in this experiment. The strip was cut, boiled and the proteins eluted loaded onto SDS-PAGE. The gels were stained with colloidal Coomassie blue prior to imaging. These gels are shown in figure 2.12. There seems to have been little effect on the three bands mentioned above and they still persist in all gel lanes. However, desalting the sample has changed the band pattern and intensity in the lanes loaded with proteins focused to the alkaline end of the strip. In gel trypsin digestion was carried out and the peptides analysed by mass spectrometry, the list of identified peptides shown in table 2.5.

On initial examination of the mass spectrometry data it can be seen that the isoelectric focusing was incomplete, causing horizontal streaking and therefore abundant proteins appeared in multiple lanes of the SDS gel. It may be that the IEF required greater than the 100 000 Vhours, or more probably, the 10kV voltage maximum of the instrument was insufficient to obtain high quality focusing. IEF is a voltage driven separation and the final sharpness of the focused bands is dependent on the applied voltage. In an 11cm IPG, focusing at 10kV, the voltage drop is approximately 1000V/cm, however, in a 24cm IPG at 10kV the voltage drop is approximately 400V/cm. An example of poor focusing on the 24cm IPG is indicated by previously identified proteins on midi-sized (11cm IPG) 2-D gels such as tropomyosin. On 11cm IPG 2-D gels, tropomyosin occurs as a single discreet spot and a second, more spread spot in the pI dimension around pH 5 (figure 2.6A & B). In this experiment,

tropomyosin is found over many pH units of the 24cm IPG strip, from lane 1 to lane 8 in figure 2.12, which is considerably more than can be accounted for by the increase in strip length. Much of the tropomyosin has focused into the dark band found in lane 7 around 37kDa (labelled 61 in the figure), but a significant amount of the protein is also found at the same molecular weight in lanes 1, 4 and 5 as shown by the number of peptides and coverage in table 2.5. A similar result is seen for Calmodulin, which is found in a band of the same molecular weight (labelled 29, 38 and 46) in lanes 3,4 and 5. Actin isoforms are found in a number of lanes, but the majority seems to be confined to two distinct bands in the SDS-PAGE, one in lanes 2, 3, 5, 7, 8, 9, 10 and 11 at approximately 41kDa (labelled 18, 27, 43, 61, 71, 77, 84 and 91), which is actin's reported molecular weight. The second actin isoform is present at approximately 85kDa in lanes 8 to 11 (labelled 66, 75, 82 and 89). These results suggest that, with increased focusing time, this method would isolate the high abundance cytoskeletal proteins to discreet bands as in 2-DGE, but provide the benefits of the 1-D SDS-PAGE/1-DLC/MS approach by including proteins below the limit of detection by current stains. Some highly alkaline histones have been bound by highly acidic molecules and focused to the acidic end of the strip. A similar result is found in chapter four with unengorged tick extract equalised by Proteominer. Both of these results could indicate incomplete reduction and alkylation of the sample and an extended focusing time is unlikely to move them back to the alkaline end of the strip.

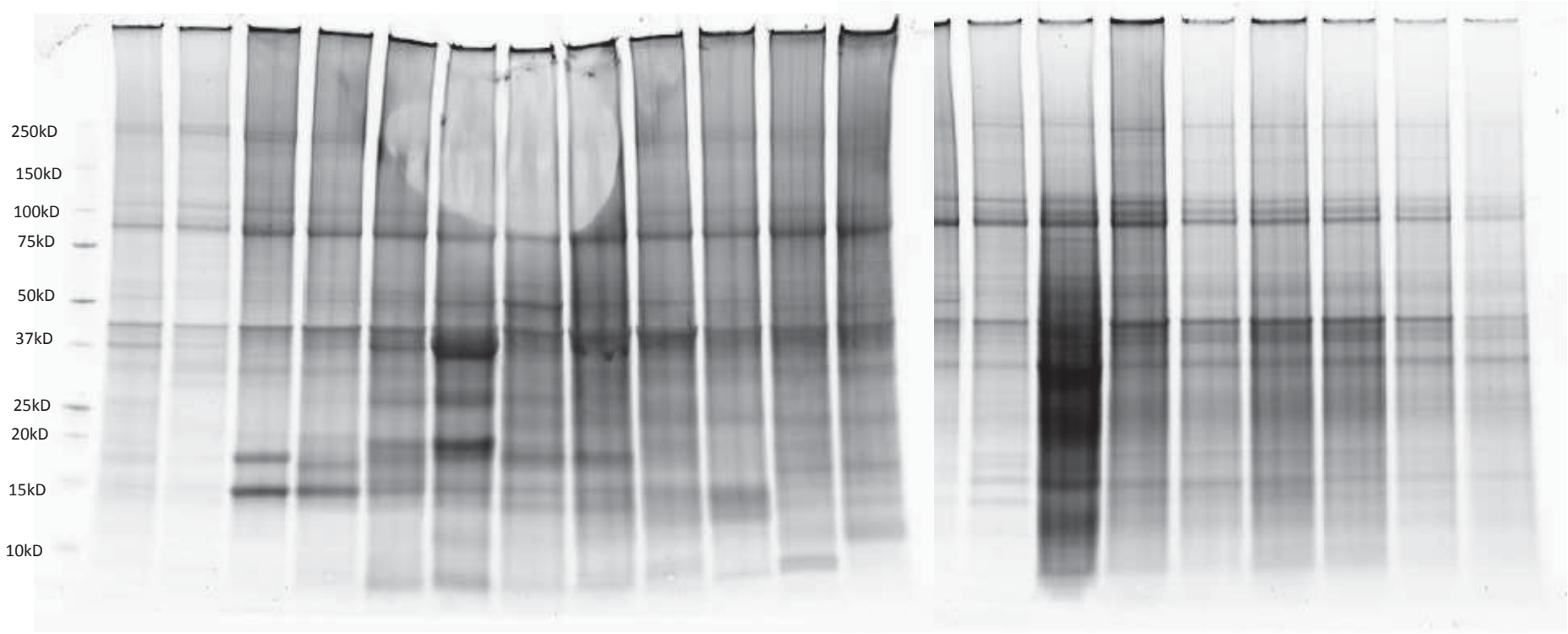

Figure 2.11: 1-D SDS-PAGE of Tris extracted female *I.holocyclus* protein separated by isoelectric focusing on a 24cm pH 3-10 IPG strip. Each well is loaded with the protein contained in an approximately 1cm piece of the IPG after boiling the IPG piece in SDS sample buffer. The strip has been cut sequentially so the pH gradient increases from left (pH 3) to right (pH 10) of the figure. Lanes are numbered to reflect the sequential 1cm dissection of the IPG from pH 3 to pH 10.

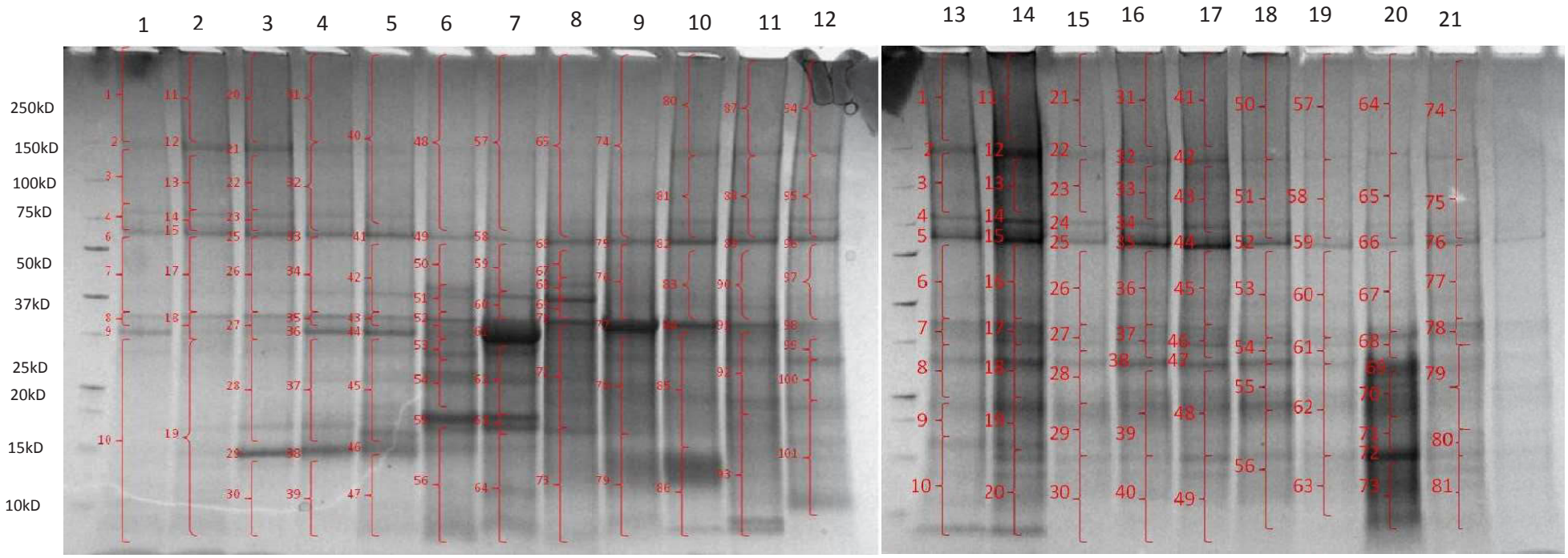

Figure 2.12: 1-D SDS-PAGE of Tris extracted female *I.holocyclus* protein separated by isoelectric focusing on a 24cm pH 3-10 IPG strip. The sample loaded onto the IPG had been desalted using a MicroBioSpin prior to passive rehydration. Each well is loaded with the protein contained in an approximately 1cm piece of the IPG after boiling the IPG piece in SDS sample buffer. The strip has been cut sequentially so the pH gradient increases from left (pH 3) to right (pH 10) of the figure. Lanes are numbered to reflect the sequential 1cm dissection of the IPG from pH 3 to pH 10.

Accession number/Description/Species	Bands found in	MW	pI	Highest Score	Matches	Unique peptides	Coverage (%)	Database	Engine
tr|Q179E8|Myosin heavy chain, nonmuscle or smooth muscle.[Aedes aegypti]	1,12,13,22,32,66,69,71,89,B15,B18,B39,B45,B50,B51	221327	5.76	348 (22)	4	4	2	trEMBL	Mascot
tr|B2ZTQ5|Muscle myosin heavy chain.[Loligo bleekeri]	11,20,21,22,80,81,B14,B32	221941	5.49	178 (11)	2	2	1	trEMBL	Mascot
sp|P05661|Myosin heavy chain, muscle.[Drosophila melanogaster]	67,68,72,75,76,80,81,82,83,84,87,88,90,91,B11,B12,B13,B15,B41,B42,B43	224328		574 (B12)	8	8	4	SwissProt	Mascot
tr|Q9U0S7|Myosin heavy chain (Fragment).[Mytilus galloprovincialis]	B31	197040	5.44	250	3	3	2		
tr|A0MQ61|Slow myosin heavy chain 1.[Danio rerio]	68,82	222828	5.51	86 (68)	1	1	0	trEMBL	Mascot
tr|Q4PMS7|Nonmuscle myosin essential light chain.[Ixodes scapularis]	46	16507	4.49	186	3	3	16	trEMBL	Mascot
sp|Q86RN8|Paramyosin.[Boophilus microplus]	4,12,13,14,21,22,32,71,72,76,77,78,80,81,82,83,87,88,89,90,B5,B11,B12,B13,B14,B15,B16,B20,B24,B32,B34,B42,B43	101930	5.53	1522 (B12)	26	22	32	SwissProt	Mascot
tr|Q5CAR2|Actin.[Ixodes ricinus]	7,8,13,17,26,27,32,33,34,36,37,41,42,43,44,45,50,51,53,59,60,61,62,64,66,67,68,69,71,73,75,77,79,81,82,85,88,89,90,B7,B8,B11,B12,B13,B14,B15,B16,B17,B18,B26,B27,B31,B34,B43,B44,B45,B46,B50,B77,B78	41797	5.3	868 (71)	44	14	57	trEMBL	Mascot
tr|Q4PKE5|Actin 5.[Aedes aegypti]	18,72,75,77,83,84,91	41795	5.3	1107 (77)	79	18	59	trEMBL	Mascot
sp|P53456|Actin-2.[Diphyllobothrium dendriticum]	22,54,B37	41746	5.39	453 (B37)	11	8	31	SwissProt	Mascot
tr|A9QUS4|Beta-actin.[Rachycentron canadum]	35	41754	5.31	584	17	17	32	trEMBL	Mascot
tr|Q172T4|Alpha-actinin.[Aedes aegypti]	88,89,B13,B15,B44	103580	5.42	141 (B15)	2	2	2	trEMBL	Mascot
tr|B0B5G3|Alpha-tubulin (Tubulin alpha-1).[Fasciola hepatica]	26,32,51,66,72,75,77,82,84,89,91	50031	4.97	91 (77)	1	1	3	trEMBL	Mascot
sp|Q8WQ47|Tubulin alpha chain (Allergen Lep d 7).[Lepidoglyphus destructor]	34,60,67,68,76,83,90	50007	5	316 (76)	4	4	13	SwissProt	Mascot
tr|Q16JS3|Tubulin alpha chain.[Aedes aegypti]	69,71	49890	5.01	610 (71)	13	9	33	trEMBL	Mascot
tr|Q8T8B0|Beta-tubulin.[Bombyx mori]	51	49782	4.73	186	2	2	7	trEMBL	Mascot
tr|Q8T8B2|Beta-tubulin.[Bombyx mori]	67,71,72,76	50182		394 (71)	7	7	19	trEMBL	Mascot
tr|A1C231|Beta-tubulin (Fragment).[Cryptocercus punctulatus]	60,68	41548	5.67	396 (68)	6	5	20	trEMBL	Mascot
tr|Q6S3D4|Beta-2 tubulin.[Laodelphax striatellus]	69	50268		1092	37	22	44	trEMBL	Mascot

sp|O97162|Tropomyosin.[Boophilus microplus]	9,18,27,34,35,36,37,43,44,45,51,53,54,55,59,60,61,62,63,64,65,66,67	32982	4.7	1141(61)	97	22	42	SwissProt	Mascot
tr|Q32VZ5|Calmodulin (Fragment).[Euchelota baker]	29	14855	4.11	211	2	1	12	trEMBL	Mascot
sp|P11121|Calmodulin (CaM).[Pyuridaeso.]	38	16801	4.06	286	4	4	55	SwissProt	Mascot
tr|Q98UH8|Calmodulin (Fragment).[Clemmys japonica]	46	15349	4.05	54	1	1	11	trEMBL	Mascot
sp|P15159|Troponin C.[Tachypleus tridentatus]	29	17422	4.11	208	4	4	7	SwissProt	Mascot
sp|Q09665|Troponin C, isoform 2.[Caenorhabditis elegans]	38	18216	4.04	60	2	1	10	SwissProt	Mascot
tr|Q4SKJ3|Chromosome undetermined SCAF14565, whole genome shotgun sequence.[Tetraodon nigroviridis] (Histone H2B)	29,79,86,820,849,861	25479	10.7	370 (861)	7	4	24	trEMBL	Mascot
tr|Q17EF0|Histone h2a.[Aedes aegypti]	29,38,46,872	13404	10.6	106 (29)	2	2	25	trEMBL	Mascot
tr|Q4PM63|Histone H2B.[Ixodes scapularis]	85,89,839,840,863,872,880	13811	10.5	175 (880)	3	2	26	trEMBL	Mascot
ens transcript:AGAP012871-RA[Anopheles gambiae] (Histone H2B)	38,46,55,56,63,819,871,873	7459	9.86	342 (46)	6	6	35	NCBInr	Mascot
tr|A7S4X9|Predicted protein.[Nematostella vectensis] (Histone H2A)	64	23541		88	2	2	7	trEMBL	Mascot
tr|Q0PXZ8|Putative 60S acidic ribosomal protein P1.[Diaphorina citri]	29,38	11451	4.24	65 (38)	2	1	14	trEMBL	Mascot
tr|Q17EE8|Histone H4.[Aedes aegypti]	30	11404	11.4	57	1	1	7	trEMBL	Mascot
tr|B2SU79|Chainlength determinant protein.[Xanthomonas oryzae pv. oryzae PXO99A]	18	31172	10.9	56	1	1	3	trEMBL	Mascot
ens transcript:ENSSART00000003513[Sorex araneus]	22	20339	10.1	69	1	1	6	NCBInr	Mascot
tr|A7C535|Putative uncharacterized protein.[Beggiatoa sp. PS]	25	8262	9.94	50	1	1	15	trEMBL	Mascot
tr|Q2F5T3|ATP synthase subunit alpha (EC 3.6.3.14).[Bombyx mori]	76,845	59621	9.21	152 (845)	2	2	6	trEMBL	Mascot
tr|Q6PTP3|ATP synthase subunit beta (EC 3.6.3.14) (Fragment).[Encope michelini]	69	45811	4.96	1063	30	17	50	trEMBL	Mascot
tr|Q17FL3|ATP synthase subunit beta (EC 3.6.3.14).[Aedes aegypti]	34,56,64,67	53940	5.02	160 (34)	2	2	5	trEMBL	Mascot
sp|Q0A4M8|ATP synthase subunit beta (EC 3.6.3.14) [Alkalilimnicola ehrlichei]	55	49544	4.92	66	1	1	4	SwissProt	Mascot
ens transcript:ENSGACT00000018465[Gasterosteus aculeatus] (ATP synthase subunit beta)	42,51,60,61,62,68,71,72,76	57348	5.22	917 (60)	21	13	33	NCBInr	Mascot
tr|Q5MCG8|Mitochondrial ATP synthase beta subunit (Fragment).[Enchytraeus buchholzi]	43	21285	5.06	95	1	1	6	trEMBL	Mascot
tr|Q5MCG6|Mitochondrial ATP synthase beta subunit (Fragment).[Theromyzon tessulatum]	54	25949	4.86	143	2	2	13	trEMBL	Mascot

tr|Q9XYC8|Vacuolar ATPase B subunit (ATP synthase subunit beta vacuolar).[Aedes aegypti]	68,69	55181	5.38	68	1	1	2	trEMBL	Mascot
tr|B0Y0Q5|ATP synthase subunit beta (EC 3.6.3.14).[Aspergillus fumigatus]	B45,B50,B52	55586		341 (B45)	5	5	14	trEMBL	Mascot
tr|A2ETL5|Putative uncharacterized protein.[Trichomonas vaginalis G3]	55	59634	5.71	62	1	1	2	trEMBL	Mascot
tr|Q4PHP7|Putative uncharacterized protein.[Ustilago maydis]	44,53	107754	4.9	59	1	1	1	trEMBL	Mascot
tr|A5BZ64|Putative uncharacterized protein.[Vitis vinifera]	72	61202	5.73	57	1	1	2		
tr|Q11AW8|2-octaprenylphenol hydroxylase (EC 1.14.13.-).[Mesorhizobium sp.]	46	59152	9.59	55	6	1	2	trEMBL	Mascot
tr|Q16PM9|Chaperonin-60kD, ch60.[Aedes aegypti]	67,68,69,76	60394	5.47	195 (67)	3	3	5	trEMBL	Mascot
tr|Q0GFD6|60 kDa chaperonin (Fragment).[uncultured Bartonella sp]	60,68,69	34085	5.07	108	1	1	5		
tr|O96783|Heat shock protein 60 (HSP60).[Plectus acuminatus]	50	61756	6.09	60	1	1	3	trEMBL	Mascot
tr|A0PA16|Heat shock protein 70kDa.[Cotumix cotumix japonica]	67	70827	5.37	338	5	5	9	trEMBL	Mascot
tr|A9BK82|Hsp70.[Cryptophyta]	97	71809	5.11	97	2	2	3	trEMBL	Mascot
sp|O73885|Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8).[Gallus gallus]	90	70783	5.47	314	6	4	12	SwissProt	Mascot
tr|B0XPW1|Molecular chaperone Hsp70.[Aspergillus fumigatus]	B52	69618	5.09	123	2	2	3	trEMBL	Mascot
tr|A0A106|Enolase (EC 4.2.1.11) (Fragment).[Semotellus sp. CD038]	51,69,83,90,91,B16,B17,B45	41139	5.82	168 (90)	3	2	14	trEMBL	Mascot
tr|A0SNX4|Enolase (EC 4.2.1.11).[Trimastix pyriformis]	91	47984	5.81	70	1	1	2	trEMBL	Mascot
tr|Q963H2|Enolase (EC 4.2.1.11) (Fragment).[Taphrorychus bicolor]	83,90	31183	5.46	164 (83)	2	2	9	trEMBL	Mascot
tr|Q7YZX3|Enolase (EC 4.2.1.11).[Onchocerca volvulus]	90	47123	6.04	156	3	2	7	trEMBL	Mascot
tr|Q8MWP3|Calreticulin.[Boophilus microplus]	51	47638	4.5	111	2	2	4	trEMBL	Mascot
tr|Q64K84|Calreticulin.[Ixodes jellisoni]	51,59	47612	4.57	66 (51)	2	1	4	trEMBL	Mascot
tr|Q7KLE5|Amphiphysin (CG8604-PA) (LD19810p).[Drosophila melanogaster]	54	65862	5.14	74	1	1	1	trEMBL	Mascot
tr|B0L419|Laminin receptor (Fragment).[Oryctolagus cuniculus]	61	28796	4.92	83	1	1	5	trEMBL	Mascot
tr|Q0V9W8|Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide.[Xenopus tropicalis]	62,72	28226	4.8	155 (62)	4	3	8	trEMBL	Mascot
sp|Q1HPS0|Myosin regulatory light chain 2 (MLC-2).[Bombyx mori]	62	22029	4.67	62	2	1	3	SwissProt	Mascot
tr|A3WG27|Putative uncharacterized protein.[Erythrobacter sp. NAP1]	62	7169	4.77	59	2	1	14	trEMBL	Mascot
tr|A5LHV9|Protein disulfide isomerase-2.[Haemaphysalis longicornis]	68	56871	4.84	135	2	2	5	trEMBL	Mascot

tr|A3EXM7|Putative ubiquilin 1,2 (Fragment).[Maconellicoccus hirsutus]	68	42519	4.68	96	1	1	5	trEMBL	Mascot
tr|B0D1E3|Predicted protein.[Laccaria bicolor]	71	45229	5.05	76	1	1	2	trEMBL	Mascot
tr|B2X122|Serpin-2.[Spodoptera exigua]	71,72,77,84,85,91,87	41063	4.74	72 (71)	1	1	2	trEMBL	Mascot
tr|Q2JXM5|Peptidase inhibitor, I4 family.[Synechococcus sp.] (Serpin-2)	83,90	48495	5.97	78 (90)	1	1	2	trEMBL	Mascot
ens transcript:ENSAPMT00000018309 [Apis mellifera]	71	42700	9.48	57	1	1	1	NCBInr	Mascot
tr|A6NVK8|DNA polymerase I (EC 2.7.7.7).[Bacteroides capillosus ATCC 29799]	71	98376	5.13	56	1	1	1	trEMBL	Mascot
tr|Q291K6|GA10074-PA (Fragment).[Drosophila pseudoobscura]	72,77,91	42742	8.66	65 (72)	2	1	3	trEMBL	Mascot
tr|Q17A09|Mitochondrial processing peptidase beta subunit.[Aedes aegypti]	76,77,83	52216	5.87	87 (77)	1	1	2	trEMBL	Mascot
tr|Q9XYC8|Vacuolar ATPase B subunit (ATP synthase subunit beta vacuolar).[Aedes aegypti]	76	55181	5.38	65	1	1	2	trEMBL	Mascot
tr|Q2T5I8|Gp51.[Burkholderia thailandensis]	76	42016	8.38	57	1	1	1	trEMBL	Mascot
tr|Q1HPK0|Vesicle amine transport protein.[Bombyx mori]	83	48790	5.85	61	1	1	2	trEMBL	Mascot
tr|A3SA77|Putative uncharacterized protein.[Sulfitobacter sp. EE-36]	89	7137	9.69	68	1	1	12	trEMBL	Mascot
tr|A7ANI1|Cytochrome c, putative.[Babesia bovis]	B81,B72,B73	12509	9.42	84 (B81	1	1	10	trEMBL	Mascot
tr|Q6B8C1|Glutathione S-transferase 2.[Ixodes pacificus]	B70	25693	7.86	72	1	1	4	trEMBL	Mascot
tr|Q8T7E4Glutathione S-transferase; EC=2.5.1.18;[Boophilus microplus]	B66	25659	8.61	69	1	1	4	trEMBL	Mascot
tr|O97117|Glutathione S-transferase.[Boophilus microplus]	B66	25688	8.22	62	1	1	4	trEMBL	Mascot
tr|A2QMH9|Malate dehydrogenase (EC 1.1.1.37).[Aspergillus niger]	B26,B36,B52,B58	35713	8.92	104 (B26)	1	1	5	trEMBL	Mascot
tr|Q5D9T1|Malate dehydrogenase (EC 1.1.1.37).[Schistosoma japonicum]	B28,B37,B54,B55	36446	8.83	175 (B37)	1	1	8	trEMBL	Mascot
tr|B0QVJ3|Malate dehydrogenase.[Haemophilus parasuis 29755]	B70	32534	5.18	66	1	1	3	trEMBL	Mascot
tr|Q6WJ29|Malate dehydrogenase (EC 1.1.1.37).[Branchiostoma belcheri tsingtauense]	B47	35289	5.8	166	2	2	9	trEMBL	Mascot
tr|A3E3L5|Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8).[Pfiesteria piscicida]	B71	17946		138	2	2	14	trEMBL	Mascot
tr|A1AWR2|Carboxyl-terminal protease precursor (EC 3.4.21.102).[Ruthia magnifica subsp. Calyptogena magnifica]	B71,B72	49803	9.35	59 (B72)	1	1	2	trEMBL	Mascot
tr|A6N9L6|40S ribosomal protein S28.[Ornithodoros parkeri]	B73	7319	10.4	165	2	2	32	trEMBL	Mascot
sp|P0AA25|Thioredoxin-1 (Trx-1) (Trx).[Escherichia coli]	B5,B8	11799	4.67	88 (B8)	1	1	14	SwissProt	Mascot
tr|Q7JPZ2|D.melanogaster ubiquitin (Fragment).[Drosophila melanogaster]	B9	8538	9.16	145	5	3	35	trEMBL	Mascot

tr|B0Y7M4|Thiazole biosynthesis enzyme.[Aspergillus fumigatus]	B11,B20,B27,B28,B30,B36,B37,B41,B45	35469	5.57	96 (B45)	1	1	4	trEMBL	Mascot
sp|P11147|Heat shock 70 kDa protein cognate 4 [Drosophila melanogaster]	B16	71087	5.36	255	4	4	7	SwissProt	Mascot
tr|B0EJ59|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12).[Dictyocaulus viviparus]	B17,B54	36408	8.16	90 (B17)	1	1	4	trEMBL	Mascot
tr|A8DX87|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (Fragment).[Phascolion strombus]	B17,B37,B45,B54,B56	34366	8.78	84	1	1	5	trEMBL	Mascot
tr|Q6QWP2|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (Fragment).[Lithobius sp. S8H266126]	B37,B46	32084	6.07	138	2	2	13	trEMBL	Mascot
tr|Q5WCG4|Glutathione-dependent formaldehyde dehydrogenase (EC 1.2.1.1).[Bacillus clausii]	B28	41260	4.87	52	1	1	2	trEMBL	Mascot
tr|Q0PPV8|Elongation factor 1-alpha (Fragment).[Pleopsidium chlorophanum]	B45,B50	44223	8.67	115 (B45)	3	2	14	trEMBL	Mascot
tr|Q0PPX3|Translation elongation factor-1 alpha (Fragment).[Physcia alpolia]	B52	28586	8.84	206	4	4	15	trEMBL	Mascot
tr|Q1E463|Formate dehydrogenase.[Coccidioides immitis]	B45,B50,B52	41036	6.29	91 (B52)	1	1	3	trEMBL	Mascot
tr|Q4CWA5|Arginine kinase, putative (EC 2.7.3.3).[Trypanosoma cruzi]	B45,B54	40172	6.29	113 (B45)	3	3	7	trEMBL	Mascot
tr|Q4PN07|Thioredoxin-dependent peroxide reductase.[Ixodes scapularis]	B48	25674	9.1	70	1	1	4	trEMBL	Mascot
sp|P07754|Alcohol dehydrogenase 3 (EC 1.1.1.1)(Alcohol dehydrogenase III) (ADH III).[Emericella nidulans]	B52	37102	6.61	146	2	1	10	SwissProt	Mascot
tr|A7E8J7|Citrate synthase, mitochondrial.[Sclerotinia sclerotiorum]	B52	50073	7.21	122	2	2	5	trEMBL	Mascot
tr|A8X135|CBR-ARD-1 protein.[Caenorhabditis briggsae]	b55	27215	8.89	85	1	1	5	trEMBL	Mascot

Table 2.5: Proteins identified from tris extracted unengorged female *I. holocylus* separated by isoelectric focusing on a 24cm pH 3-10 IPG strip. Following focusing, the strip was cut into 1cm length, boiled in SDS sample buffer and the protein separated by SDS-PAGE on a 4-12% Criterion XT gel.

2.3.6: Western blots of unengorged ticks probed with hyperimmune human sera.

The work presented up to this point has employed a non-hypothesis driven approach to experimental design. The experiments have not been designed to look for specific proteins or protein subtypes but have employed techniques to maximise the number of tick proteins able to be identified. This non-hypothesis driven approach has been used to determine which combination of fractionation, separation and analysis techniques are the most appropriate for this sample. If sample was not limiting, the MCE/2-DGE approach provides the highest resolution of the proteins able to be solubilised from unengorged ticks. The 1-D SDS-PAGE/LC/MS approach has the attraction of a reduction in mass spectrometer analysis time, but the poor quality of bioinformatic resources available makes its use limited. By keeping the proteins intact until isolated to a single spot of good intensity, the mass spectrometry data collected from the tryptic peptides can be guaranteed to come from a single protein isoform. This data can be archived for future analysis when the bioinformatic resources improve.

The non-hypothesis driven approach has provided a limited list of proteins that could be considered as vaccine candidates (section 2.4). However this approach is limited in its capacity to extend this list. A hypothesis driven approach is now required, employing a biological “filter” to the sample to identify other vaccine candidates. One such biological filter is the use of immune sera from host animals. In this work, two immune sera are available, a human hyperallergic sera and a dog hyperimmune sera. Each of these will recognise different proteins due to the way the immune response is occurring in each host.

In the work of Dorey (1991), the probing of Western blotted, 1-D SDS-PAGE separated tick proteins with human allergic sera and an anti-human IgE secondary antibody

revealed a number of protein bands that are likely to be tick allergens (section 1.3.2). However, at the time the tools were not available to identify the proteins bound by the human sera. As the serum was still available, the experiment was repeated by probing both 1-D and 2-D gels of the tick sequential protein extracts. The transfer protocol was changed from a Towbin wet blot system to a Khyse-Anderson semi-dry system (1984) as the semi-dry system is found to be quicker and as reliable in our lab. Otherwise, the use of nitrocellulose membrane and 0.1% Tween-20 in TBS as a blocking reagent was unchanged from Dorey's work. Two nitrocellulose membranes were used to "catch" protein passing through the first membrane. The second membrane was stained with Direct Blue 71 to use as a landmark to determine spot position. Figure 2.13 shows Western blots of male and female sequential extracts separated on a 1-D gel probed with human allergic sera and a control blot probed only with anti-human IgE conjugated to alkaline phosphatise (AP). A smear with a number of protein bands in both male and female Tris extracts is detected, but no cross-reaction with the secondary is observed. This is not the case for the UTC7Tris extract where five distinct bands are present in both blots. A similar result has been observed when the commercially available dog anti-tick sera is used to probe Western blots except in that case the secondary antibody is an anti-dog IgG-AP conjugate (data not shown). The non-specific bands are of similar MW and intensity to those seen in Dorey's work where the tick protein extract was unfractionated. Thus, it may be that the result seen in Dorey's work was due to non-specific binding of the secondary antibody rather than the hyperimmune sera. There is no blot showing probing with the secondary antibody alone in Dorey's work and no reference is made to one being performed.

Figure 2.14A shows the Western blotted 2-D gel of the male Tris extract probed with human allergic sera and the anti-human IgE-AP conjugate. The male extract was used in this experiment because all of the female extracts had been used for other experiments. The blot presented in figure 2.13 indicates that, although the protein load is slightly lower, the pattern of bands is similar when both male and female Tris extracts are probed with the human allergic sera. On the blot of the 2-D gel (figure

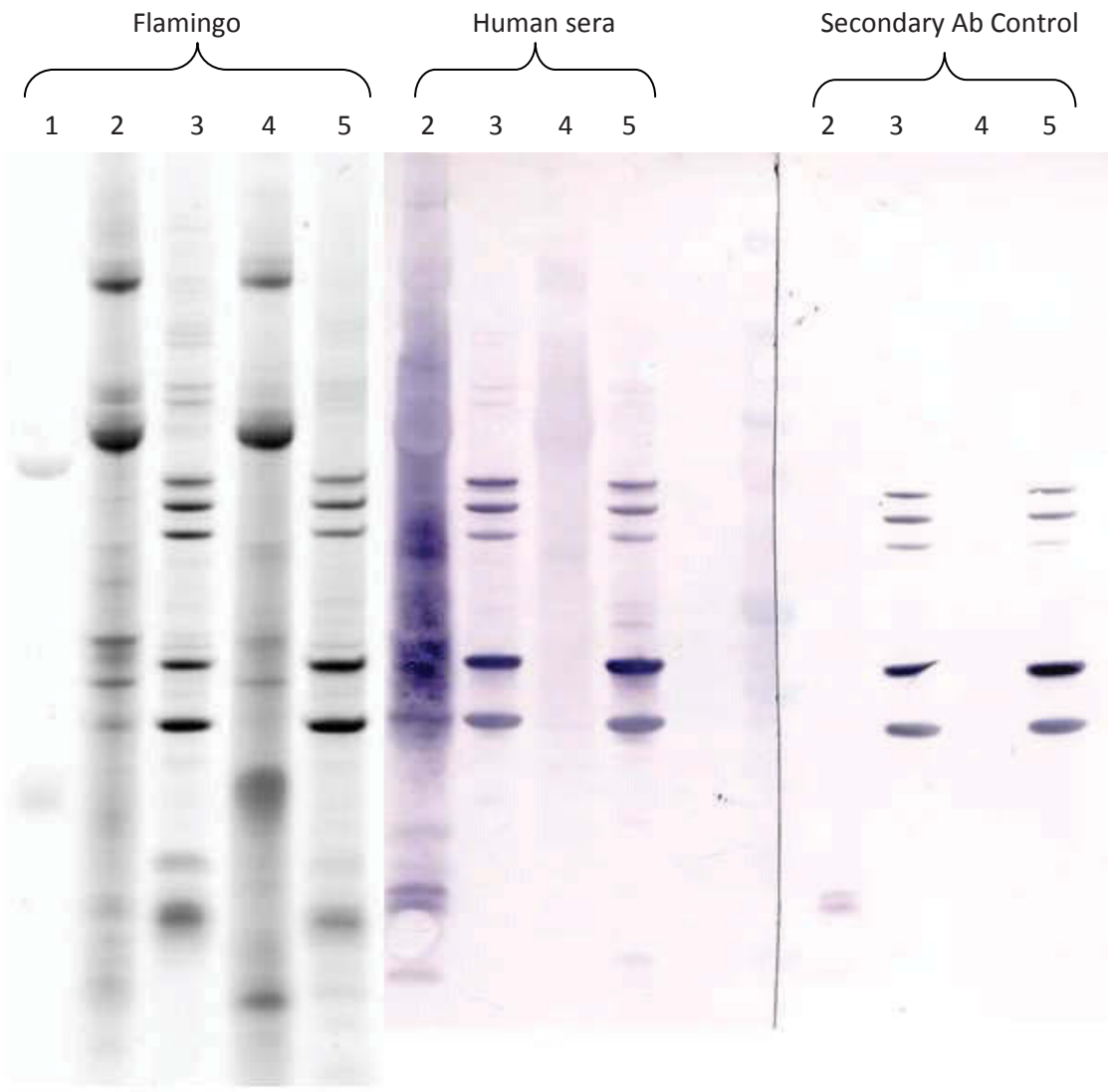

Figure 2.13: Western blot of sequentially extracted male and female *I.holocyclus* proteins separated by 1-D SDS-PAGE. The proteins were blotted to nitrocellulose. Blot A (left) was probed with 10% human allergic sera in TBS/0.05% Tween-20 overnight followed by probing with a anti-human IgE-AP conjugate for two hours. Blot B (right) was probed with anti-human IgE-AP alone for two hours. Both blots were incubated with BCIP/NBT (SigmaFAST) until bands were visible.

1) MW markers
2) Female Tris extract
3) Female UTC7 extract
4) Male Tris extract
5) Male UTC7 extract

2.14A) a number of spots are present, but the high background makes discerning them difficult. Some unstained or negative stained parts of the membrane line up with some intense spots on the Flamingo gel and the support membrane stained with Direct Blue 71 (figure 2.14C). The main set of these spots is labelled “actin spots” as they are in a similar position to actin identified in other gels (figure 2.6). The support membrane was included during blotting, as it is often the case with nitrocellulose membrane that some abundant proteins will pass through the membrane during transfer. These proteins are “caught” by the support membrane, which can then be stained for total protein and used as a reference image for spot cutting from a Flamingo stained gel run in parallel. A control blot probed with secondary antibody alone showed no cross reactivity (blot not shown).

By overlaying the Western blot probed with human sera (figure 2.14A) and the Flamingo stained gel (figure 2.14B) in Adobe Photo Shop, spots detected in the blot by the human sera were able to be located in the gel. Fourteen spots were located, excised from the gel and analysed. The homologous proteins identified are listed in table 2.6. The detection of tropomyosin by the human hyperallergic serum further confirms its status as a pan-allergen as described in section 2.4.3. Of the other proteins identified, paramyosin, protein disulphide isomerase and serpins have all been described as possible vaccine candidates in other blood feeding organisms (see section 2.4). At least one spot visualised on the Western blot cannot be seen on the Flamingo stained gel in a similar region. This is not surprising considering the amplifying nature of alkaline phosphatase conjugated antibodies. As the enzyme consumes the substrate and forms the purple precipitate, the detection limit is increased 10-100 fold above the best fluorescent stains. The protein would be present in the gel, but it is well below the limit of stain detection and analysis by mass spectrometry.

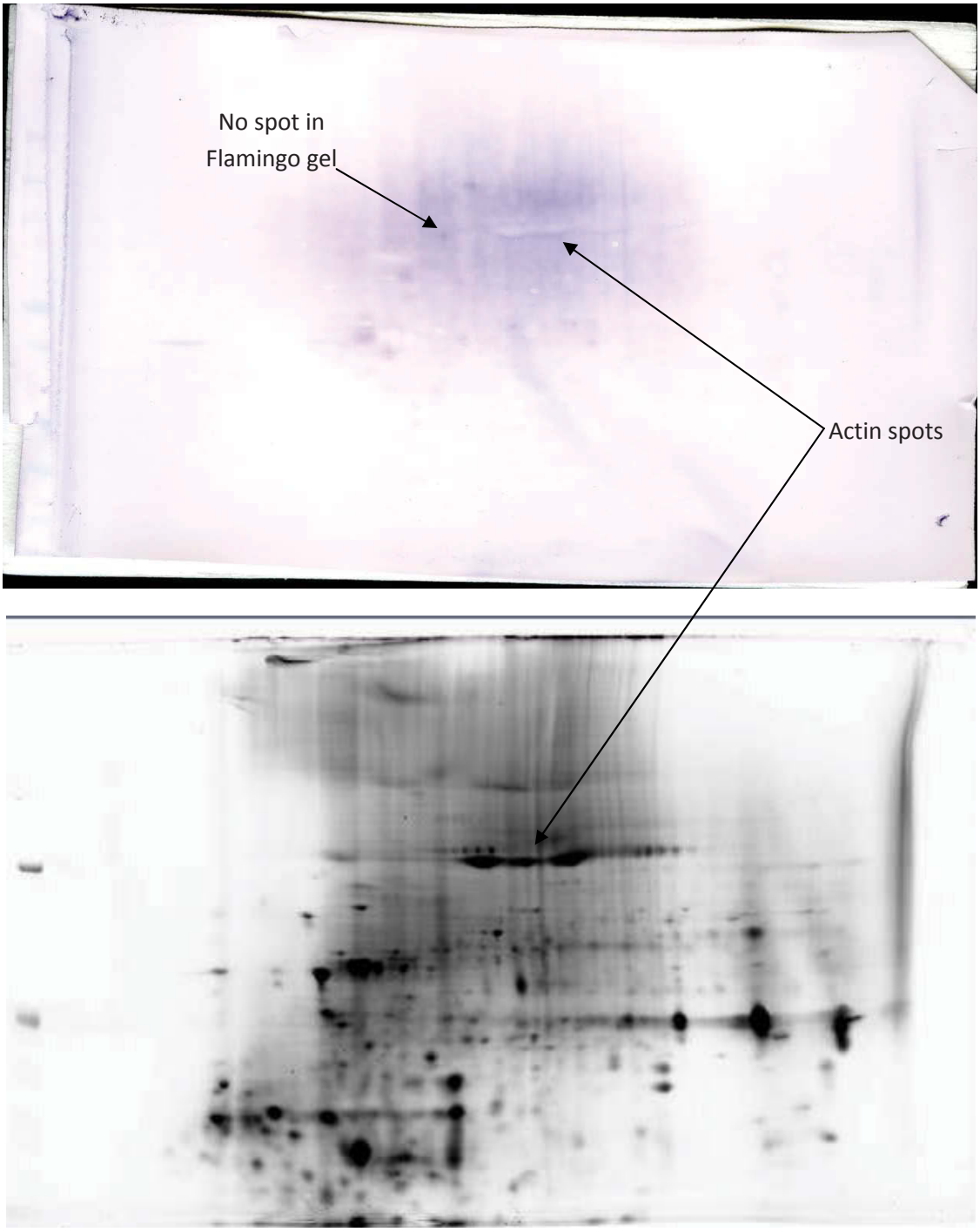

Figure 2.14: Western blot of Tris extracted male I.holocyclus proteins separated by 2-D SDS-PAGE (A: upper gel). The proteins were blotted to nitrocellulose and probed with 10% human allergic sera in TBS/0.05% Tween-20 overnight followed by probing with a anti-human IgE-AP conjugate for two hours. The blot was then incubated with BCIP/NBT (SigmaFAST) until bands were visible. The lower gel (B) is the same sample separated at the same time and in the same way, but the gel was stained with Flamingo showing the total protein in the sample.

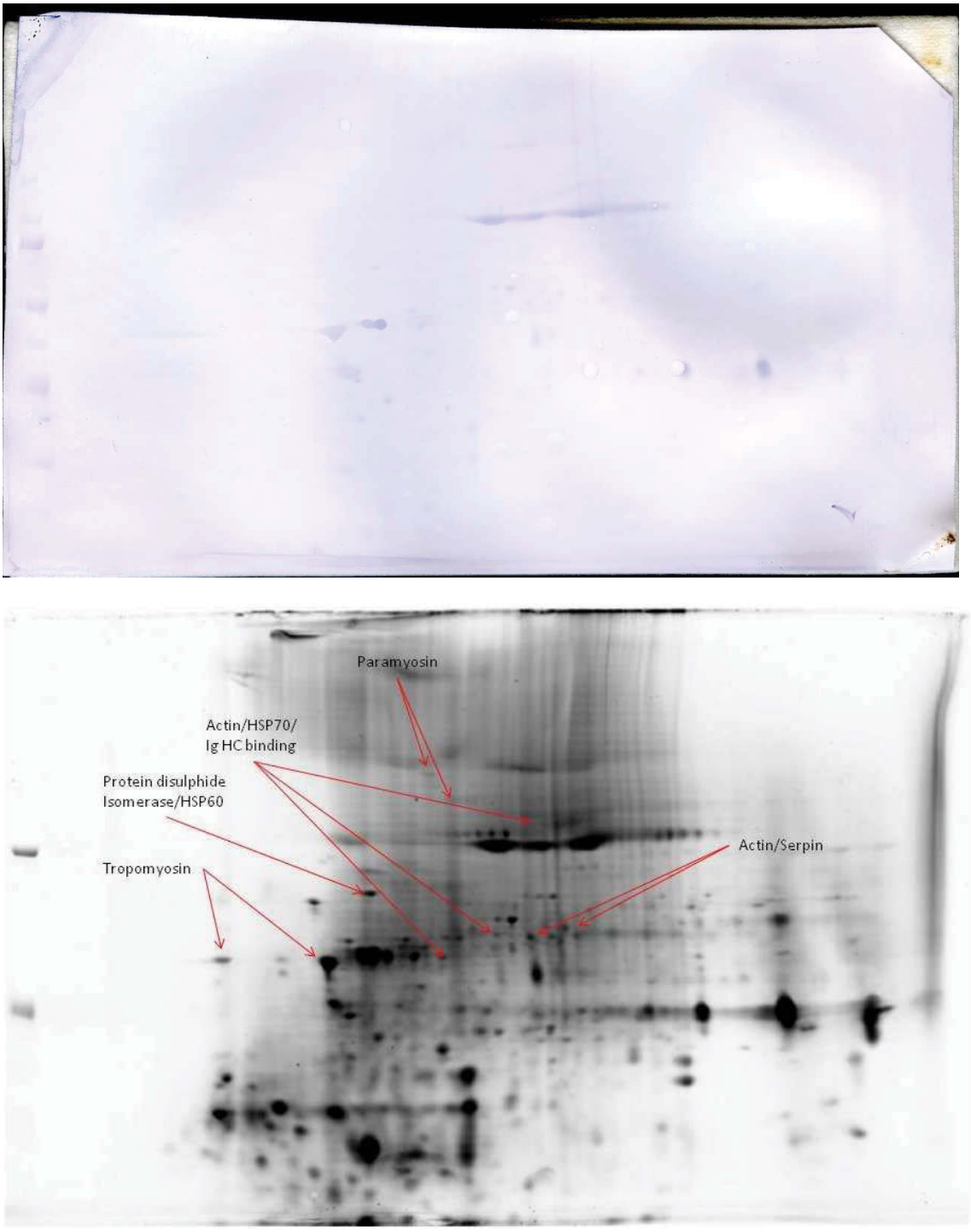

Figure 2.14: Nitrocellulose support membrane used during Western blotting of Tris extracted male *I.holocyclus* proteins. These proteins have been separated by 2-D SDS-PAGE stained for total protein with Direct Blue 71 (C – upper gel). This blot "catches" protein passing through the first membrane (A: previous page). The lower gel (D) is the Flamingo stained gel shown on the previous page (B) now annotated with proteins identified in the Western blot (A).

Spot #	Accession number/Description/Species	pI	MW	Score	Matches	Unique peptides	Coverage (%)	Database	Engine
1	sp|O97162|Tropomyosin [Boophilus microplus]	4.7	32982	605	14	11	26	SwissProt	Mascot
2	sp|O97162|Tropomyosin [Boophilus microplus]	4.7	32982	932	81	17	34	SwissProt	Mascot
3	tr|Q4PKE5| Actin 5 [Aedes aegypti]	5.3	41795	705	18	13	51	trEMBL	Mascot
	tr|A0PA16| Heat shock protein 70kDa [Coturnix coturnix japonica]	5.37	70827	151	2	2	4	trEMBL	Mascot
	tr|Q2489[1]| Immunoglobulin heavy chain binding protein; Fragment [Eimeria maxima]	5.36	38130	151	2	2	7	trEMBL	Mascot
4	tr|Q5CAR2|Actin [Ixodes ricinus]	5.3	41797	275	5	4	21	trEMBL	Mascot
	tr|A0PA16| Heat shock protein 70kDa [Coturnix coturnix japonica]	5.37	70827	129	2	2	4	trEMBL	Mascot
	tr|Q2489[1]| Immunoglobulin heavy chain binding protein; Fragment [Eimeria maxima]	5.36	38130	129	2	2	7	trEMBL	Mascot
5	tr|Q5CAR2|Actin [Ixodes ricinus]	5.3	41797	163	3	3	17	trEMBL	Mascot
	tr|Q06B75| Serpin-1 Precursor [Ixodes ricinus]	6.2	43906	71	1	1	2	trEMBL	Mascot
7	sp|Q86RN8| Paramyosin [Boophilus microplus]	5.53	101930	679	10	9	13	SwissProt	Mascot
9	sp|Q86RN8| Paramyosin [Boophilus microplus]	5.53	101930	1284	21	18	24	SwissProt	Mascot
12	tr|Q7Q270| AGAP004002-PA [Anopheles gambiae] (60kDa Chaperonin)	5.55	60740	280	3	3	6	trEMBL	Mascot

Table 2.6 – Homologous proteins identified from spots matched to spots detected in Western blot probed with human sera (figure 2.12A). Spots were excised from the Flamingo stained gel shown in figure 2.12B.

2.4: Identification of potential vaccine candidates.

The list of proteins found in *I.holocyclus* is a long one, but by no means exhaustive. However, only a subset of these proteins is of interest as they have been reported as vaccine candidates in other organisms. These proteins thus have the potential to be used as vaccine candidates or drug targets, not only for the control of *I. holocyclus*, but of other ticks. Thus the next section will only discuss these selected proteins.

2.4.1: Glutathione-S-Transferase.

Glutathione-S-Transferase (GST) is found in all eukaryotes and catalyses the reaction of reduced glutathione with a variety of compounds with an electrophilic site, such as xenobiotics including insecticides. Elevated expression of GSTs have been correlated with insect resistance to insecticides (Ketterman *et al.*, 2001, Wei *et al.*, 2001), particularly organophosphate resistance (Huang *et al.*, 1998, Ranson *et al.*, 1997, Wei *et al.*, 2001). Parasite GSTs may also be involved in digestive processes (Douch and Buchanan, 1978), prostaglandin synthesis (Meyer and Thomas, 1995) and protecting cellular molecules from oxidative attack (Feng *et al.*, 1999, Sharp *et al.*, 1991). A GST from *R. microplus* has been purified, its gene sequence determined and its presence localised to the salivary glands and gut in adult females (Freitas *et al.*, 2007, He *et al.*, 1999, Rosa de Lima *et al.*, 2002). The gene shows high homology to mammalian mu class GSTs but is very divergent from most insect GSTs. With this level of homology to potential mammalian host GST, the effectiveness of tick GST as a vaccine is unclear, although they have been proposed as vaccine targets in helminths (Balloul *et al.*, 1987, Yang *et al.*, 1997). Its inhibition is likely to be crucial is reversing acaricide resistance. The *I. holocyclus* ticks used in these experiments were collected mainly from native animals and it could be speculated that GST levels may be low due to the low probability of being exposed to acaricides. Another role for GST in *R. microplus* that has been proposed is its involvement in oxidative stress protection (Freitas *et al.*, 2007). GST activity increases significantly

during embryonic development and senescence of the larvae. In embryo development, the increase directly correlated with increased oxygen consumption and metabolism. This leads to higher concentration of reactive oxygen species, which are dealt with by GST. In larvae the reverse was found with oxygen consumption decreasing which may be related to metabolic depression and starvation to maximise their survival time, as found in other diverse organisms (Hochachka and Lutz, 2001). GST levels increase to prevent oxidative stress (Davey and Cooksey, 1989, Labruna *et al.*, 2003). Thus the disruption of tick GSTs could be an effective means of their control.

2.4.2: Calmodulin.

Calmodulin, the most intensively studied member of the E-F-hand family of calcium sensors, is a calcium binding protein and mediates a number of processes in many different cell types and organelles (Chin and Means, 2000). The protein binds to other proteins that cannot bind calcium and thus use calmodulin as a calcium sensor and signal transducer. Calmodulin is expressed in all eukaryotic cells, constituting 0.1% of total cellular protein, and is highly conserved.

2.4.3: Tropomyosin and tropinin complex.

As expected, many of the components of the actin-myosin muscle contraction system were found including tropomyosin, tropinin, various actin and myosin isoforms. In fact actin isoforms were the most abundant proteins in both the soluble and insoluble fraction. Tropomyosin is an actin-binding protein important in muscle contractions, which, along with the tropinin complex, regulates muscle contractions by regulating myosin binding to actin. It is a pan-allergen and highly conserved across species (Jeong *et al.*, 2006). The tropinin complex consists of three proteins, tropinin T or tropomyosin-binding tropinin, tropinin I, or inhibitory tropinin and tropinin C that binds calcium and is a calmodulin-like protein. The work of You *et al* (2001) in *H. longicornis* has described

tropinin T as an antigen expressed during blood feeding as it was discovered by probing western blots with hyperimmune sera of rabbits fed on three consecutive times by 200 adult ticks.

2.4.4: Calreticulin.

Calreticulin is a multifunctional, highly conserved lectin that is inactivated by binding calcium. It has been characterised as a molecular chaperone, an extracellular lectin, an intracellular mediator of integrin function, an inhibitor of steroid hormone-regulated gene expression and a C1q-binding protein (Coppolino and Dedhar, 1998). C1q-binding implies that calreticulin interferes with C1q-mediated inflammatory processes and the gene sequence isolated from *R. microplus* contains six human-like C1q binding sites (Ferreira *et al*., 2002b). In ticks, calreticulin has been found to be secreted into the host and has been suggested to modulate the host immune system (Jaworski *et al.*, 1995). As whole ticks were used in the sample preparation of *I. holocyclus*, it is impossible to determine whether or not the calreticulin identified here would be secreted. Immunisation with recombinant *R. microplus* calreticulin does not produce a humoral response in infested cattle, but the recombinant protein is recognised by sera from *R. sanguineus* infested dogs (Ferreira *et al*., 2002b). Calreticulin has been suggested as a vaccine candidate in *Schistosoma* and hookworms (Pritchard *et al.*, 1999) Using human serum IgE to probe Western blots and cDNA libraries of the hookworm, *Necator americanus*, calreticulin was identified strongly by the sera in both methods. Pritchard *et al* (1997) have hypothesised that due to calreticulin-like molecules associating with components of the complement system, parasite calreticulin is forming complexes with these components. These pseudo-immune complexes would lead to aberrant antigen presentation and push T cell development towards Th2. As stated in the introduction, shifting the immune response to Th2 favours blood feeding in ticks (Schoeler *et al.*, 1999). Immunising with calreticulin can cause disruption of feeding in multi-host ticks such as *A.*

americanium (Jaworski *et al.*, 1995) and *I. scapularis* (Sanders *et al.*, 1999) but not in one-host ticks such as *R. microplus* (Ferreira et al., 2002b). Thus it may be a useful vaccine candidate in *I. holocyclus* as it is a three-host tick.

2.4.5: Thioredoxin-dependent peroxide reductase.

Thioredoxin-dependent peroxide reductases or peroxiredoxins are part of the anti-oxidant system. In mammalian hosts, selenium-containing glutathione peroxidases provide most of the self-protection against hydroperoxides, but these are not found in many pathogenic organisms (Jaeger and Flohe, 2006). Thus components of parasite anti-oxidant systems are attractive vaccine candidates and drug targets because of this evolutionary diversification although it may mean that individual vaccines or drugs are required for each pathogen or group of pathogens. Peroxiredoxins obtain their reducing equivalents from the thioredoxin reductase/thioredoxin redox system and transfer the reducing equivalent to hydrogen peroxide (Krnajski *et al.*, 2001) and they do not depend on redox cofactors such as metals or prosthetic groups (Kim *et al.*, 1988).

2.4.6: Arginine kinase.

Arginine kinase was found by both 2-DGE and 1-D SDS-PAGE/LC/MS. This enzyme is a type of phosphagen kinase that catalyzes the transfer of a phosphoryl group from a guanidino phosphagen, to ADP, producing a molecule of ATP. They are found throughout invertebrates and are usually present in tissues with high energy demands requiring continuous delivery of ATP or tissues that require short bursts of energy demand such as muscle fibres (Kotlyar *et al.*, 2000, Lang *et al.*, 1980). Other cellular and tissue processes that experience high energy demand include monovalent ion transport, acid-base balance, nitrogen excretion and gas exchange (Kotlyar *et al.*, 2000). A number of high energy demand processes occur in ticks. During blood feeding, Fielden *et al* (1999) observed a switch from discontinuous to continuous CO_2 release in *D. variabilis* was correlated to an

increase in metabolic expenditure associated with blood meal digestion. Arginine kinase was found amongst the twenty most abundant proteins on 2-DGE by Untalan *et al* (2005) in *R. microplus*. The spot from which arginine kinase was identified in the Tris-extract (figure 3.3A) in these experiments is of reasonable intensity, but not amongst the twenty most-abundant. In the 1-D SDS-PAGE/1-DLC/MS experiment, this protein was found in both fractions in the top third of both lists (Table 3.4C and D), supporting Untalan's findings on its abundance. In vertebrates, only one out of the seven different guanidino phosphagen kinases is present, creatine kinase. Thus if tick arginine kinase is exposed to immune attack it could be an effective vaccine as no homolog is present in potential hosts. Kiel *et al* (2007) identified arginine kinase as one of 28 immuno-reactive molecules from Western-blotted 2-D gels of the sheep gastrointestinal nematode, *Trichostrongylus colubriformis,* probed with plasma from resistant sheep, suggesting it is an exposed antigen.

2.4.7: Enolase.

Enolase, also known as 2-phospho-D-glycerate hydrolyase, is a metalloenzyme catalysing the conversion of 2-phosphoglycerate to phosphoenolpyruvate in the penultimate step of glycolysis. It is present in all tissues and organisms capable of glycolysis or fermentation. Enolase is also found on the surface of prokaryotic and eukaryotic cells, but it possesses no signal sequence or membrane anchoring motif (Pancholi, 2001). They are known to be secreted before reassociating with the membrane and exhibit ligand binding, non-enzymatic functions that play roles in colonisation and invasion of some pathogens and parasites. In *S. pneumonia*, a secreted α-enolase, which has been established as a receptor and activator of plasminogen, was found to increase plasminogen binding activity (Bergmann *et al.*, 2001). Enolase has been found to have the same plasminogen and plasmin-binding activity in *Candida albicans*, indicating that the interaction may contribute to tissue invasion (Jong *et al.*, 2003). Plasminogen activation is

responsible for the degradation of intravascular clots (Plow *et al.*, 1995, Redlitz *et al.*, 1995), an activity that would be very beneficial to blood feeding ticks. Kiel *et al* (2007) identified enolase as one of 28 immuno-reactive molecules from Western-blotted 2-D gels of the sheep gastrointestinal nematode, *Trichostrongylus colubriformis,* probed with plasma from resistant sheep. The determination of whether tick enolase is surface associated and which cell type it is associated with would require the use of immunohistochemical techniques

2.4.8: Aspartic protease.

Spot A8 from the Tris-extract has homology to an aspartic protease-like protein of the peptidase A1 family. In general these enzymes have two highly conserved aspartates in their active site and have an acidic pH optimum. In eukaryotes, the group includes pepsins, cathepsins and renins. An aspartic protease isolated from *R. microplus* eggs, *Boophilus* Yolk Cathepsin (BYC), has been tested as a vaccine candidate in cattle inducing a significant immune response (Nascimento-Silva *et al.*, 2008). BYC is different to other aspartic endopeptidases in that it has only one catalytic aspartate residue, which results in very slow digestion of its natural substrate, vitellin, a hemeprotein found in tick egg yolk. Another protein with aspartic protease activity found in *R.microplus* is THAP or Tick Heme-binding Aspartic Proteinase (Sorgine et al., 2000). THAP possesses a heme-binding site that is believed to work as a docking site to recognise heme on protein surfaces, thus increasing the specificity of THAP towards hemeproteins and possibly regulated vitellin degradation. Other work has also suggested aspartic proteases play a key role in haemoglobin proteolysis (Boldbaatar *et al.*, 2006, Sojka *et al.*, 2008) and disrupting these enzymes from degrading vitellin may be an effective method of tick control (Pohl *et al*, 2008).

2.4.9: Protein disulphide isomerise.

Protein disulphide isomerise (PDI) is a multifunctional protein whose role in the endoplasmic reticulum (ER) involves folding, assembly and post translational modification of proteins by catalysing the oxidation and isomerisation of disulphides (Wilkinson and Gilbert, 2004). PDIs are one of the most abundant ER proteins with millimolar concentration (Lyles and Gilbert, 1991) and have been found to be expressed in high concentration in lysates of *H. longicornis* (Liao *et al.*, 2007). In the same study, two PDI genes were found to be upregulated in *Babesia* infected larvae suggesting a role in pathogen transmission. Further study in *H. longicornis* with RNA interference (RNAi) of the PDI genes (Liao *et al.*, 2008) showed severe damage to the midgut and cuticle along with significant reduction of disulphide bonded vitellogenin expression. The results indicate involvement of PDIs in blood feeding, viability and oocyte development. PDI has also been found to be abundantly expressed in the salivary glands of *A. variegatum* and is a potential vaccine target (Knizetova *et al.*, 2006). Kiel *et al* (2007) identified protein disulpide isomerase as one of 28 immuno-reactive molecules from Western-blotted 2-D gels of the sheep gastrointestinal nematode, *Trichostrongylus colubriformis,* probed with plasma from resistant sheep.

2.4.10: Paramyosin.

Paramyosin is a known allergen in mites (Jeong *et al.*, 2006) and is highly immunogenic in extracts of adult schistosomes (Gobert and McManus, 2005). In addition, vaccination/challenge studies with native paramyosin produced a protective immune response against schistosomes. Paramyosin has been shown to be immunogenic in a number of helminths and is being explored as a vaccine, however work to date has shown only modest reductions in worm and/or egg burdens. Indirect immunofluorescence experiments using monoclonal antibodies produced against native schistosome paramyosin has shown surface expression of the protein in the muscle layer of the worm making the protein readily available to the host immune system. A number of non-muscle

functions have been suggested for paramyosin (Gobert and McManus, 2005) and the situation could be similar in ticks. Paramyosin from *S. mansoni* has been shown to bind the collagen-like reagent of complement protein C1q, inhibiting complement C1 function thus possibly inhibiting the complement cascade (Laclette *et al.*, 1992). Another complement inhibitor protein from *S. mansoni* has been characterized as a surface-exposed form of paramyosin (Deng *et al.*, 2003) while another study (Loukas *et al.*, 2001) has shown paramyosin acting as an Fc receptor on the surface of *S. mansoni*. Akpek *et al*. (2002) found another immunomodulatory function for paramyosin in the binding of calgranulin, a protein released by neutrophils involved in host defense. All of this evidence suggests the surface expressed form of paramyosin inhibits the complement cascade, modulating the immune system in favour of this blood feeding parasitic worm (Gobert and McManus, 2005). Previous work in *R. microplus* has shown that recombinant paramyosin can bind both collagen and IgG (Ferreira *et al.*, 2002a) and the protein has been localized to the fat body and salivary glands but not in saliva itself. However there is no evidence that the protein is exposed to the host immune system (Ferreira *et al*, 2002a) which needs to be ascertained before it can be explored as a vaccine candidate.

2.5: Conclusions.

The results presented in this chapter show that simply fractionating and cataloguing the proteins present in an organism's proteome is unable to provide any real insight into what proteins would be useful vaccine candidates. The data presented shows how relatively simple it is to prepare a sample and resolve approximately 900 protein spots using 2-D gels when the MCE and narrow range strips are combined. The use of 1-D SDS-PAGE/1-DLC/MS/MS is also useful and provided complimentary data to the 2-D gels, mainly due to the more comprehensive treatment of the separated proteins. By dissecting the gel lane completely, proteins are included and analysed that would be too faint to be seen as a spot on 2-DGE. This is supported by the observing that the extra proteins found

in the 1-D SDS-PAGE/1-DLC/MS/MS experiments were identified by one high scoring peptide. These methods can now be combined with other novel techniques and assays to reveal proteins that are possible vaccine candidates and gain insight into the tick-host relationship. However, it is clear that the key to identifying the maximum number of proteins is not only to start with much larger amounts of starting material, but to fractionate it in such a way that high abundance proteins are removed from the analysis. Cytoskeletal proteins have been the main high abundance protein needing to be dealt with during this work and this led to the variation and combination of a number of techniques. The results of these experiments varied and a single solution was not found. The laborious dissection of the tick and the isolation of organs may reduce the abundance of cytoskeletal proteins by removing the contribution of those in the cuticle, but this is not likely to significantly affect their presence as a high abundance proteins.

The MCE, in concert with narrow range IPG strips, doubled the number of protein spots visible. However, to properly exploit the loading capacity of the MCE, 500 unengorged female ticks would be required for liquid nitrogen grinding. This would greatly increase spot intensity and consequently the quality of MS/MS data obtained. The disadvantage to starting with 500 ticks is that cytoskeletal proteins, especially actin, will also be increased and will continue to interfere with analysis. This was demonstrated in the 1-D SDS-PAGE/1-DLC/MS methodology where actin was present in high abundance in the top half of the gel and easily dominated the list of identified proteins. By including a pH 6.5 membrane into the MCE it may be possible to isolate actin in the pH 5-6.5 chamber as much of the actin found have pIs in this range. However the analysis of any low abundance protein focused into this chamber would be very difficult. The use of 24cm IPG strips as the isoelectric fractionators in place of the MCE was also evaluated. As mentioned, the focusing time needs to be extended to ensure proteins are more focused to their pI. This is indicated by the presence of actin, tropomyosin, paramyosin and myosin

heavy chain across a wide range of pI. However, the interference of these high abundance proteins with the analysis of other proteins has been reduced by their containment in discrete bands in the second dimension SDS-PAGE. Continuous elution preparative electrophoresis proved to be of limited use compared to the other methodologies employed. This is mainly due to the great number of additional handling and concentration steps required. This method would be of far greater use when combined with the MCE to fractionate milligram quantities of sample that would cause problems associated with overloading if used with the other methodologies.

A great number of protein spots were not able to be identified by MS/MS ion searching of the available databases or by *de novo* sequencing. This result reflects the results found by other researchers using both proteomic and transcriptomic approaches to studying ticks. Francischetti *et al* (2008) while commenting on tick salivary transcriptome studies made the point that, common to all the studies, is that the vast majority of the identified proteins have no confirmed or known function. The solution to this is in two parts. Firstly a complete tick genome sequence is necessary. This is soon to be provided in the form of the *I. scapularis* genome sequencing project (Hill and Wikel, 2005, Pagel Van Zee *et al.*, 2007). In the case of *I. holocyclus*, the usefulness of using the *I. scapularis* genome to search data will need to be investigated, keeping the work of Mans and Neitz (2004) in mind. As stated in section 2.1.4, the authors point out that tick species may have diverged at a very early point in their evolution and blood feeding may have evolved independently. If this is indeed the case, proteins with a certain function in one family could have quite a different amino acid sequence to the protein with a similar function in another family. As *I. holocyclus* is only known to be found on the east coast of Australia it is possible it diverged and was isolated at this very early evolutionary point. Thus its genome may be different enough to *I. scapularis* to greatly impede searches with MS/MS data as the

current software cannot deal with sequence variation that affects peptide parent and fragment masses.

The second part to the solution is the annotation of the available databases. In this work, a number of matches were to EST sequences present in the available tick gene indices (compbio.dfci.harvard.edu/tgi/). Although these ESTs have been assembled into tentative consensus sequences, BLAST searching of larger databases, such as NCBI non-redundant, returns matches with hypothetical proteins of no known function. This issue is a recurrent one when working on organisms with only limited study outside the highly characterised model organisms.

Two dimensional gel electrophoresis is still the highest resolving proteomic technique available, however it does have some practical issues that need consideration. The ability to resolve 400 spots from a sample on one gel using a pH 3-10 IPG strip needs to be balanced with the difficulty of excising all of these protein spots, trypsin digesting and analysing the peptides by mass spectrometry. It is impractical to devote 400 hours of instrument time to analysing all the spots on one gel, not to mention the increasing chance of contaminating the samples when manually performing 400 trypsin digests. The solution to this is the use of differential display, as in the work of Rachinsky *et al* (2007a & b), or using methodologies that target subsets of proteins. In this work, a differential display experiment could not be designed as no comparative sample was available. The design of a meaningful differential display experiment must be very carefully undertaken as the samples cannot be radically different as the proteomes cannot be compared. In addition, mixed proteomes present a challenge, especially if one of the proteomes is especially dominant in its abundance, as in engorged ticks. The use of a targeted

methodology is an attractive alternative, and the use of fluorescent zymograms to determine enzyme activity is one example which will be presented in the next chapter.

The use of sera from hyperallergic hosts is an effective tool to be used in a targeted approach. This has been previously used with Western blots of 2-DGE separated protein samples as in the work of Rachinsky *et al* (2007a & b). A hyperimmune dog sera raised against feeding *I.holocyclus* was available for use in this work and was trialled with Western blotted 2-DGE separated sequentially extracted proteins from unengorged females. However cross reactivity of the anti-dog IgG secondary antibody with the blotted tick proteins prevented this approach being useful in this work (data not shown). The solving of this non specific binding problem or the sourcing of an alternative secondary antibody may not enable Western blotted tick proteins to be a useful tool as many researchers would expect. The probing of Western blotted proteins with specific primary antibodies and secondary antibodies conjugated to alkaline phosphatase or horseradish peroxidise and their localisation by visual or chemiluminescent staining is 1000-10000 times more sensitive than current visual and fluorescent protein stains (Ben Herbert, personal communication). Thus the detection of a protein spot on a blot does not guarantee that it will able to be identified and thus fractionation and enrichment techniques need still to be optimised to obtain sufficient amounts of low abundance proteins to be identified by mass spectrometry. This is the case for the Western blots presented in figure 2.13 where some spots are present on the blot, but not on the Flamingo stained gel.

The cross-reaction of the anti-dog IgG secondary antibody was a major limiting factor to fulfilling one of the goals of this work, the discovery of proteins in unengorged ticks that cause an immune response. Cross reaction was also observed with the anti-human IgE

secondary antibody, but this was limited to proteins in the UTC7Tris extract allowing probing of Western blotted Tris extracted proteins with hyperimmune human sera to be carried out. However, a solution is needed to prevent the cross reactivity of both secondary antibodies. One solution is to not use a secondary antibody at all and biotinylate all of the protein present in the hyperimmune human and dog sera. This biotinylated can be used to probe Western blots which are subsequently probed with avidin conjugated to alkaline phosphatase (AP) or horse radish peroxidise (HRP). This would produce the same level of sensitivity as AP or HRP labelled secondary antibodies as the signal amplification occurs through the conjugated enzymes. In addition, the proteins of the anti-sera are labelled with multiple biotin molecules, producing an additional level of signal amplification.

The use of biotinylated anti-sera would also circumvent another issue specific to the dog anti-tick sera. This sera is isolated from dogs that are continuously fed on by ticks and as a result of the continual feeding, the dogs become hyperimmune and do not suffer from paralysis. However, anecdotal reports show that if the hyperimmune dogs are kept tick-free for a period of over six months they are no longer immune to the neurotoxin and will suffer paralysis if fed on by a single female tick [http://www.peg.apc.org/~ullavet/tick.html]. This would imply that it is in fact IgM that provides protection against holocyclotoxin in hyperimmune animals continually fed on by *I. holocyclus* and modulation of the immune system by the tick prevents a switch to a memory, IgG based immune response. This hampers the use of the dog anti-tick sera in Western blot analysis due to the difficulty of obtaining appropriate anti-dog IgM and IgE secondary antibodies. Biotinylation of the anti-sera makes the sourcing of the secondary antibodies unnecessary. A final advantage to the biotinylated anti-sera is that the class of antibody binding to a specific protein spot on a blot maybe determined by simply cutting the relevant piece of blot, trypsin digesting the protein and antibody and analysing the

resulting peptides by mass spectrometry. These methods were unable to be tested in this work due to a lack of time and tick sample for Western blots.

An alternative approach to Western blots is to use the sera bound to cyanogen bromide activated Sepharose for affinity chromatography. The bound protein can then be eluted from the beads and further separated by 1-D or 2-D SDS-PAGE, excised from the gel if detected by Coomassie staining and identified by mass spectrometry. This methodology was trialled in this work and problems were confronted. The majority of the proteins in the hyperimmune sera are serum proteins rather than antibodies and numerous attempts were made to isolate high molecular weight immunoglobulin from the serum proteins. This was to ensure that the affinity column contained the maximum amount of immunoglobulin to bind the maximum amount of tick proteins causing a host immune response rather than non-specific binding of tick proteins to host serum proteins. Protein A/G isolation of IgG from the sera was also considered and rejected due to anecdotal evidence that it is in fact IgM that provides protection against holocyclotoxin in hyperimmune animals continually fed on by *I. holocyclus* as described in the previous paragraph [http://www.peg.apc.org/~ullavet/tick.html]. The retention of any IgE present was also considered when rejecting the use of Protein A/G. Tris extracted tick proteins was reacted with the immobilised immunoglobulin as the addition of the chaotrope rich UTC7 extract would denature the antibody binding sites and also affect the structure of the agarose beads. After eluting the bound protein with 0.2M glycine-HCl pH 3.0, the eluate was analysed by 1-D SDS-PAGE. No protein was found in the eluate (gel image not shown).

The work presented in this chapter represents the first comprehensive proteomic analysis of *Ixodes holocylus* to be reported. Table 2.7 lists the homologous proteins identified by

the three main methods used in this chapter, 2D-PAGE, 1D-PAGE/1D-LC/MS/MS and 24cm IPG/1D-PAGE. In addition, ten proteins were identified that have been shown in other tick species to have promise as either potential vaccine candidates or drug targets. Several proteins, such as paramyosin, calreticulin, and tropomyosin, have been suggested as vaccine candidates in other organisms such as *Schistosoma*. Others proteins, such as the peroxiredoxins, have been proposed as drug targets while the protein disulphide isomerise transcript has been shown to be a very effective target for RNA interference. However, the aim of this work is not to ascertain which will be an effective vaccine candidate, but to develop methodologies to discover possible candidates and assist study of the tick-host interaction. The next chapters described the development and adaption of methodologies to help facilitate these studies.

	Method found in		
Accession number/Description/Species	2D-PAGE	1D-PAGE	24cm IPG
tr|Q179E8|Myosin heavy chain, nonmuscle or smooth muscle.[Aedes aegypti]		X	X
tr|B2ZTQ5|Muscle myosin heavy chain.[Loligo bleekeri]			X
sp|P05661|Myosin heavy chain, muscle.[Drosophila melanogaster]	X	X	X
tr|Q9U0S7|Myosin heavy chain (Fragment).[Mytilus galloprovincialis]			X
tr|A0MQ61|Slow myosin heavy chain 1.[Danio rerio]	X		X
tr|Q4PM87|Nonmuscle myosin essential light chain.[Ixodes scapularis]	X	X	X
tr|B0FL77|Myosin light chain 2.[Bombyx mandarina]	X		
sp|Q86RN8|Paramyosin.[Boophilus microplus]	X	X	X
tr|Q5CAR2|Actin.[Ixodes ricinus]	X		X
tr|Q4PKE5|Actin 5.[Aedes aegypti]	X	X	X
sp|P53456|Actin-2.[Diphyllobothrium dendriticum]			X
tr|A9QUS4|Beta-actin.[Rachycentron canadum]	X		X
tr|Q172T4|Alpha-actinin.[Aedes aegypti]		X	X
tr|B0B5G3|Alpha-tubulin (Tubulin alpha-1).[Fasciola hepatica]			X
sp|Q8WQ47|Tubulin alpha chain (Allergen Lep d ?).[Lepidoglyphus destructor]			X
tr|Q16JS3|Tubulin alpha chain.[Aedes aegypti]		X	X
tr|Q8T8B0|Beta-tubulin.[Bombyx mori]			X
tr|Q8T8B2|Beta-tubulin.[Bombyx mori]		X	X
tr|A1C231|Beta-tubulin (Fragment).[Cryptocercus punctulatus]			X
tr|Q0PHP0|Beta-1 tubulin (Fragment).[Aedes aegypti]	X		
tr|Q6S3D4|Beta-2 tubulin.[Laodelphax striatellus]			X
tr|Q9U7Q1|Beta tubulin.[Trichuris trichiura]		X	
sp|O97162|Tropomyosin.[Boophilus microplus]	X	X	X
tr|Q32VZ5|Calmodulin (Fragment).[Eucheilota bakeri]			X
sp|P11121|Calmodulin (CaM).[Pyuridae sp.]			X
tr|Q98UH8|Calmodulin (Fragment).[Clemmys japonica]			X
sp|P02595|Calmodulin (CaM).[Patinopecten sp.]	X	X	
sp|P15159|Troponin C.[Tachypleus tridentatus]	X	X	X
sp|Q09665|Troponin C, isoform 2.[Caenorhabditis elegans]			X
tr|A8E4K0|Troponin T.[Haemaphysalis qinghaiensis]	X	X	
tr|Q4SKJ3|Chromosome undetermined SCAF14565, whole genome shotgun sequence.[Tetraodon nigroviridis] (Histone H2B)		X	X
tr|Q17EF0|Histone h2a.[Aedes aegypti]		X	X
tr|Q4PM63|Histone H2B.[Ixodes scapularis]		X	X
ens transcript:AGAP012871-RA[Anopheles gambiae] (Histone H2B)			X
tr|A7S4X9|Predicted protein.[Nematostella vectensis] (Histone H2A)			X
tr|Q4PMB4|60S acidic ribosomal protein P0.[Ixodes scapularis]		X	
tr|Q0PXZ8|Putative 60S acidic ribosomal protein P1.[Diaphorina citri]			X

Table 2.7: List of homologous proteins identified by the three main analysis methods used in this chapter, 2D-PAGE, 1D-PAGE/1D-LC/MS/MS and 24cm IPG/1D-PAGE.

tr|Q4PMC0|Ribosomal protein LP2.[Ixodes scapularis]		X	
tr|Q4PM16|60S ribosomal protein L23.[Ixodes scapularis]		X	
sp|Q4PM54|60S ribosomal protein L17.[Ixodes scapularis]		X	
sp|Q4PM04|60S ribosomal protein L18.[Ixodes scapularis]		X	
tr|Q4PMD6|Ribosomal protein L30.[Ixodes scapularis]		X	
sp|Q4PMD1|60S ribosomal protein L38.[Ixodes scapularis]		X	
tr|A6N9Z6|60s ribosomal protein L10.[Ornithodoros parkeri]		X	
tr|A6N9L6|40S ribosomal protein S28.[Ornithodoros parkeri]		X	
tr|Q4PM82|Ribosomal protein S25.[Ixodes scapularis]		X	
tr|Q4PM10|Ribosomal protein S14.[Ixodes scapularis]		X	
tr|Q4PM11|40S ribosomal protein S13.[Ixodes scapularis]		X	
tr|Q4PMD3|40S ribosomal protein S10.[Ixodes scapularis]		X	
tr|Q4PM03|40S ribosomal protein S9.[Ixodes scapularis]		X	
sp|Q4PMB3|40S ribosomal protein S4.[Ixodes scapularis]		X	
tr|A6N9Z4|40S ribosomal protein S3.[Ornithodoros parkeri]		X	
tr|Q4PM67|Ribosomal protein S16.[Ixodes scapularis]		X	
tr|A6N9R2|Ribosomal protein S18.[Ornithodoros parkeri]		X	
tr|Q4PM30|Ribosomal protein S19.[Ixodes scapularis]		X	
tr|Q5DJT2|Histone H3 (Fragment).[Limnadopsis birchii]		X	
tr|A5Y4F5|Histone H3a (Fragment).[Ostracoberyx dorygenys]		X	
tr|Q17EE8|Histone H4.[Aedes aegypti]			X
tr|B2SU79|Chain length determinant protein.[Xanthomonas oryzae pv. oryzae PXO99A]			X
ens transcript:ENSSART00000003513[Sorex araneus]			X
tr|A7C535|Putative uncharacterized protein.[Beggiatoa sp. PS]			X
tr|Q2F5T3|ATP synthase subunit alpha (EC 3.6.3.14).[Bombyx mori]	X		X
tr|Q6PTP3|ATP synthase subunit beta (EC 3.6.3.14) (Fragment).[Encope michelini]	X	X	X
tr|Q17FL3|ATP synthase subunit beta (EC 3.6.3.14).[Aedes aegypti]	X		X
sp|Q0A4M8|ATP synthase subunit beta (EC 3.6.3.14) [Alkalilimnicola ehrlichei]			X
ens transcript:ENSGACT00000018465[Gasterosteus aculeatus] (ATP synthase subunit beta)			X
tr|Q1HRQ7|Mitochondrial ATP synthase alpha subunit (ATP synthase alpha subunit mitochondrial).[Aedes aegypti]		X	
tr|Q5MCG8|Mitochondrial ATP synthase beta subunit (Fragment).[Enchytraeus buchholzi]			X
tr|Q5MCG6|Mitochondrial ATP synthase beta subunit (Fragment).[Theromyzon tessulatum]			X
tr|Q9XYC8|Vacuolar ATPase B subunit (ATP synthase subunit beta vacuolar).[Aedes aegypti]			X
tr|B0Y0Q5|ATP synthase subunit beta (EC 3.6.3.14).[Aspergillus fumigatus]			X
tr|Q4PM92|ATP synthase D chain.[Ixodes scapularis]	X		
tr|A2ETL5|Putative uncharacterized protein.[Trichomonas vaginalis G3]			X
tr|Q4PHP7|Putative uncharacterized protein.[Ustilago maydis]			X

Table 2.7: List of homologous proteins identified by the three main analysis methods used in this chapter, 2D-PAGE, 1D-PAGE/1D-LC/MS/MS and 24cm IPG/1D-PAGE.

tr|A5BZ64|Putative uncharacterized protein.[Vitis vinifera]			X
tr|Q11AW8|2-octaprenylphenol hydroxylase (EC 1.14.13.-).[Mesorhizobium sp.]			X
tr|Q16PM9|Chaperonin-60kD, ch60.[Aedes aegypti]		X	X
tr|Q0GFD6|60 kDa chaperonin (Fragment).[uncultured Bartonella sp]			X
tr|O96783|Heat shock protein 60 (HSP60).[Plectus acuminatus]			X
tr|A0PA16|Heat shock protein 70kDa.[Coturnix coturnix japonica]		X	X
tr|A9BK82|Hsp70.[Cryptophyta]			X
sp|O73885|Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8).[Gallus gallus]			X
tr|B0XPW1|Molecular chaperone Hsp70.[Aspergillus fumigatus]			X
tr|A0JMA1|Heat shock protein 90kDa alpha (Cytosolic), class A member 1, gene 2.[Xenopus tropicalis]		X	
tr|A0A106|Enolase (EC 4.2.1.11) (Fragment).[Semiotellus sp. CD038]	X		X
tr|A0SNX4|Enolase (EC 4.2.1.11).[Trimastix pyriformis]			X
tr|Q963H2|Enolase (EC 4.2.1.11) (Fragment).[Taphrorychus bicolor]		X	X
tr|Q7YZX3|Enolase (EC 4.2.1.11).[Onchocerca volvulus]			X
tr|Q8MWP3|Calreticulin.[Boophilus microplus]			X
tr|Q64K84|Calreticulin.[Ixodes jellisoni]	X		X
tr|Q7KLE5|Amphiphysin (CG8604-PA) (LD19810p).[Drosophila melanogaster]			X
tr|B0L419|Laminin receptor (Fragment).[Oryctolagus cuniculus]			X
tr|Q0V9W8|Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide.[Xenopus tropicalis]			X
sp|Q1HPS0|Myosin regulatory light chain 2 (MLC-2).[Bombyx mori]			X
tr|A3WG27|Putative uncharacterized protein.[Erythrobacter sp. NAP1]			X
tr|A5LHV9|Protein disulfide isomerase-2.[Haemaphysalis longicornis]			X
tr|A3EXM7|Putative ubiquilin 1,2 (Fragment).[Maconellicoccus hirsutus]			X
tr|B0D1E3|Predicted protein.[Laccaria bicolor]			X
tr|B2X122|Serpin-2.[Spodoptera exigua]		X	X
tr|Q06B75|Serpin-1 precursor.[Ixodes ricinus]	X		
tr|Q2JXM5|Peptidase inhibitor, I4 family.[Synechococcus sp.] (Serpin-2)			X
ens transcript:ENSAPMT00000018309 [Apis mellifera]			X
tr|A6NVK8|DNA polymerase I (EC 2.7.7.7).[Bacteroides capillosus ATCC 29799]			X
tr|Q291K6|GA10074-PA (Fragment).[Drosophila pseudoobscura]			X
tr|Q17A09|Mitochondrial processing peptidase beta subunit.[Aedes aegypti]			X
tr|Q9XYC8|Vacuolar ATPase B subunit (ATP synthase subunit beta vacuolar).[Aedes aegypti]			X
tr|Q2T5I8|Gp51.[Burkholderia thailandensis]			X
tr|Q1HPK0|Vesicle amine transport protein.[Bombyx mori]			X
tr|A3SA77|Putative uncharacterized protein.[Sulfitobacter sp. EE-36]			X
tr|A7ANI1|Cytochrome c, putative.[Babesia bovis]		X	X
tr|Q6B8C1|Glutathione S-transferase 2.[Ixodes pacificus]			X

Table 2.7: List of homologous proteins identified by the three main analysis methods used in this chapter, 2D-PAGE, 1D-PAGE/1D-LC/MS/MS and 24cm IPG/1D-PAGE.

tr|Q8T7E4Glutathione S-transferase; EC=2.5.1.18;[Boophilus microplus]			X
tr|O97117|Glutathione S-transferase.[Boophilus microplus]	X	X	X
tr|Q4PLZ0|Mitochondrial malate dehydrogenase (Fragment).[Ixodes scapularis]		X	
tr|A2QMH9|Malate dehydrogenase (EC 1.1.1.37).[Aspergillus niger]			X
tr|Q5D9T1|Malate dehydrogenase (EC 1.1.1.37).[Schistosoma japonicum]			X
tr|B0QVJ3|Malate dehydrogenase.[Haemophilus parasuis 29755]			X
tr|Q6WJ29|Malate dehydrogenase (EC 1.1.1.37).[Branchiostoma belcheri tsingtauense]		X	X
tr|Q4PLY3|Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8).[Ixodes scapularis]		X	
tr|A3E3L5|Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8).[Pfiesteria piscicida]			X
tr|A1AWR2|Carboxyl-terminal protease precursor (EC 3.4.21.102).[Ruthia magnifica subsp. Calyptogena magnifica]			X
tr|A6N9L6|40S ribosomal protein S28.[Ornithodoros parkeri]		X	X
sp|P0AA25|Thioredoxin-1 (Trx-1) (Trx).[Escherichia coli]		X	X
tr|Q7JPZ2|D.melanogaster ubiquitin (Fragment).[Drosophila melanogaster]			X
tr|B0Y7M4|Thiazole biosynthesis enzyme.[Aspergillus fumigatus]			X
sp|P11147|Heat shock 70 kDa protein cognate 4 [Drosophila melanogaster]			X
tr|B0LJ59|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12).[Dictyocaulus viviparus]	X		X
tr|A8DX87|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (Fragment).[Phascolion strombus]	X		X
tr|Q6QWP2|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (Fragment).[Lithobius sp. SBH266126]		X	X
tr|Q5WCG4|Glutathione-dependent formaldehyde dehydrogenase (EC 1.2.1.1).[Bacillus clausii]			X
tr|Q9BNV9|Elongation factor 1-alpha (Fragment).[Amblyomma sp. 'Amb2']		X	
tr|Q0PPV8|Elongation factor 1-alpha (Fragment).[Pleopsidium chlorophanum]			X
tr|Q0PPX3|Translation elongation factor-1 alpha (Fragment).[Physcia aipolia]			X
tr|Q1E463|Formate dehydrogenase.[Coccidioides immitis]			X
tr|Q5XXS7|Putative arginine kinase.[Oncometopia nigricans]		X	
tr|Q4CWA5|Arginine kinase, putative (EC 2.7.3.3).[Trypanosoma cruzi]			X
ISGI-TC11435 similar to UP|Q38EU9_9TRYP (Q38EU9) Arginine kinase , partial (52%)	X		
tr|Q4PN07|Thioredoxin-dependent peroxide reductase.[Ixodes scapularis]	X		X
sp|P07754|Alcohol dehydrogenase 3 (EC 1.1.1.1) (Alcohol dehydrogenase III) (ADH III).[Emericella nidulans]			X
tr|A7E8J7|Citrate synthase, mitochondrial.[Sclerotinia sclerotiorum]			X
tr|A8X135|CBR-ARD-1 protein.[Caenorhabditis briggsae]			X
ISGI-G893P536FK9 similar to UP|Q6P6X6_BRARE (Q6P6X6) Nol5 protein, partial (29%)	X		
ISGI-TC8810 similar to UP|Q692U7_IXOSC (Q692U7) Salivary gland protein, partial (44%)	X		
tr|Q2VCI9|Aspartic protease-like.[Solanum tuberosum]	X		
AVGI-TC149	X		

Table 2.7: List of homologous proteins identified by the three main analysis methods used in this chapter, 2D-PAGE, 1D-PAGE/1D-LC/MS/MS and 24cm IPG/1D-PAGE.

tr|Q0MR22|RAD52-like protein.[Penicillium marneffei]	X		
ISGI-TC15028 homologue to UP|Q6P7N2_XENTR (Q6P7N2) MGC75629 protein (Eukaryotic translation initiation factor 2), partial (6%)	X		
ISGI-TC627 similar to UP|Q692U7_IXOSC (Q692U7) Salivary gland protein, partial (51%)	X		
ISGI-TC7696 similar to UP|Q5C5Z8_SCHJA (Q5C5Z8) SJCHGC09463 protein (Fragment), partial (14%)	X		
ISGI-G894P545FB3	X		
BMGI-TC12871	X		
ISGI-DN973124 UP|Q8FZR3_BRUSU (Q8FZR3) Oxidoreductase FAD-binding, partial (2%)	X		
ISGI-TC12609 similar to UP|Q5UAQ6_BOMMO (Q5UAQ6) Ribosomal protein L32, complete	X		
ISGI-TC4720 similar to UP|Q692U7_IXOSC (Q692U7) Salivary gland protein, partial (48%)	X		
tr|Q1HQR4|RAB family GTPase (Rab11).[Aedes aegypti]	X	X	
ISGI-DN970843 homologue to UP|Q4PMM2_IXOSC (Q4PMM2) Salivary secreted serine protease, partial	X		
ISGI-G893P564FK12 UP|Q8DQS9_STRR6 (Q8DQS9) Serine acetyltransferase , partial (5%)	X		
tr|A0SHR2|Protein disulfide isomerase (EC 5.3.4.1).[Amblyomma variegatum]		X	
ISGI-TC3271 weakly similar to UP|Q70JV3_BRUMA (Q70JV3) Protein disulphide isomerase precursor	X		
RAGI-TC1455 similar to SP|Q96B10|ABE1_HUMAN ATP-binding cassette sub-family E member 1 (RNase L inhibitor)	X		
sp|Q1HR36|14-3-3 protein zeta.[Aedes aegypti]	X	X	
ISGI-G894P559FA16 homologue to UP|Q3R2U8_XYLFA (Q3R2U8) Glycoside hydrolase, family 24, partial (9%)	X		
tr|Q8T5G9|Triosephosphate isomerase (EC 5.3.1.1) (Fragment).[Archaeopotamobius sibiriensis]		X	
tr|O35079|Polyubiquitin.[Cricetulus griseus]		X	
tr|Q6W975|Sodium/potassium ATPase alpha subunit (Fragment).[Garypus californicus]		X	
tr|B0FYV2|Eukaryotic initiation factor 4A.[Plutella xylostella]		X	
tr|Q170J7|Moesin/ezrin/radixin.[Aedes aegypti]		X	
tr|Q29KW9|Phosphoglycerate kinase (EC 2.7.2.3) (Fragment).[Drosophila pseudoobscura]		X	
tr|Q178U8|Fructose-bisphosphate aldolase (EC 4.1.2.13).[Aedes aegypti]		X	
tr|Q0TR25|Putative alpha-N-acetylgalactosaminidase.[Clostridium perfringens]		X	
tr|B0WRR2|Vacuolar protein sorting 13D.[Culex quinquefasciatus]		X	
tr|Q4PMD5|Succinate dehydrogenase Ip subunit.[Ixodes scapularis]		X	
tr|B0W1R7|Prohibitin-2.[Culex quinquefasciatus]		X	
tr|A8MLG2|Methionine aminopeptidase (EC 3.4.11.18).[Alkaliphilus oremlandii]		X	
tr|B1B544|Vitellogenin-C.[Haemaphysalis longicornis]		X	
sp|Q7PPA5|Calcium-transporting ATPase sarcoplasmic/endoplasmic reticulum (EC 3.6.3.8).[Anopheles gambiae]		X	
sp|P13395|Spectrin alpha chain.[Drosophila melanogaster]		X	
sp_vs|P54385-2|(Gdh)Isoform C of P54385.[Drosophila melanogaster]		X	
tr|Q5MM89|ADP ribosylation factor 79F.[Aedes aegypti]		X	

Table 2.7: List of homologous proteins identified by the three main analysis methods used in this chapter, 2D-PAGE, 1D-PAGE/1D-LC/MS/MS and 24cm IPG/1D-PAGE.

tr|Q0PZI5|Eukaryotic initiation factor 4A.[Callinectes sapidus]		X	
tr|Q8IT76|Alpha-2-macroglobulin precursor splice variant 1.[Ornithodoros moubata]		X	
tr|A2CEW4|5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase.[Danio rerio]		X	
tr|Q6B864|Cytochrome b5.[Ixodes pacificus]		X	
tr|A8P0E9|Isocitrate dehydrogenase (NAD) subunit alpha, mitochondrial, putative (EC 1.1.1.41).[Brugia malayi]		X	
tr|Q6WNX0|Ferritin.[Ixodes scapularis]		X	
tr|Q16PZ7|Cytoplasmic dynein light chain.[Aedes aegypti]		X	
tr Q2XW13 Phospholipid-hydroperoxide glutathione peroxidase [Boophilus microplus)		X	
tr|Q6B8D1|Putative salivary secreted peptide.[Ixodes pacificus]		X	
tr|Q178I1|Vinculin.[Aedes aegypti]		X	
tr|A5EQ81|Hydrogenobyrinic acid a,c-diamide cobaltochelatase (EC 6.6.1.2).[Bradyrhizobium sp.]		X	
tr|Q4PLZ7|Probable microsomal signal peptidase 22 kDa subunit.[Ixodes scapularis]		X	
tr|Q09JF2|Cytochrome c oxidase subunit Va.[Argas monolakensis]	X		
tr|A1V933|Iron-sulfur cluster-binding protein, Rieske family.[Burkholderia mallei]	X		
tr|B0UTS7|AMP-dependent synthetase and ligase.[Haemophilus somnus]	X		
tr|Q9GPM1|9.8 kDa basic protein.[Amblyomma hebraeum]		X	
tr|Q11AW8|2-octaprenylphenol hydroxylase (EC 1.14.13.-).[Mesorhizobium sp.]		X	
tr|Q5MM89|ADP ribosylation factor 79F.[Aedes aegypti]		X	
tr|A9QQC2|Cofilin.[Lycosa singoriensis]		X	

Table 2.7: List of homologous proteins identified by the three main analysis methods used in this chapter, 2D-PAGE, 1D-PAGE/1D-LC/MS/MS and 24cm IPG/1D-PAGE.

Experiment	Peptide matches above identity threshold	Decoy matches above identity threshold	False discovery rate	Number of homologous proteins identified
Tris extract 200ug 1D	808	17	2.1%	285
UTC7 extract 200ug 1D	1215	20	1.65%	446
24cm IPG + 1D	528	6	1.14%	104

Table 2.8: Number of homologous proteins found in shotgun experiments (section 2.3.3 and 2.3.5) and the false discovery rates at a significance threshold of $p<0.01$. False discovery rate was determined by Mascot performing a decoy database search.

Chapter three – Zymography with fluorescent substrates.

Preface.

The work presented in this chapter was carried out concurrently with the work presented in chapter two, where a number of non-hypothesis driven fractionation techniques were assessed to ascertain what work flow would produce the greatest number of protein identifications. The results showed that the techniques are complimentary and each provides identifications that are unique to the technique. In addition, the complementarity is shown by the strengths and weaknesses of each technique. The combination of the MCE with 2-DGE using narrow range IPG strips provides unrivalled resolution of nearly 1000 spots with high abundance cytoskeletal proteins well isolated from other spots. However, the excision, tryptic digestion and analysis by LC/MS/MS of those 1000 gel spots is an almost impossible task without the use of robotics for the digestion. The use of 1-D SDS-PAGE to fractionate the sample is a robust technique that utilises the solubilising power of SDS to extract proteins not soluble in the reagents used for isoelectric focusing. However, the technique is sensitive to high abundance proteins which smear through the gel lane at high sample loads. These high abundance proteins then produce high abundance peptides which are selected first for fragmentation by the mass spectrometer.

One way to avoid these high abundance proteins is to use techniques that ignore their presence and focus on other sub sections of the proteome, such as enzymes. This work was carried out after reading a paper by Zhao and Russell (2003) in which enzymatically cleavable substrates were co-polymerised into SDS-PAGE gels and the substrate cleaved by enzymes subjected to separation by 2-DGE. The attractiveness of this technique was that it can be used with unfractionated samples as fluorescent spots in the gel can only be

the result of a protein possessing an enzymatic activity specific for the substrate. In contrast to the traditional gelatine of casein zymogram, the fluorescent spots can be directly excised from the gel, washed, trypsin digested and the resulting peptides analysed by LC/MS/MS without contamination from gelatine or casein. The assay was used to ascertain the presence of serine proteases in tick proteins fractionated by solubility. In this way, the fractionation techniques developed in the non-hypothesis driven approach were applied to a hypothesis driven approach.

3.1.1: Fluorescent zymograms.

As described in the introduction to chapter two, one way to fractionate a specific subtype of proteins is to use a specific property of the protein subtype, such as enzymatic activity. Traditional solution based assays have used whole unfractionated saliva, salivary gland extracts (SGE) or mid guts, or chromatographic subfractionation of these samples and this can consume significant quantities of sample while appropriate separation and assay conditions are investigated. Mendiola *et al* (1996) used 2000 excised mid guts from *R. microplus* which were further fractionated by anion exchange chromatography to study the proteolytic activities present, finding aspartic and cysteine proteases activities at acid pHs. However, the majority of enzymes found were discovered through searching cDNA libraries with primers designed from highly conserved regions of the enzymes rather than isolating and characterising the protein possessing the enzymatic activity. Localisation of the enzyme within the tick was determined by RT-PCR analysis of dissected tick organs (Mulenga *et al*., 2003, Mulenga *et al*., 1999a).

An alternative is to perform enzymatic assays on samples separated by isoelectric focusing and SDS-PAGE, adapting the zymogram method of Zhao and Russell (2003). Proteins are sequentially fractionated as described in chapter two (section 2.2.1) but not reduced and alkylated. These samples are then focused on pH 3-10 IPG's and separated by SDS-PAGE in gels co-polymerised with specific peptide substrates containing an enzymatically cleavable methyl courmaryl amide group (MCA). The focused and separated proteins are renatured in the gel with a physiological buffer before incubation at 37^{O}C. Renatured proteins with enzymatic activity against the substrate cleave the MCA group which fluoresces under UV light. Thus spots are seen where enzymatic activity is occurring and these spots can be excised from the gel and analysed by mass spectrometry as described in section 2.2.6.

This work is focused on finding protein spots with serine protease-like activity. The work of Mulenga *et al* (2003) has previously described the presence of enzymes of this activity in fed-tick midgut using RT-PCR. The choice of serine protease activity was due to the availability of substrate and availability of trypsin as a control. Serine proteases are potential tick vaccine candidates as they are involved in regulating several physiological and developmental processes in a number of organisms (Krem and Di Cera, 2001, Neurath, 1984, Neurath, 1986, Rao *et al.*, 1998) and they are generally secreted proteases and thus likely to be exposed to the host immune system (Allingham *et al.*, 1992, Ben-Yakir, 1989, Vaughan and Azad, 1988). Recent work by Miyoshi *et al* (2008) in *H. longicornis* has shown that silencing of three serine protease by RNA interference leads to a significant reduction in engorged tick body weight.

3.2: Methodology.

Please note: this work was commenced and completed before it was decided to repeat and expand on the work presented in chapter two and thus the sample preparation approach is not as rigorous as employed in chapter two.

3.2.1: Sample preparation.

Figure 3.1 diagramatically outlines the sequence of experimental methods used in the work presented in this chapter. Whole unengorged female ticks were collected from bushland in the Lennox Head area of NSW. Whole unengorged female ticks (50) were snap frozen in liquid nitrogen and ground to a fine powder with a mortar and pestle. Unengorged male ticks (50) were also prepared in parallel. The powder was transferred to a centrifuge tube and 2mL of 40mM Tris-HCl pH 8.8 added. The sample was then sonicated directly with an ultrasonic probe for 6 x 30 seconds on ice. The sample was then centrifuged at 40 000g for 10 minutes at room temperature. The supernatant containing

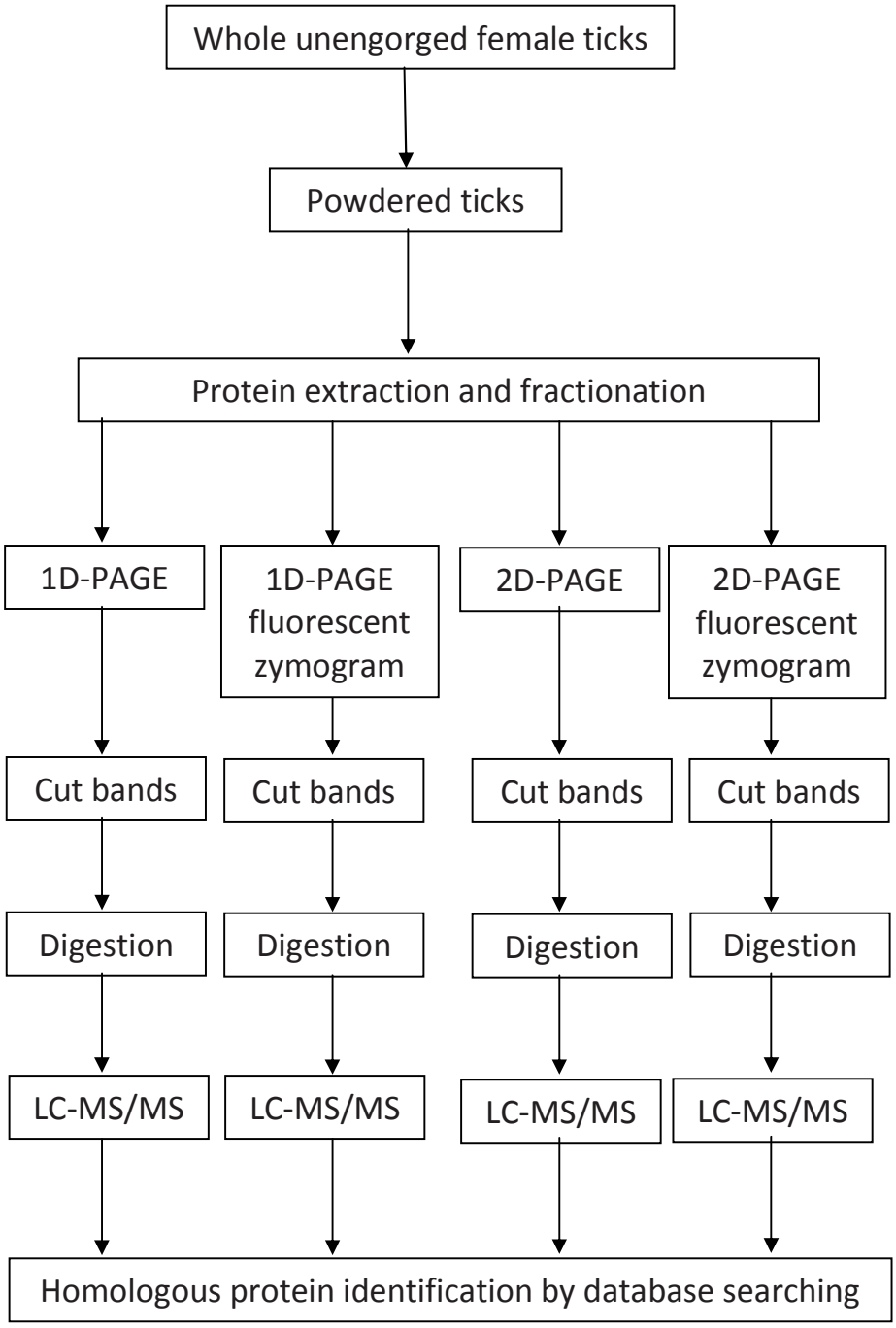

Figure 3.1: Diagramatic outline of the experimental steps carried out in this chapter. Exact methods are described in detail in the body text of the chapter.

soluble proteins was removed. In later experiments, this process was repeated three times (fraction one). The pellet was resuspended in 2mL of 7M urea, 2M thiourea, 1% C7BzO and 40mM Tris-HCl pH 8.8 and subjected to the same sonication and centrifugation procedure. The supernatant was removed and labelled fraction two. Fractions were buffer exchanged into 7M urea, 2M thiourea, 1% C7BzO and the conductivity checked with a Horiba B-173 conductivity meter to ensure it was below 200µS.

3.2.2: 1-D SDS-PAGE/Fluorescent zymography.

After adding 2x sample buffer containing SDS, 15µL of the Tris-HCl and UTC7Tris-HCl extracts were loaded into the wells of a 10% polyacrylamide Tris-HCl mini-gel that had been co-polymerised with 200µM of peptide-MCA substrate specific for the enzyme type being analysed. For enzymes with trypsin-like activity, Boc-Gln-Ala-Arg-4-methyl-coumaryl-7-amide was used and 62.5ng of porcine trypsin loaded as a positive control. Gels not containing peptide substrate were also performed to ascertain the amount of protein loaded. The gel was run at 200V until the bromophenol blue dye front had been run of the gel. The gel was then washed for 7 x 5 minutes with cold water and soaked in 6.3mM Bicine, 0.1M NaCl, pH 9.0 for 30 minutes at 37^{O}C. The gel was then imaged with a Bio-Rad ChemiDOC XRS CCD camera system using 365nm emission wavelength bulbs and a coumarin (480nm) band pass filter. Fluorescent spots indicating enzymatic activity were manually excised from the gel and the gel incubated at room temperature overnight. The gel was then reimaged and previously undetected fluorescent spots excised manually or using a Bio-Rad EXQuest spot cutter. Gels not containing peptide substrate were fixed with 40% methanol, 10% acetic acid for 60 minutes before staining with Bio-Rad Flamingo fluorescent protein stain for 90 minutes. Gels were then imaged using a Bio-Rad PharosFX Plus Molecular Imager utilising a 532nm excitation laser and 605nm band pass filter. Following this, the gels were overstained with colloidal Coomassie blue stain overnight and spots excised with a Bio-Rad EXQuest spot cutter.

3.2.3: Isoelectic focusing.

Bio-Rad 11cm pH 3-10 IPG Readystrips were partially rehydrated with 100µL of the buffer exchanged fraction one or two for 30 minutes at room temperature. A further 150µL of sample was applied using active loading by pipetting the sample under the IPG strip which had been laid facedown and covered in paraffin oil in a Proteome Systems IsoelectrIQ focusing tray. The sample is then drawn into the strip by applying a slowly increasing voltage to the strip. The IsoelectrIQ instrument was programmed in the following way to focus the proteins in the sample. Step 1: 100-3000V in a convex ramp over 5 hours with current limited to 75µA/strip. Step 2: 3000-10000V in a linear ramp over three hours with current limited to 50µA/strip. Step 3: 10000V with a current limit of 35µA/strip until focusing had reached at least 100000 V hours.

3.2.4: 2-D SDS-PAGE/Fluorescent zymography.

Focused IPG strips were equilibrated for 25 minutes in 7M urea, 250mM Tris-HCl pH 8.5, 2% SDS, with bromophenol blue. Once equilibrated, the strips were laid on top of an 11cm midi-sized 10% polyacrylamide Tris-HCl gel that had been copolymerised with 200uM of peptide-MCA substrate specific for the enzyme type being analysed. The gels were run and zymography carried out as described in section 4.2.2

3.2.5: Protein digestion, mass spectrometry and data analysis

Please refer to section 2.2.5, 2.2.6 and 2.2.7 for details of these methodologies.

3.3: Results.

3.3.1: Fluorescent zymograms.

Figure 3.2 shows the fluorescent zymogram of the extracts separated by 1-D SDS-PAGE. This was performed to ascertain whether the assay worked, using porcine trypsin as a positive control, and if the tick extract indeed contained proteins with serine protease activity. The trypsin worked so well that the lane of the gel in which it was loaded had to be removed as the brightness, due to high amounts of substrate cleavage, was interfering with the sensitivity of the assay. Some of the substrate cleaved by the trypsin can be visualised on the right of the zymogram in figure one. A number of bands are clearly visible in the zymogram with the majority being in the UTC7Tris-HCl extract. The band pattern in the zymogram is the same for both male and female ticks, even though there are differences in the band patterns in the Flamingo stained gels which is to be expected. As the male tick does not feed directly on the host, but from the female tick, it was not used in further work as its proteins would never been seen by the host immune system.

Eighteen bands in total were cut from both Flamingo stained and zymogram gels representing nine bands showing enzymatic activity. The corresponding band was cut from the Flamingo gel to determine whether the zymogram methodology adversely affected trypsin digestion of the protein. Only three bands produced identifications and these were for the cytoskeletal proteins actin, paramyosin and tropomyosin. It is possible that these proteins are co migrating in the gel with the proteins possessing the serine protease activity breaking down the substrate.

Figure 3.3 shows the 2-D gels with and without enzymatic substrate for sequentially extracted fractions one (Tris) and two (UTC7Tris-HCl) from female ticks. As these protein

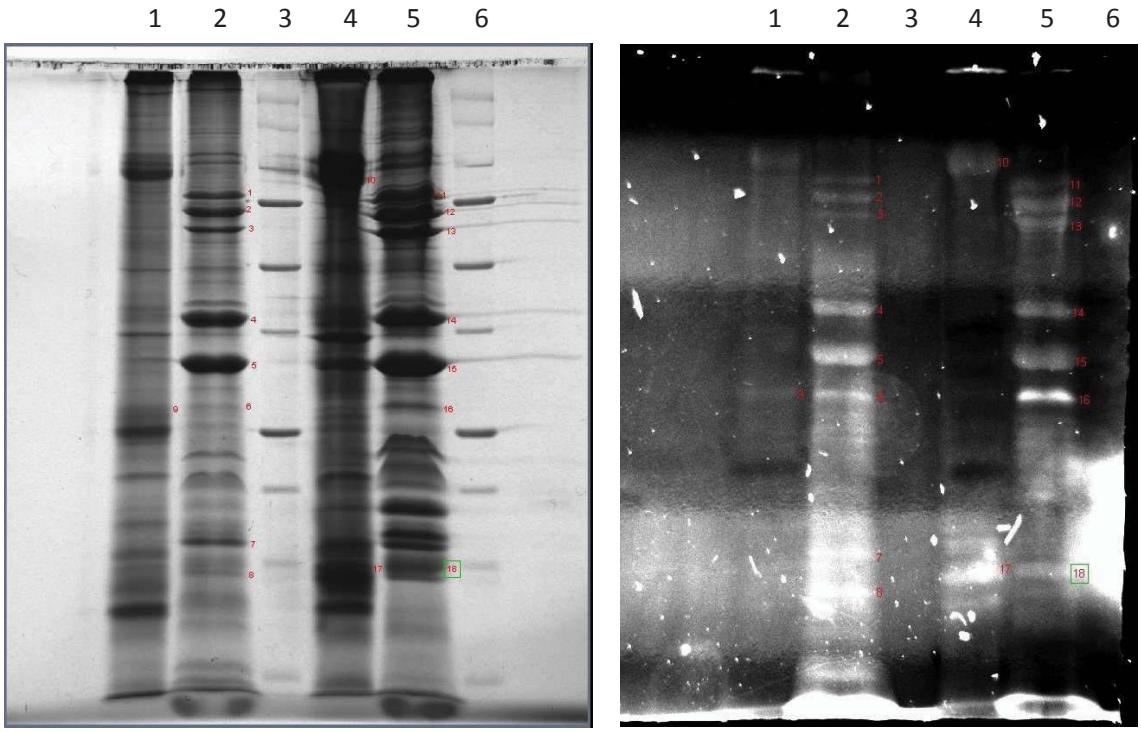

Figure 3.2: Fluorescent zymogram of unengorged male and female *I.holocyclus* separated by 1-D SDS-PAGE. The left gel (A) has been stained with Flamingo fluorescent stain and the right gel (B) was copolymerised with Boc-Gln-Ala-Arg-4-methyl-coumaryl-7-amide to reveal the presence of enzymes with serine protease activity. The numbers beside individual gel bands indicate corresponding bands cut from both gels for trypsin digestion. The large bright spot on the right side of the zymogram (right gel) is caused by 10ng of porcine trypsin acting as a positive control.

1) Male Tris-HCl extract
2) Male UTC7Tris-HCl extract
3) MW markers
4) Female Tris-HCl extract
5) Female UTC7Tris-HCl extract
6) MW markers

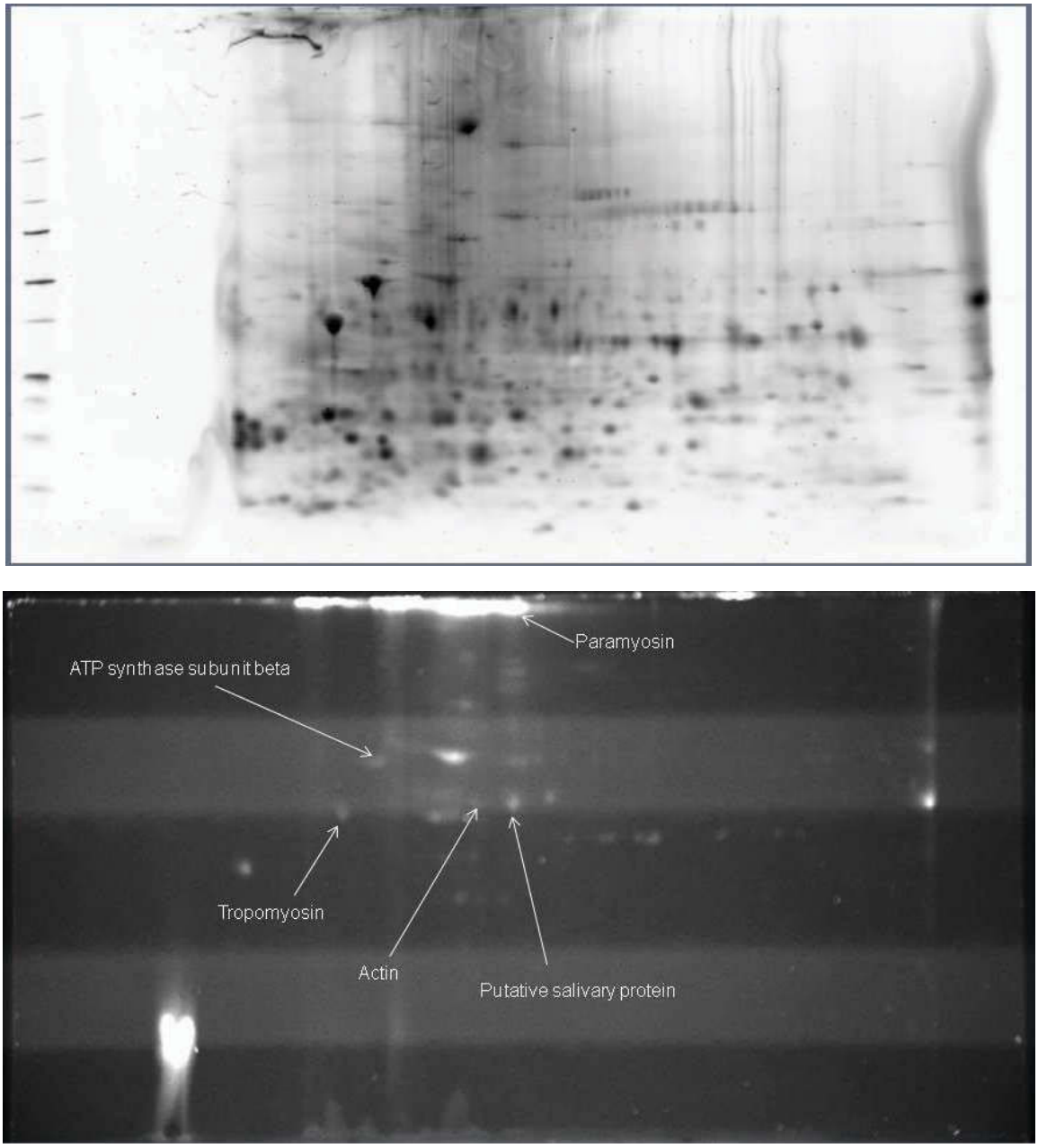

Figure 3.3: 2-D fluorescent zymogram gels of *Ixodes holocyclus* protein extracted with 40mM Tris-HCl pH 8.8 (fraction 1). The upper gel (A) has been stained with Flamingo fluorescent stain showing total protein in the sample while the lower gel (B) has been copolymerized with the serine protease substrate Boc-Gln-Ala-Arg-4-methyl-coumaryl-7-amide and spots reflect where serine protease activity is present.

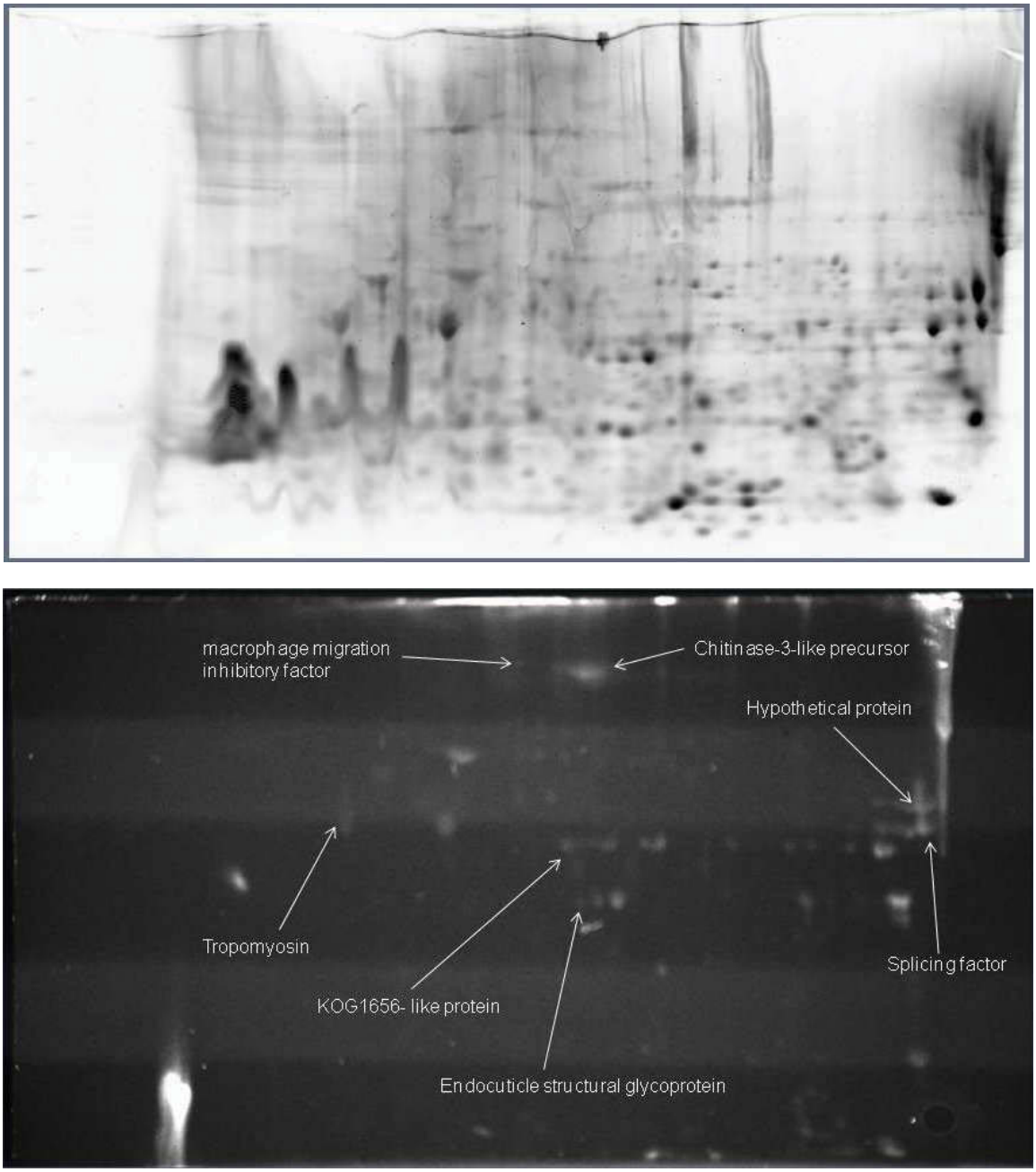

Figure 3.4: 2-D fluorescent zymogram gels of *Ixodes holocyclus* protein extracted with 7M urea, 2M thiourea, 1%C7BzO and 40mM Tris-HCl pH 8.8 (fraction 2). The upper gel (C) has been stained with Flamingo fluorescent stain showing total protein in the sample while the lower gel (D) has been copolymerized with the serine protease substrate Boc-Gln-Ala-Arg-4-methyl-coumaryl-7-amide and spots reflect where serine protease activity is present.

samples were not reduced and alkylated prior to focusing, the resolution of individual spots is poor when compared to gels where proteins have been reduced and alkylated prior to focusing (figure 2.6A and B). Fraction two does contain a few spots also seen in fraction one, however this is likely due to only one Tris extraction being done. Further work has shown a second Tris extraction removes these proteins from fraction two (data not shown). Spots of similar pI and MW are also observed in all of the gels copolymerised with enzymatic substrates, especially one located around pH 4 where no other spots are present. However in both cases the identity of this spot was unable to be made with all data analysis methods used.

By comparing the Flamingo-stained gels, displaying the total protein present in the extractions, with the gels containing the MCA substrate, it is apparent that only a small proportion of the protein spots are enzymes with serine protease-like activities. In the Tris-fraction, 22 fluorescent spots were excised and the UTC7Tris fraction had 28 spots excised for identification from the trypsin substrate gels. All of spots generated tryptic peptides that were able to be fragmented by MS/MS. However, only nine spots in the Tris-fraction and one (Tropomyosin) in the UTC7Tris-fraction matched proteins in the databases searched with any confidence (Table 3.1). These identifications are shown in table one.

3.3.2: Protein identifications.

The majority of the identifications were proteins that are not known to have trypsin-like activity, which was surprising. In fact, few enzymes were found. A protein with homology to an aspartic protease from *Haemaphysalis longicornis* was found in spot 8 of the Tris-extract, but the score given by Peaks ProteinID (71%) indicates that this identification needs to be verified. Aspartic proteases are most active at acid pH and preferentially cleave at aromatic containing amino acids. It may be that the extended incubation time of the gel allowed this protease to degrade the substrate at a reduced rate that would have not been detectable with a shorter incubation time. A suitable

substrate for the aspartic protease, renin, is available and the assay will be repeated using this substrate to better characterise this spot. Spot 10 was tentatively identified as a Midgut cysteine proteinase from *R. appendiculatus* by Peaks ProteinID with a score of 54%, but again this must be verified. Some cysteine proteases are known to have a broad specificity so this spots activity against the serine-protease substrate is not unlikely. However with a score of 54%, this spot needs further investigation to confirm its identity.

The BLAST searching of the ESTs matching two spots from the UTC7Tris extract, spots 22 and 24, identified two enzyme types that were unexpected, phosphoenolpyruvate carboxykinase and fructose 1,6-bisphosphate aldolase. These enzymes are involved in gluconeogenesis and glycolysis respectively and are unlikely to cleave the peptide substrate used in the assay. However, as the gel was incubated overnight at room temperature it maybe that these enzymes can catalyse the breakdown of the substrate at a greatly reduced rate than some of the other enzymes present in spots showing fluorescence. The BLAST scores would indicate that the EST used in the search is related to the proteins identified. Spot 24 had an Peaks ProteinID score of 77% with 7% coverage and 3 peptide matches to the EST that has homology to fructose 1,6-bisphosphate aldolase, giving some confidence in the identification. In contrast, spot 22 has only a score of 13% and thus this identification is questionable.

Spot 28 has homology to uncharacterised salivary proteins from *I.scapularis* (Guilfoile and Packila, 2004), *I.ricinus* (Leboulle *et al.*, 2002) and *I.pacificus* (Francischetti *et al.*, 2005) whose genes are upregulated during feeding. These identifications are curious because the ticks used in this work were unfed and may represent constitutive expression of the protein. Spot 30 and 31 match to different ESTs in ISGI, but both have homology to the same uncharacterised salivary protein from *I.scapularis* (Guilfoile and Packila, 2004). The mRNA for the salivary protein identified in spot 31 was also found to be expressed during blood feeding.

Spot #	Description/Species	pI	MW	Score	Matches	Coverage (%)	Database	Engine
3	Cytochrome b - *O.moubata*			90%	2	3	Acari	InChorus
4	unknown - *ixodes scapularis*		41891	90%	3	9	Acari	InChorus
	eukaryotic translation initiation factor 1A domain						NCBInr	BLASTp
5	ATP synthase subunit beta - *Encope michelini*	4.95	45882	428	7	22	TrEMBL	Mascot
	ATP synthase subunit beta - *Encope michelini*	4.95	45882	99%	7	22	TrEMBL	InChorus
	cytochrome b - *Ornithodoros moubata*		41891	90%	3	5	Acari	InChorus
6	Aspartic Protease - *Haemaphysalis longicornis*		41971	44%	1	2	Acari	InChorus
7	Actin - *Ixodes ricinus*		41563	99%	3	10	Acari	InChorus
	Actin	5.24	38759	116	3	11	VISDB	Mascot
	Aspartic Protease - *Haemaphysalis longicornis*		41971	84%	5	2	Acari	InChorus
8	Aspartic Protease - *Haemaphysalis longicornis*		41971	71%	4	5	Acari	InChorus
10	Midgut cysteine proteinase - *R. appendiculatus*		37332	56%	4	3	Acari	InChorus
11	Tropomyosin - *Haemaphysalis longicornis*	4.69	32853	721	14	34	VISDB	Mascot
	Tropomyosin - *Haemaphysalis longicornis*		32872	99%	21	35	Acari	InChorus
18	Troponin T - *Haemaphysalis qinghaiensis*		121665	95%	5	6	BMGI EST	InChorus
19	Paramyosin - *Rhipicephalus (Boophilus) microplus*		102285	1622	40	38	SwissProt	Mascot
	Paramyosin - *Rhipicephalus (Boophilus) microplus*		101992	99%	33	44	Acari	InChorus
20	Myosin heavy chain, muscle - *Drosophila melanogaster*	5.75	222606		10	5	SwissProt	Mascot
	Paramyosin - *Rhipicephalus (Boophilus) microplus*		101992	99%	14	14	SwissProt	Mascot
	Actin - *Ixodes ricinus*		41563	99%	12	26	Acari	InChorus
	Paramyosin - *Rhipicephalus microplus*		101992	99%	14	14	Acari	InChorus
21	Paramyosin - *Rhipicephalus (Boophilus) microplus*	5.53	102285	1613	36	38	SwissProt	Mascot
	Paramyosin - *Rhipicephalus (Boophilus) microplus*		101992	99%	78	45	Acari	InChorus
22	Paramyosin - *Rhipicephalus (Boophilus) microplus*		102285	676	15	21	SwissProt	Mascot
	Paramyosin - *Rhipicephalus (Boophilus) microplus*		101992	99%	23	26	Acari	InChorus
7-F2*	*Haemaphysalis longicornis* tropomyosin mRNA		32853	400	10	23	VISDB	Mascot
	Tropomyosin - *Haemaphysalis longicornis*		33002	99%	10	28	Acari	InChorus

Table 3.1: MS/MS ion search results for spots showing flucrescence in zymographic gel copolymerised with the trypsin-like substrate Boc-Gln-Ala-Arg-4-methyl-coumaryl-7-amide.

Spot#	EST ID	BLAST result	NCBI gene	Total Score	coverage	Identity	E-value
Tris-Extract							
3	ISGI-TC8398 similar to UP|Q4RH48_TETNG Chromosome undetermined	*Ixodes ricinus* mRNA for actin	gi:60730229	1380	93	95	0
10	BMGI-TC6898	*Colwelliopsychrerythraea* 34H, complete genome	gi:71143482	53.6	4	86	0.002
UTC7-Extract							
5	BMGI-TC12106	*Haemaphysalis qinghaiensis* troponin T mRNA	gi:65332125	1175	67	85	0
15	ISGI-GB94P541F31 similar to UP|Q2F5J2_BOMMO Prohibitin protein	*Taeniopygia guttata* prohibitin variant 1-like mRNA	gi:76158349	181	53	72	3.00E-42
18	BMGI-TC12024	NADH dehydrogenase subunit 5 -*Rhodactis sp*	gi:108515256	48.2	1	81	
		adenosine triphosphatase 8 & 6 -*Apis mellifera ligustica* (mitochondria)	gi:336293	48.2	1	81	
22	BMGI-TC12304	PREDICTED: *Ornithorhynchus anatinus* similar to Phosphoenolpyruvate carboxykinase, cytosolic GTP mRNA	gi:149639780	646	75	70	0
24	BMGI-TC12043	PREDICTED: *Nasonia vitripennis* similar to putative fructose 1,6-bisphosphate aldolase (LOC100118253), mRNA		556	72	72	8.00E-155
	ISGI-TC7517 similar to UP|Q75PQ3_ANTYA Fructose 1,6-bisphosphate aldolase	*Aedes aegypti* fructose-bisphosphate aldolase mRNA	gi:157111183	464	83	75	2.00E-127
28	BMGI-TC11621	*Ixodes ricinus* mRNA for hypothetical protein (ORF1)	gi:12054697	362	52	82	2.00E-44
30	ISGI-TC5991 homologue to UP|Q692U7_IXOSC Salivary gland protein	*Ixodes scapularis* hypothetical salivary protein mRNA		1316	65	88	3.00E-176
31	ISGI-GB93P55RA5 similar to UP|Q692U7_IXOSC Salivary gland protein	*Ixodes ricinus* partial mRNA from salivary gland induced during blood feeding	gi:5911723	670	62	74	2.00E-65

Table 3.2: BLAST search results for fluorescent spots matched to EST sequences by Peaks InChorus homology search.
BMGI – *B.microplus* Gene Index.
ISGI – *I.scapularis* Gene Index.

any of the spots exhibiting enzymatic activity were identified as proteins associated with the cytoskeleton. Tropomyosin is found at the same spot in both extracts with high confidence by both Mascot and Peaks ProteinID. It is an important allergen found in shellfish, other invertebrates and arthropods, having a highly conserved amino acid sequence (Jeong *et al.*, 2006). Tropomyosin, along with troponin C, I, and T subunits, regulates the contraction of skeletal muscle by binding actin monomers in a thin filament. It is also found in non-muscle cells and is believed to function by providing mechanical support of the cytoplasmic membrane and in the transportation of other molecules. However, there is no documented evidence that tropomyosin has any enzymatic activity. Paramyosin, found only in the Tris-extract, is also an invertebrate muscle protein, being described as multi-functional whose roles may include host immunomodulation (Ferreira *et al*., 2002a). Once again, paramyosin has not been reported to have serine protease activity.

Spot 4 in the Tris-extract was identified as Unknown found in *I. scapularis*. A BLAST search showed high homology with eukaryotic translation initiation factor 1A domain, which is required *in vitro* for maximal rates of protein synthesis in mammalian systems functioning primarily by dissociating ribosomes and stabilizing 40 S pre-initiation complexes (Wei *et al.*, 1995).

Spot 22 in the Tris-extract was identified as ATP synthase subunit beta from a variety of species. The subunit is part of a multi-subunit non-phosphorylated ATPase found in the mitochondria and thylakoid membrane. The beta subunit is part of the extramembrane part of the complex.

The searching of tick EST libraries with MS/MS data and Peaks ProteinID, followed by BLAST searches of the EST assigned identifications to 11 other protein spots (Table 3.1 and 3.2) as well as confirming proteins identified in Mascot and ProteinID searches of

available protein databases (data not shown). Some of these were cytoskeletal proteins (actin and tropinin T) whereas others had very high homology to salivary protein mRNA from other ticks. One of great interest is spot 31 from the UTC7-extract having homology to an *I. ricinus* mRNA from salivary gland found in a subtractive methodology between unengorged and engorged ticks, being upregulated during blood feeding (Guilfoile and Packila, 2004).

The remaining spots showing enzymatic activity did not match any entries in the protein databases searched by MS/MS ion searches. A lack of quality MS/MS spectra prevented further characterisation of the protein in the spots using *de novo* sequence determination.

3.4: Discussion.

The subfractionation of proteomes, as stated in chapter two, is becoming a powerful tool to remove high abundance proteins and also focus on specific types of proteins. In the work presented in this chapter, the whole unengorged adult female tick proteome has been subfractionated only by solubility and then analysed to display proteins with specific enzymatic activity. This workflow thus combines the non-hypothesis driven approach of fractionating by solubility and combines it with the hypothesis driven approach of assaying for specific enzymes subtypes. Activity was identified by adapting a novel fluorescent assay performed in a two-dimensional SDS-PAGE experiment that avoids the necessity of purifying the proteins in numerous steps as many other studies have done. Spots displaying enzymatic activity are then excised from the gel and identified by LC/MS/MS. Whole ticks were used because of the ease of preparation of the sample and the fact that successful tick vaccines to date have not been salivary proteins, but concealed antigens such as Bm86 (Nuttall *et al*., 2006). The most attractive target or targets for a vaccine is a protein present at the onset of feeding so the immune system can react immediately after attachment, leading us to

test this technique on unengorged whole ticks. Also, the work of Mulenga *et al* (2003) localised serine protease activity to tick midgut rather than salivary glands. However, the use of whole ticks as starting material was driven more by time constraints preventing organ dissection and a lack of a tick colony in the laboratory. Unengorged ticks have the advantage of not being contaminated with host proteins with protease activity that would also produce fluorescent spots with possibly far greater intensity. Finally, the combination of a zymogram measuring a specific enzyme subtype with the resolving power of 2-DGE should negate the need for the complex fractionation steps needed when looking at larger proteomes rather than the enzyme subset examined here.

The aim of this study was to use this novel methodology to identify protein spots with enzymatic activity that could be potential vaccine candidates. However many of the proteins identified are not known to have enzymatic activity. This was surprising and unexpected. The difference in spot patterns observed between the Tris-soluble and UTC7Tris-soluble extracts and between the different substrates with the same sample suggests that the fluorescent spots are a result of enzymatic breakdown of the substrate and not an artefact. It is possible that due to the samples being prepared under native conditions, proteins may be complexed or aggregated and remain bound throughout electrophoresis or that the use of broad range IPG's (pH 3-10) has not sufficiently resolved the proteins. However examination of MS/MS data for the spots containing tropomyosin and paramyosin did not indicate the presence of peptides belonging to other proteins, although proteins with enzymatic activity could be co-migrating at a concentration below the limit of detection. These two proteins have only been reported to bind to other cytoskeletal components such as actin, the troponin complex and myosin, so it is not known whether tropomyosin and paramyosin bind other proteins with serine protease activity in a specific or non-specific manner.

A small number of enzymes were found using this methodology, but not with the activity expected. Proteins with homology to an aspartic protease from *H. longicornis* (discussed in section 2.4.9) were identified in spots 6, 7 & 8 and a cysteine proteinase from *R. appendulatus* in spot 10. This enzyme, known as longespin (Boldbaatar *et al.*, 2006) has been shown to be expressed in the midgut and salivary glands of *H. longicornis* with expression increasing during feeding. It is able to hydrolyse Haemoglobin under acidic conditions (pH 3.5) and aspartic proteases from other blood-feeding organisms, such as *Plasmodium* and *Schistosoma* are to play a key role to haemoglobin proteolysis. In *R. microplus* it has been found that protease activity in midgut extracts is optimum at pH 3.0 with no activity found above pH 6.0 and aspartic and cysteine activities were identified using specific protease inhibitors (Mendiola *et al.*, 1996). Unfortunately there are a limited number of substrates available for these enzymes. There is a substrate available for the assay that need to be examined, Suc-Arg-Pro-Phe-His-Leu-Leu-Val-Tyr-MCA, a substrate for renin and proteinase A which are both aspartic proteases.

In a recent publication by Miyoshi *et al* (2008), recombinant versions of serine proteinases from *H. longicornis* were shown to have a pH optimum of 5.0. This is in contrast to serine proteinases from other sources which have pH optimums in the alkaline range and would explain why no serine proteinases were identified when conducting the assay at alkaline pH. The different pH optimum could be due to the enzyme's localization in the midgut lumen which is believed to have a pH between 6.3-6.5 during feeding (Coons *et al.*, 1989). In addition, optimum activity of tick intracellular digestive enzymes is thought to be acidic (Reich and Zorzopulos, 1978). Therefore, the assay needs to be repeated at acidic pH (5-6) with the Boc-Gln-Ala-Arg-4-methyl-coumaryl-7-amide substrate.

Of the 50 spots excised, only 10 were identified by the widely used method of MS/MS ion searching with algorithms such as Mascot and Peaks Peaks ProteinID using

available protein databases. Eleven more were identified from tick EST databases. This result is similar to those obtained in chapter two and demonstrates the deficiencies of MS/MS ion based searching and the datasets available to tick researchers. MS/MS ion based searching cannot deal with mutations in protein sequence as amino acid substitutions changes not only the precursor ion match, but also the fragmentation spectra. A single amino acid substitution can mean that a peptide is not able to be matched. The researcher must then resort to *de novo* peptide sequencing and the subsequent use of algorithms such as SPIDER (Han *et al*., 2005), or BLAST searching the *de novo* determined sequences. This is time consuming as fragmentation spectra must be manually validated to ensure the quality of the sequence used for searching. This work resorted to *de novo* sequencing when other methods failed to identify the protein present in a spot. No proteins were identified by *de novo* sequencing because either the MS/MS data was not of sufficient quality to obtain reliable sequence or the sequence matched no entries when BLAST searched against NCBI non-redundant database.

In summary, a novel, gel-based assay has been used to separate and characterise tick proteins possessing enzymatic activity. Twenty-one proteins were identified. However, these proteins were mainly cytoskeletal and not reported to have enzymatic activity and no proteins with serine protease activity were found although some spots matched to tick salivary proteins whose functions have not been characterised. However, with the recent findings of Miyoshi *et al* (2008) showing acidic pH optimums for *H. longicornis* serine proteinases, the assay needs to be repeated at acidic pH. Further work with this assay would involve the trialling of different substrates to quickly identify protein spots with specific types of enzyme activities. It is likely that repeating the assay with partially engorged ticks would reveal a different pattern of spots containing serine protease activity as the work of Mulenga *et al* (2003) showed much greater expression in fed ticks. This sample would be contaminated with host proteins and enzymes which will also produce spots as they degrade the substrate. The

work presented in the next chapter focuses on using equaliser technology (Boschetti *et al*., 2007) to reduce the amount of host protein and enrich tick proteins present in preparations of engorged ticks, thus overcoming the problem. The methodology maintains the protein in its native form and thus it can be assayed with the described gel based enzymatic assay.

Chapter Four – Equalisation technology and its application to the *Ixodes holocyclus* proteome.

Preface.

The previous chapter detailed a fluorescent zymogram assay that utilised the resolving power of 2-DGE to isolate single protein spots with a certain enzymatic activity. However, the majority of the proteins identified were the high abundance cytoskeletal proteins that caused problems in chapter two. Their identification, rather than enzymes being identified, was likely due to two reasons. Firstly, Miyoshi *et al's* (2008) recent finding (published online in late Sept, 2008) that recombinant *H. longicornis* serine proteases have a pH optimum of 5.0 means the assay is being performed at the wrong pH. Secondly, the resolving power of 2-DGE will be compromised by not being able to reduce and alkylate the proteins. Thus some strong protein-protein interactions may still be present and thus, co migration of the cytoskeletal proteins with enzymes may be occurring. The work of this thesis has thus returned to the problem of removing the high abundance cytoskeletal proteins. If there are important tick proteins bound to cytoskeletal proteins, they may be depleted by any technique designed to remove abundant proteins. This will therefore, be a test of the protein-protein interaction hypothesis.

Our laboratory was fortunate to be asked by Bio-Rad to beta-test a new sample preparation technology called Proteominer. Proteominer is an equalisation method that seeks to "normalise" the levels of all proteins in the same sample, reducing the concentration of high abundance proteins and enriching for lower abundance proteins. This global treatment of the sample is therefore a non-hypothesis driven technique in the same way many of the techniques presented in chapter two are. The work presented in this chapter details the use of Proteominer to equalise tick protein samples. One sample that would test the limits of the equalisation technology is engorged ticks, which, in protein concentration and dynamic range terms, must be one

of the most extreme examples of a mixed proteome as the great majority of the sample is host blood rather than tick proteins. However, if this sample could be equalised, the result may not only reveal what tick proteins are present during feeding but also what host proteins related to the host immune response are taken up.

The technology at present has the main constraint of needing the proteins in the sample to be soluble. This has implications for the types of sample preparation and prefractionation techniques that can be applied. Additionally, the protein-protein interactions mentioned above could affect the “equality” of the equalisation. Some novel methods to increase sample solubility and disrupt the interactions were attempted with mixed results. Finally, the equalised engorged tick sample was analysed using the fluorescent zymogram technique presented in chapter four to see if more enzymes are identified after equalisation.

4.1: Low Abundance Proteins.

As stated in chapter two, the enrichment of low abundance proteins and removal of high abundance proteins is the focus of much research in proteomics. This is mainly due to the simple fact that the abundant proteins are well characterised and proteins that responsible for changes in an organism are often of low abundance (less than 5000 copies per cell, (Ghaemmaghami *et al.*, 2003). The effect of not removing or dealing with high abundance proteins was graphically shown in section 2.3.3 where 1-D SDS-PAGE and 1-DLC/MS was used as an alternative to 2-DGE to analyse sequentially extracted unengorged female *I. holocylcus*. Cytoskeletal proteins were greatly over represented in this analysis.

As the detection limits of protein stains and mass spectrometry get lower and lower, many scientists are finding that in most proteomes, the dynamic concentration range could be as great as 10 orders of magnitude (Boschetti *et al*., 2007). This makes the detection of the "rare proteome", which is believed to comprise at least 50%, by individual protein species, of a given proteome, very difficult as the high abundance proteins that are well characterised mask the existence of less characterised or novel low abundance proteins. The data presented in chapter two and three shows that ticks also have a wide dynamic range of protein concentration. In the case of blood feeding, the protein content from the blood of the host, predominantly haemoglobin, is much greater than the protein belonging to the tick. Thus to study the proteins expressed by the blood feeder during feeding, the host protein must be removed or somehow depleted. Previous efforts in our lab to deplete and remove proteins in the blood meal have only been moderately successful and involved many chromatographic steps resulting in the loss of already low amounts of proteins as the number of purification steps increases (Thurn, 1994).

A number of prefractionation techniques based on affinity, electrophoretic and chromatographic properties have appeared over the last few years to deal with this

problem of high abundance proteins (Righetti *et al*., 2005). Of these, the multicompartment electrolyser (Herbert and Righetti, 2000) was used in chapter two to resolve the proteome of unengorged ticks. Electrophoretic techniques attempt to isolate the high abundance protein to a particular isoelectric fraction. Affinity techniques are rarely global, affecting different subsets of the proteome rather than the whole (Azarkan *et al*., 2007). In addition, both techniques are limited in the amount of sample that can be loaded. As no amplification technique for proteins exists, the isolation of low abundance proteins requires a very large amount of starting material to recover enough of the low abundance protein to be detected and identified using tandem mass spectrometry (Herbert and Harry, 2008).

A recent technical advance to deal with high abundance proteins is Equalizer Technology (figure 4.1). This approach uses a combinatorial hexapeptide ligand library bound to chromatographic beads and aims to "embrace" all proteins present in a sample (Boschetti *et al*, 2007). Each single bead has millions of copies of a single, unique ligand structure and each bead, potentially, has a different ligand from every other bead. Depending on the number of amino acids used, a hexapeptide library can contain millions of different ligands, up to 64 million if 20 amino acids are used. Each individual bead acts as an affinity column for a certain protein. When a complex protein sample is exposed to the library of beads, the beads with affinity for the abundant proteins will rapidly become saturated and the majority of the abundant protein will remain unbound and removed on washing of the beads. Thus the lower abundance proteins are enriched while the abundant protein's concentration is reduced. The equalized sample can then be eluted from the beads by a variety of means including the use of high salt concentrations (1M), chaotropes and high or low pH. The equalized sample can also be subfractionated during elution by using different elution conditions (Boschetti *et al*, 2007). This technology has been recently commercialized by Bio-Rad as Proteominer. The only real limitation of the technology is that sample adsorption must occur under native physiological conditions (Righetti

and Boschetti, 2007) limiting the technology to proteins that are soluble in buffers such as PBS. Experiments being conducted in our lab are trialling different reagents, such as organic solvents and alkylating reagents containing quaternary ammonium, to overcome this limitation.

Equalizer technology has been applied to a number of samples with great success. From previous studies of human serum, HUPO has generated a core set of serum and plasma proteins containing 889 entries and it was thought that this greatly underestimated the number of proteins that should be present. Anderson and Anderson (2002) have written that most if not all human proteins should be found in serum, as well as bacterial and viral proteins. The dynamic range of protein concentration in serum covering 10 orders of magnitude, from 1ng/mL for some interleukins to 30mg/mL for albumin, was one of the main samples in mind when developing equalizer technology. In 2004, Anderson *et al* published a compilation of 1175 non-redundant proteins in human serum collated from several sources. This was superseded by the HUPO Plasma Proteome Project (PPP) using the combined efforts of 35 laboratories to generate a core set of 3020 proteins. In the work of Sennels *et al* (2006), 3661 unique gene products were found using equalizer technology in a single experimental protocol. The technology has also been applied with success to platelet lysate (Guerrier *et al*., 2007), serum of never-smoked lung cancer patients (Au *et al*., 2007), human urine (Castagna *et al*., 2005), *E.coli* (Thulasiraman *et al*., 2005), yeast (Righetti and Boschetti, 2007), cerebrospinal fluid, saliva proteins, and egg white proteins (Boschetti *et al*, 2007).

In this chapter, equalizer technology was applied to both engorged and unengorged *Ixodes holocyclus*, the Australian Paralysis Tick. The previous chapters have shown that, as with most eukaryotic organisms, the dynamic concentration range of proteins is quite large in tick protein samples. This is especially true for engorged ticks where we estimate at least 75% of the protein in the tick is blood meal from its host.

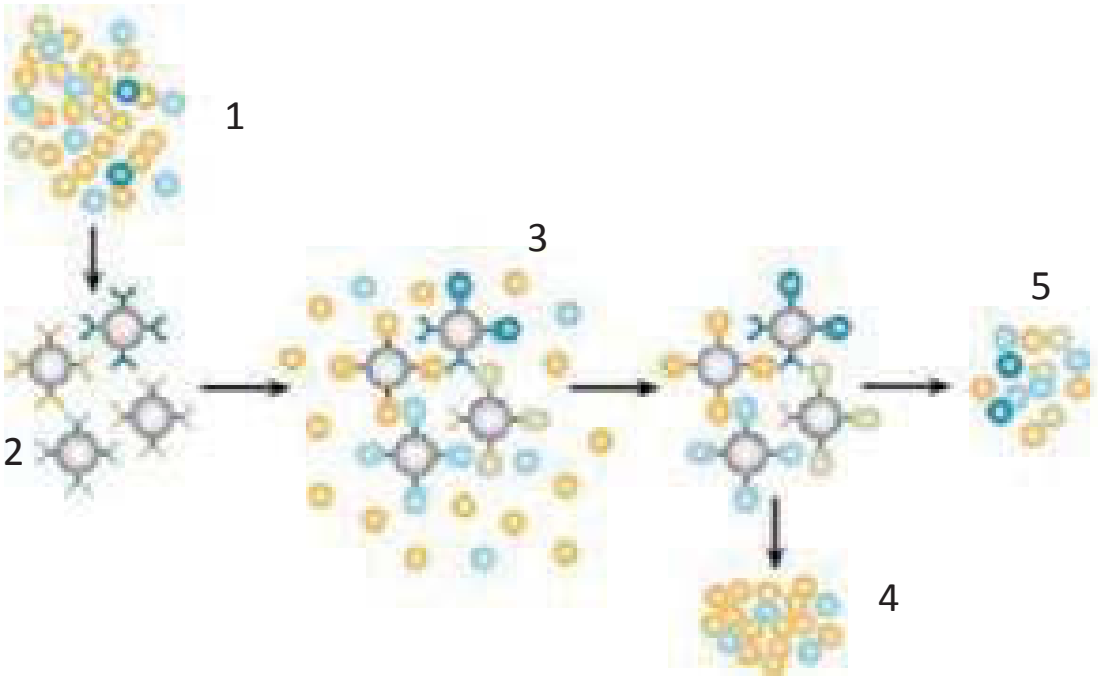

Figure 4.1: Graphical representation of protein equalisation using a combinatorial hexapeptide library (Proteominer). The solubilised protein (1) is mixed with the Proteominer beads (2). To each bead is bound a separate type of hexapeptide. The soluble proteins bind to the appropriate hexapeptide with the high abundance proteins saturating their hexapeptide ligand (3). Unbound protein is then washed away (4) and the equalised protein eluted from the beads for analysis (5).

As shown in chapter two, sequential extraction of ground ticks (40mM Tris-HCl pH 8.8 followed by 7M Urea, 2M Thiourea, 1% C7BzO, 40mM Tris-HCl pH 8.8, both supplemented with 150mM LiCl) and two-dimensional gels has shown hundreds of protein spots are present in a wide concentration range. However, only a small number of these spots, approximately 20%, have been identified by in-gel trypsin digestion and tandem mass spectrometry. This is due to two reasons. Firstly, often the spots are too faint and thus the peptide recovery is too low to identify the protein. Secondly, the available amount of sequence data for ticks and especially *I.holocyclus* is relatively small, and *de novo* peptide sequencing and BLAST homology searching becomes necessary. This approach relies on high quality MS/MS data which is difficult with the low amount of peptides recovered from gel plugs of low intensity spots. Thus the application of equalizer technology to these samples was very attractive as not only should it increase the relative concentration of low abundance tick proteins, but by removing high abundance host protein, the equalization may reveal host proteins produced in response to tick feeding. There is no published literature about using Proteominer with mixed proteomes. Not only are engorged ticks an example of a mixed proteome, they are a mixed proteome where one proteome, the host, is in far greater abundance than the tick proteome. Much of the work in this chapter is about optimising methodologies using Proteominer on mixed proteomes.

4.2: Methods.

4.2.1: Preliminary experiment equalising protein from fully engorged ticks.

Figure 4.2 diagrammatically shows the sequence of experimental methods used in the work presented in this chapter. Fully engorged female ticks (fed on dogs and native animals and collected by veterinarians and stored at -20^{O}C until use) were snap frozen in liquid nitrogen and ground to a fine powder in a mortar and pestle as described in the chapter two. The powder was transferred to 50mL tubes and 5mL PBS added. Following 3 x 30 second pulses with an ultrasonic probe (50% power), samples were centrifuged to remove insoluble material and the supernatant removed. The pellet was re-extracted with 5mL PBS and added to the first extract. The contents of a Proteominer column was added to the supernatant in a 50mL tube and incubated at room temperature on a rotating wheel for 2 hours. The supernatant and Proteominer beads were then passed back through the empty Proteominer column, retaining the beads and allowing the unbound material to flow through. The beads were then washed with PBS until the beads were as white as possible. The bound equalised protein was then sequentially eluted with UTC7Tris-HCl pH 8.8 and UTC7Citric acid pH 4.5.

4.2.2: Unengorged tick protein equalisation.

100 unengorged female ticks stored at -80C were ground to powder with liquid nitrogen in mortar and pestle. The powder was transferred to a 50mL tube, suspended in 5mL of PBS and ultrasonicated 3 x 30 seconds with an ultrasonic probe at 50% power. The sample was then centrifuged at 3000g to pellet insoluble material and supernatant removed. 3mL of PBS was added to the pellet and ultrasonicated with an ultrasonic probe to aid resuspension. The sample was then centrifuged and the extraction repeated once more. The third extraction was clear and it was assumed that this extract did not contain significant additional protein so it was not added to the other extracts. The remaining insoluble pellet was extracted three times with UTC7Tris-HCl pH 8.8 (6mL total) and frozen for use in other experiments as being in

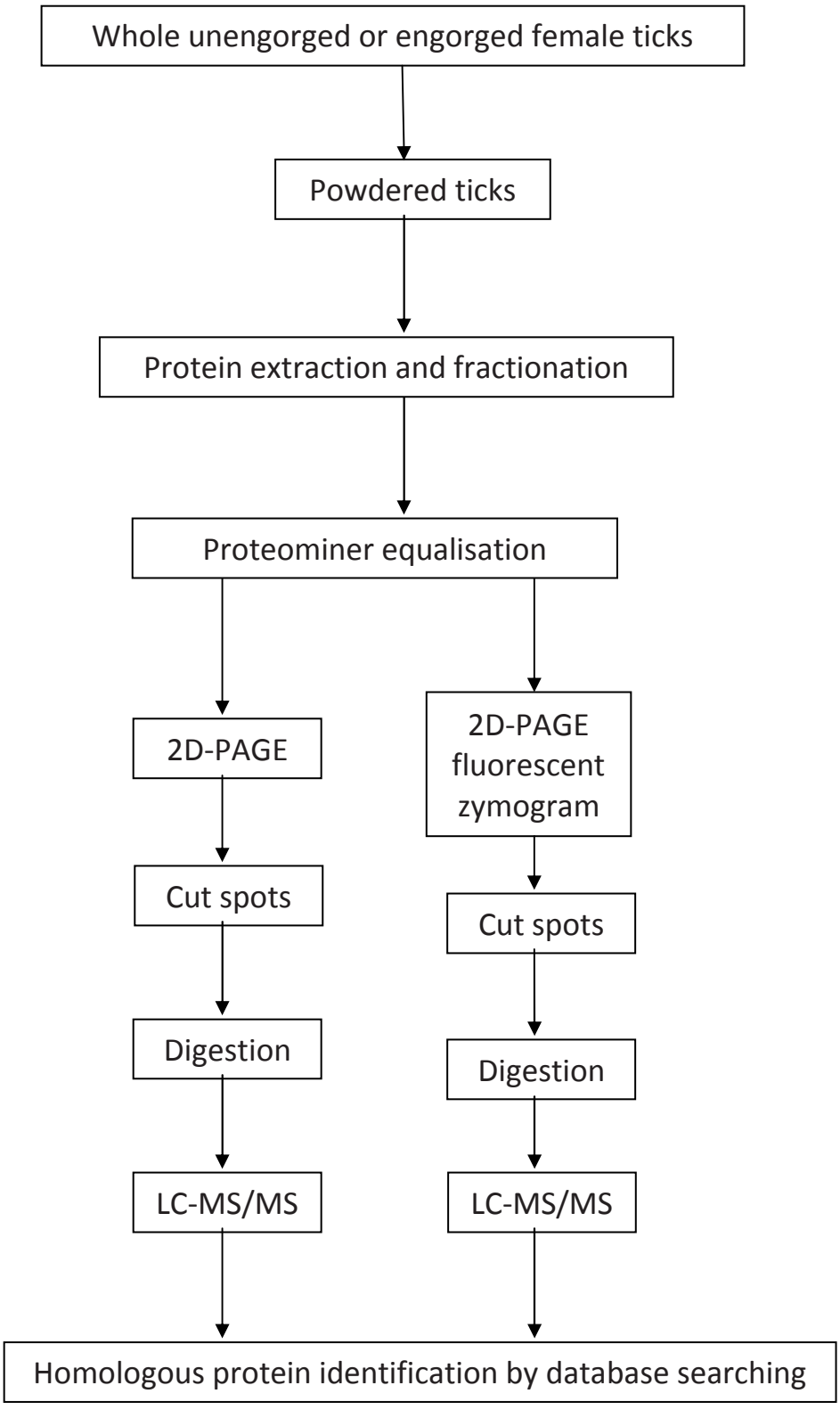

Figure 4.2: Diagrammatic outline of the experimental steps carried out in this chapter. Exact methods are described in detail in the body text of the chapter.

UTC7 makes it unsuitable for equalisation experiments with Proteominer. One vial of Proteominer beads was prepared according to the manufacturer's instructions. Once prepared, the beads were removed from the spin column provided and added to sample (approximately 7.5mL volume) in a 15mL tube. The combined sample and beads were then incubated at room temperature on rotating wheel for two hours. The sample and beads were then loaded back into the spin column and washed three times with 1mL of wash buffer (provided in Proteominer kit). The bound, equalised protein was then eluted from the beads with 2 x 500µl of unbuffered UTC7 followed by 500µL of unbuffered 6M Guanidine HCl. The pH of the UTC7 elutions was adjusted with 1.5M Tris-HCl pH 8.8, to a final concentration of 40mM before the proteins were reduced and alkylated as described in section 2.2.1. The guanidine-HCl elution was adjusted to 100mM NH_4HCO_3 with 1M NH_4HCO_3 before reduction and alkylation as described in chapter two. All samples were desalted into UTC7 with BioSpins as described in section 2.2.1 before isoelectric focusing.

4.2.3: Partially engorged tick protein equalisation.

Forty-five partially engorged female ticks were snap frozen in liquid nitrogen and then ground to a powder with a mortar and pestle. The powder was resuspended in 10mL of PBS and sonicated with an ultrasonic probe at 50% power for 3 x 30 seconds. The sample was then centrifuged at 3000g for 10 minutes, the supernatant removed, another 10mL of PBS added and the extraction procedure repeated a total of three times. The three extracts were combined and a 1mL aliquot removed (non equalised sample). The remaining sample was added to the removed beads of one Proteominer spin column in a 50mL tube. The sample and the beads were incubated on a rotating wheel for three hours at room temperature. The beads were centrifuged to the bottom of the 50mL tube at 1000g for 2 minutes and the unbound material removed to a new 50mL tube. The beads were resuspended in 1mL of PBS and returned to the original spin column and washed twice with 1mL of PBS, wash solutions being removed into a collection tube by centrifuging the spin column at 1000g for 2 minutes. Bound, equalised protein was eluted with 2 x 1mL of unbuffered

UTC7 followed by 2 x 500µL of unbuffered 6M guanidine-HCl. Any remaining protein was eluted by adding 500µL of 2 x SDS sample buffer. The pH of the UTC7 and guanidine elutions was adjusted to 8.8 by adding 1.5M Tris-HCl to a final concentration of 50mM before reduction and alkylation as described in section 2.2.1. Samples were then acetone precipitated and precipitation was assisted by the addition of citric acid until the precipitate was visible. Samples were resuspended in UTC7 and a 7.5µL aliquot loaded onto SDS-PAGE as described in section 2.2.2. This experiment was repeated with two columns of beads to produce an equalised sample that could be examined with the fluorescent zymogram method described in chapter four. However the sample was not reduced and alkylated so that the proteins could be renatured in the gel and retain their enzymatic activity. All samples were desalted into UTC7 with BioSpins as described in section 2.2.1 before isoelectric focusing.

4.2.4: Fully engorged tick protein equalisation.

The equalisation experiment with PBS soluble protein from engorged female ticks was repeated to produce an equalised sample that could be examined with the fluorescent zymogram method described in section 3.2.2. Thirty-six fully engorged female ticks were snap frozen in liquid nitrogen and then ground to a powder with a mortar and pestle. The powder was halved and resuspended in PBS (20mL each) in centrifuge tubes and sonicated with an ultrasonic probe for 6 x 30 seconds on ice. The sample was centrifuged at 38000g for 10 mins and the supernatant removed. The pellet was resuspended in PBS (15mL each), sonicated and centrifuged again. The supernatant was removed and the pellet resuspended in PBS (10mL each), sonicated and centrifuged. An aliquot (1mL) of each PBS extract was removed and stored at -20^{O}C until analysis. The remaining pellet was then resuspended in UTC7 + 50mM Tris-HCl pH 8.8 + 50mM LiCl (10mL) and sonicated with the probe. The sample was centrifuged at 38000g for 10 mins and the supernatant removed before adding more UTC7 + 50mM Tris-HCl pH 8.8 + 50mM LiCl and repeating the extraction process. A total of 3 x 10mL extractions were performed and the extract frozen at -20^{O}C. An aliquot (1mL) of each extract was removed and stored at -20^{O}C until analysis. All PBS

extracts were then combined. Due to the large sample volume, it was decided to perform the equalisation in a column format. The sample was recirculated with a peristaltic pump over the combined beads of two columns at 100µL/min overnight. The beads were washed with 400mL PBS at 250µL/min for a further 24 hours, all at 4^{O}C. Periodically the beads had to be stirred up in the column to allow efficient flow of both sample and PBS wash buffer. Following washing, the beads were removed from the column and returned to their original spin column. The beads were further washed with 6 x 1mL of PBS until the PBS wash solution was clear indicating no further non specifically bound protein was being washed off. Protein bound to the beads was eluted with 10 x 500µL of unbuffered UTC7 and then 5 x 500µL of 6M guanidine-HCl. The large number of elutions were performed to ensure all protein was removed by each elution buffer before moving on to the next one. All collected fractions and the beads were frozen at -20^{O}C until analysis. The efficiency of the sequential extraction and equalisation was examined by separating 7.5µL of the samples and fractions by SDS-PAGE as described in section 2.2.2. All samples were desalted into UTC7 with BioSpins as described in section 2.2.1 before isoelectric focusing. Fluorescent zymograms were performed as described in section 3.2.2.

4.2.5: Fully engorged tick protein equalisation after reduction and alkylation.

Thirty-five fully engorged female *I. holocyclus* stored at -20^{O}C were snap frozen in liquid nitrogen and then ground to a powder with a mortar and pestle. The powder was then transferred to a 50mL tube and 50mL of ice-cold 10% trichloroacetic acid (TCA) in acetone added to precipitate as much as the protein as possible. The sample was left to precipitate at -20^{O}C for 30 minutes before being centrifuged at 1500g for 10 minutes. The TCA/acetone was poured off and the pellet resuspended, washed and centrifuged in 50mL of ice-cold acetone twice to remove any remaining TCA. The pellet was then solubilised in 7M urea, 2M thiourea, 1% C7BzO (UTC7), 50mM Tris-HCl pH 8.8 + 150mM LiCl before the proteins were reduced and alkylated as described in section 2.2.1. To ensure complete reduction and alkylation of all haemoglobin multimers, a second amount of reagents was added after 45 minutes. The reagents were then

removed by precipitating the protein with five volumes of acetone at room temperature for 30 minutes. An attempt was made to resuspend the precipitated protein with 30% trifluoroethanol (TFE) in PBS, but only a small proportion of the protein would solubilise in this solution.

To improve protein solubility, the sample preparation was repeated but the proteins were alkylated with (3-acrylamidopropyl)-trimethylammonium chloride, an alkylating reagent containing a quaternary ammonia group, at a final concentration of 20mM. This would provide an additional charged group to the proteins and was theorised to improve solubility in aqueous solutions such as PBS. Three lots of twelve engorged female *I. holocyclus* were frozen and ground as previously described. The powder in three separate 50mL tubes was resuspended in 10% TCA/acetone, sonicated in a waterbath for 10 minutes and left at -20^{O}C for 30 minutes for the proteins to precipitate. The sample was centrifuged at 1500g for 5 mins and the acetone discarded. The entire procedure was repeated at least ten times until a sufficient layer of white protein was present on top of a dark red pellet. This white protein was removed with a spatula and washed with ice-cold acetone to remove the TCA. The powder was allowed to dry before being halved. To one half was added 30% TFE in PBS which was then alternately heated to 50^{O}C for 30 minutes and then ultrasonicated in a waterbath. The sample was then centrifuged to remove insoluble proteins. Other work in our laboratory has shown that heating the sample in 30% TFE in PBS to 50^{O}C increases the amount of protein solubilised compared to solubilising at room temperature (Sivell, 2008). The second half of the sample was resuspended in UTC7Tris-HCl pH 8.8 + 150mM LiCl before being reduced and alkylated with 5mM TBP and 20mM (3-acrylamidopropyl)-trimethylammonium chloride for 90 minutes. Reagents were removed by acetone precipitation and the precipitated protein solubilised in 30% TFE in PBS as described above.

Solubilised protein was then incubated for two hours on a rotating wheel with a column of Proteominer beads prepared as per manufacturer's instructions. After washing away unbound protein with five washes of 30% TFE in PBS, the equalised protein was eluted by boiling the beads in 2x SDS sample buffer. This was done because by alkylating the proteins with a reagent containing a quaternary ammonia functional group the isoelectric point of the protein is altered. The alkylated proteins would then focus to a more alkaline point in the IPG strip and there is a risk that some would focus to the very end of the strip and not be resolved. Thus, the equalised protein was loaded onto SDS-PAGE for analysis using a 1-D SDS-PAGE/1-DLC/MS experiment as described in chapter two, section 2.3.3.

To disrupt haemoglobin structure, liquid nitrogen ground engorged ticks (50) were resuspended in a buffer containing ammonium chloride, which has been shown to burst red blood cells (Maren and Wiley, 1970). Following disruption with an ultrasonic probe in this buffer (50mL per 10 ticks), the insoluble material was pelleted by centrifugation at 38 000g for 10 minutes and the protein in the supernatant precipitated with ice-cold 10% TCA in acetone. The pellet would not solubilise in PBS and thus TFE was added to a concentration of 30% and the solution alternately ultrasonicated and heated at 50^{o}C to solubilise the protein. Approximately 250mL of supernatant was combined in a device constructed from five 50mL tubes. One column of Proteominer beads were prepared as per manufacturer's instructions and added to the supernatant, which was then incubated overnight at room temperature on a rotating wheel. The beads were then recovered and returned to their original spin column for washing (5 x 1mL of 30% TFE in PBS). The equalised protein was eluted from the beads by boiling them for 10 minutes in SDS sample buffer and loaded onto SDS-PAGE for analysis using a 1-D SDS-PAGE/1-DLC/MS experiment as described in chapter two, section 2.3.3.

4.2.6: Gel electrophoresis and mass spectrometry.

For all Proteominer experiments except the quaternary ammonium alkylation, the PBS extracted unequalised sample, the equalised sample and the unbound sample were analysed by 2-DGE using pH 3-10 11cm IPG's as described in section 2.2.3. All gels, including 1-D-SDS-PAGE of the quaternary ammonium alkylated sample, were stained and imaged as described in section 2.2.4. Selected spots from 2-DGE and entire 1-D SDS-PAGE lanes were excised, trypsin digested and identified by mass spectrometry and database searching as described in sections 2.2.5, 2.2.6 and 2.2.7.

4.3: Results and Discussion.

4.3.1: Preliminary experiment.

Figure 4.3 shows gel images of PBS extracted whole engorged ticks that have been equalized and the beads sequentially eluted with UTC7Tris (A) and then UTC7Citric acid (B). The multiple ascending "trains" of spots in Gel A show that the reduction and alkylation of the sample has not been complete as these "trains" are all isoforms of canine haemoglobin multimers. This result is not unusual when dealing with high abundance serum proteins. What is apparent when comparing the two gels is the strikingly different patterns of protein spots. Dog haemoglobin is clearly visible at the alkaline end of both gels, but the citric acid has eluted a number of more alkaline proteins focused right at the end of the IPG. All tick proteins that were identified in this sample were previously identified in unengorged ticks in chapter two, but many other spots were unable to be identified with the database search methods used.

4.3.2: Unengorged tick protein equalisation.

Figure 4.4 shows an SDS-PAGE of the fractions obtained from the Proteominer beads when using them to equalise unengorged female *I. holocylcus*. The first comparison of interest is that the initial PBS extracted proteins and the protein not bound by the beads has a protein pattern that is almost identical in both bands present and band intensity. This is to be expected as this technique is not performing a complete depletion, but rather binding a small amount of the high abundance proteins and concentrating the low abundance proteins, which are not able to be seen in the original extract. Lanes 4-6 show the unbound protein being washed away and it is evident that additional washes could have been performed. Once washed, the beads were then sequentially eluted with increasingly chaotropic and ionic buffers and finally, the beads boiled in SDS sample buffer. Lanes 7 and 9 show the protein that is eluted from the beads with UTC7 while lanes 10 and 11 show the proteins eluted with 6M guanidine-

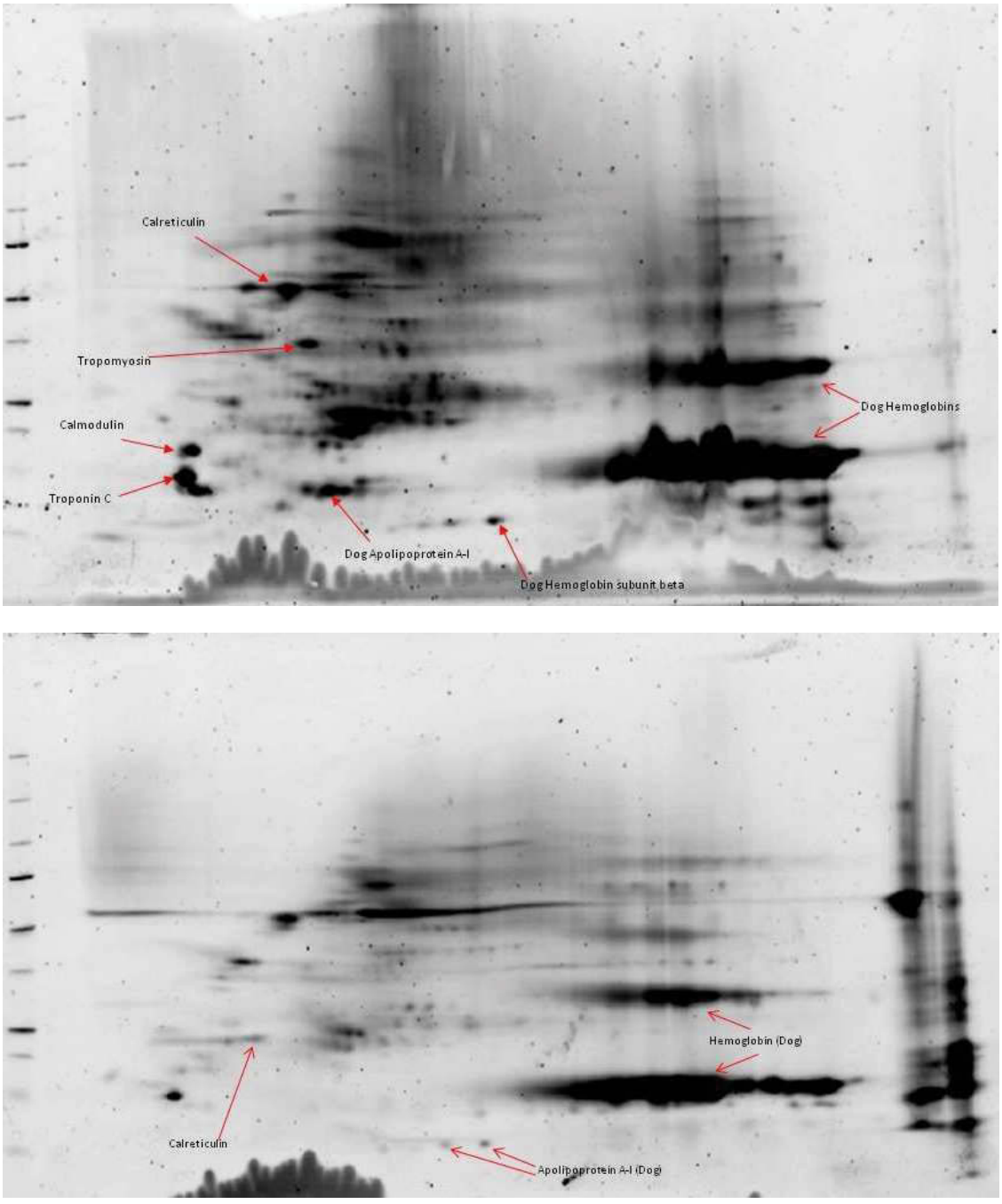

Figure 4.3: Preliminary experiment. PBS extracted protein of engorged female *I.holocyclus* subjected to equalisation using a Proteominer hexapeptide library and then eluted with UTC7Tris-HCl pH 8.8 (Gel A) and UTC7Citric Acid pH 4.5. Labelled proteins are those identified using Mascot searching of NCBI non-redundant database.

UTC7Tris-HCl pH8.8 elution of beads						
Spot #	Protein accession number/Protein identification/Species	Score	Matches	Coverage (%)	pI	MW
1	sp|P15159|Troponin C.[*Tachypleus tridentotus*]	194	12	11	4.11	17422
4	sp|P02595|Calmodulin (CaM).[*Patinapecten sp.*]	315	5	29	4.04	16802
5	sp|P02648|Apolipoprotein A-I precursor (Apo-AI) (ApoA-I).[*Canis familiaris*]	83	2	7	5.2	30178
5	apolipoprotein A-I precursor - ens|ENSCAFP00000019630 [*Canis familiaris*]	424	10	28	5.28	30163
7	sp|P02648|Apolipoprotein A-I precursor (Apo-AI) (ApoA-I).[*Canis familiaris*]	497	10	32	5.2	30178
8	sp|P02648|Apolipoprotein A-I precursor (Apo-AI) (ApoA-I).[*Canis familiaris*]	154	3	12	5.2	30178
13	tr|Q64K84|Calreticulin.[*Ixodes jellisoni*]	343	6	15	4.57	48038
14	Tropomyosin - tr|A5Z1D9|Putative uncharacterized protein.[*Haemaphysalis qinghaiensis*]	384	6	18	4.7	32893
15	apolipoprotein A-I precursor - ens|ENSCAFP00000019630 [*Canis familiaris*]	761	24	45	5.28	30163
16	apolipoprotein A-I precursor - ens|ENSCAFP00000019630 [*Canis familiaris*]	671	15	43	5.28	30163
17	apolipoprotein A-I precursor - ens|ENSCAFP00000019630 [*Canis familiaris*]	580	12	33	5.28	30163
19	sp|P60524|Hemoglobin subunit beta (Hemoglobin beta chain) (Beta-globin).[*Canis familiaris*]	148	3	18	6.69	19093
21	apolipoprotein A-I precursor - ens|ENSCAFP00000019630 [*Canis familiaris*]	1250	33	64	5.28	30163
22	apolipoprotein A-I precursor - ens|ENSCAFP00000019630 [*Canis familiaris*]	1305	60	64	5.28	30163
23	apolipoprotein A-I precursor - ens|ENSCAFP00000019630 [*Canis familiaris*]	1403	60	68	5.28	30163
24	apolipoprotein A-I precursor - ens|ENSCAFP00000019630 [*Canis familiaris*]	1161	31	62	5.28	30163
25	apolipoprotein A-I precursor - ens|ENSCAFP00000019630 [*Canis familiaris*]	360	6	18	5.28	30163
26	Hemoglobin subunit beta - ens|ENSCAFP00000021230 [*Canis familiaris*]	212	4	29	7.83	16259
27	apolipoprotein A-I precursor - ens|ENSCAFP00000019630 [*Canis familiaris*]	246	4	18	5.28	30163
28	Hemoglobin subunit beta - ens|ENSCAFP00000021230 [*Canis familiaris*]	363	6	49	7.83	16259
31	apolipoprotein A-I precursor - ens|ENSCAFP00000019630 [*Canis familiaris*]	632	12	41	5.28	30163
AT1	HBDG - hemoglobin beta chain - *Canis familiaris*	158	2	17	7.95	16128
AT3	MMACTBR2 NID: - *Mus musculus* (b-actin)	116	3	10	5.78	39516
AT4	LPDGA1 - apolipoprotein A-I precursor - *Canis familiaris*	672	12	42	5.2	30178
	ens transcript:ENSCAFT00000021138 [*Canis familiaris*]	757	14	46	5.28	30163
AT5	LPDGA1 - apolipoprotein A-I precursor - *Canis familiaris*	218	4	15	5.2	30178
	ens transcript:ENSCAFT00000021138 [*Canis familiaris*]	271	5	18	5.28	30163
AT8	LPDGA1 - apolipoprotein A-I precursor - *Canis familiaris*	352	7	28	5.2	30178
	ens transcript:ENSCAFT00000021138 [*Canis familiaris*]	402	8	31	5.28	30163
AT9	HBDG - hemoglobin beta chain - *Canis familiaris*	228	4	30	7.95	16128

AT10	HBDG - hemoglobin beta chain - *Canis familiaris*	156	3	23	7.96	15128
AT11	Q64K84_9ACAR - Calreticulin - *Ixodes jellisoni*	421	8	16	4.57	48038
AT12	LPDGA1 - apolipoprotein A-I precursor - *Canis familiaris*	131	2	8	5.2	30178
	ens transcript:ENSCAFT00000021138 [*Canis familiaris*]	178	4	12	5.28	30163
BT1	LPDGA1 - apolipoprotein A-I precursor - *Canis familiaris*	593	9	34	5.2	30178
	ens transcript:ENSCAFT00000021138 [*Canis familiaris*]	651	10	37	5.28	30163
BT2	Q45YD8_9BILA - ATP synthase beta subunit (Fragment) - *Amphiporus angulatus*	683	12	31	5.22	45116
	ens transcript:ENSCAFT00000021138 [*Canis familiaris*]	707	12	42	5.28	30163
BT3	SJHUA - spectrin alpha chain - human	91	2	0	4.98	282305
	ens transcript:ENSCAFT00000018512 [*Canis familiaris*]	92	2	0	5.04	288238
BT4	A5LHV9 - Protein disulfide isomerase-2.[*Haemaphysalis longicornis*]	61	1	2	4.84	57368
	GB93P56UKD16 - weakly similar to UP PDIA1_RABIT Protein disulfide-isomerase precursor	91%	2	7		72772
BT5	HBDG - hemoglobin beta chain - *Canis familiaris*	163	2	17	7.96	15128
BT6	HBDG - hemoglobin beta chain - *Canis familiaris*	200	3	23	7.96	15128
BT7	HBDG - hemoglobin beta chain - *Canis familiaris*	190	3	23	7.96	15128
BT8	AB090854 NID: - *Canis familiaris* (serum albumin)	871	13	16	5.52	71046
BT9	AB090854 NID: - *Canis familiaris* (serum albumin)	759	15	22	5.52	71046
BT10	AB090854 NID: - *Canis familiaris* (serum albumin)	933	16	24	5.52	71046
BT11	AB091831 NID: - *Camelus dromedarius* (mRNA for immunoglobulin heavy chain mu)	100	2	3	5.48	60638
	AVDGGM - Ig heavy chain V region (Gom) - *Canis familiaris* (tentative sequence)	87	1	16	5.24	12564
	ens transcript:ENSCAFT00000029302 [*Canis familiaris*]	340	5	7	5.89	68791
BT12	HBDG - hemoglobin beta chain - *Canis familiaris*	178	2	17	7.96	15128
	ens transcript:ENSCAFT00000029302 [*Canis familiaris*]	321	7	8	5.89	68791
CT1	HBDG - hemoglobin beta chain - *Canis familiaris*	171	2	17	7.96	15128
CT2	HBDG - hemoglobin beta chain - *Canis familiaris*	199	3	23	7.96	15128
CT3	AB090854 NID: - *Canis familiaris* (serum albumin)	390	5	11	5.52	71046
CT4	AB090854 NID: - *Canis familiaris* (serum albumin)	395	5	8	5.52	71046
CT5	HBDG - hemoglobin beta chain - *Canis familiaris*	129	2	15	7.96	15128

UTC7 Citric Acid elution pH 4.5						
Spot #	Protein accession number/Protein identification/Species	Score	Matches	Coverage (%)	pI	MW
AC8	LPDGA1 - apolipoprotein A-I precursor - *Canis familiaris*	517	13	44	5.2	30178

UTC7 Citric Acid elution pH 4.5						
Spot #	Protein accession number/Protein identification/Species	Score	Matches	Coverage (%)	pI	MW
AC8	LPDGA1 - apolipoprotein A-I precursor - dog	517	13	44	5.2	30178
AC9	LPDGA1 - apolipoprotein A-I precursor - dog	481	9	39	5.2	30178
AC10	LPDGA1 - apolipoprotein A-I precursor - dog	517	10	43	5.2	30178
AC12	tr|Q64K84|Calreticulin.[Ixodes jellisoni]	105	2	7	4.57	48038
DC5	P60524 sp|P60524|Hemoglobin subunit beta [Canis familiaris]	450	8	63	7.83	16259

Table 4.1: Identifications of proteins from fully engorged female Ixodes holocyclus subjected to equalisation using a Proteominer hexapeptide library, eluted sequentially with UTC7Tris-HCl pH 8.8 and UTC7Citric Acid pH 4.5, and separated by isoelectric focusing and SDS-PAGE (Preliminary experiment). Protein spots were excised from the gel, reduced and alkylated, then digested overnight with trypsin. The resulting peptides were then separated and analysed by LC/MS/MS as described in chapter three. Data was then searched against MSDB and NCBI non redundant databases using the Mascot search engine. Gel images showing the location of all spots are presented in appendix.

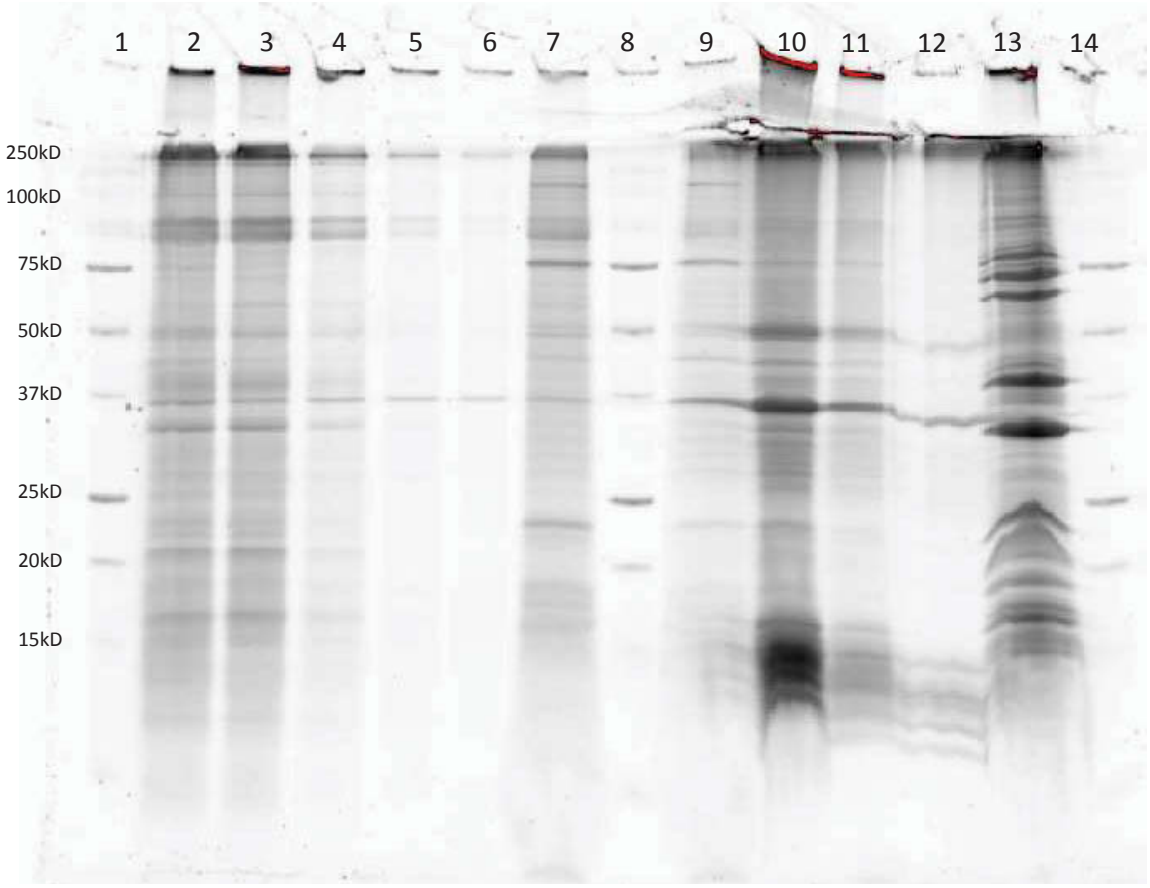

Figure 4.4: SDS-PAGE of PBS extracted proteins from whole unengorged female I.holocyclus that has been subjected to Proteominer equalisation. 10µL of each sample was loaded onto the gel.

1) Molecular weight markers.
2) PBS extracted protein.
3) Unbound protein after two hour incubation with beads.
4) PBS Wash 1.
5) PBS Wash 2.
6) PBS Wash 3.
7) Elution 1 (UTC7).
8) Molecular weight markers.
9) Elution 2 (UTC7).
10) Elution 3 (GuHCl).
11) Elution 4 (GuHCl).
12) Elution 5 (SDS).
13) Insoluble tick protein remaining after PBS extraction subsequently extracted with UTC7Tris-HCl pH 8.8.
14) Molecular weight markers.

HCl. It was expected, from the literature (Boschetti *et al.*, 2007) that the guanidine would strip all remaining protein from the beads. Fraction 10, representing the protein eluted when the beads are boiled in SDS sample buffer, shows that this is not the case. What is evident is that the protein bands are of a more uniform intensity and some protein bands not seen in the original PBS extract are visible in the elutions. This would imply that equalisation has occurred.

The gels in figure 4.5 are two-dimensional gels (isoelectric focusing and SDS-PAGE) showing PBS extracted whole unengorged ticks prior to equalization (A), following equalization (B) and the protein not bound by the beads during equalization and subsequently washed away before eluting (C). All three gels show similar spot patterns and spot intensity for abundant proteins such as tropomyosin and actin isoforms of approximately 40kDa. Using Proteomeweaver to determine spot numbers in the gels, gel A contains approximately 370 spots while gel B contains approximately 410 spots, an increase of 40. This increase in spot numbers is not as extreme as that seen in equalised human plasma (Sennels *et al.*, 2007) and red blood cell lysate (Simo *et al.*, 2008) but this is likely due to the protein concentration and amount being too low to enrich the low abundance proteins above the fluorescent stain detection limit. This experiment could not be repeated due a lack of more unengorged ticks to prepare sample.

Comparing gels A and C shows that the abundant proteins are relatively unchanged in their intensity, but a number of spots are absent or of lower intensity in gel C. Comparison of gels A and C with gel B show a group of spots at the acidic end of the gels whose intensity has increased while the intensity of some spots at the alkaline end has reduced. This would imply that the equalization is not equal. This observation has been made when applying Proteominer to other samples, including plasma, in our laboratory (data not shown). As the sample preparation has occurred without any chaotropes or denaturants, protein-protein interactions normally present are not disrupted.

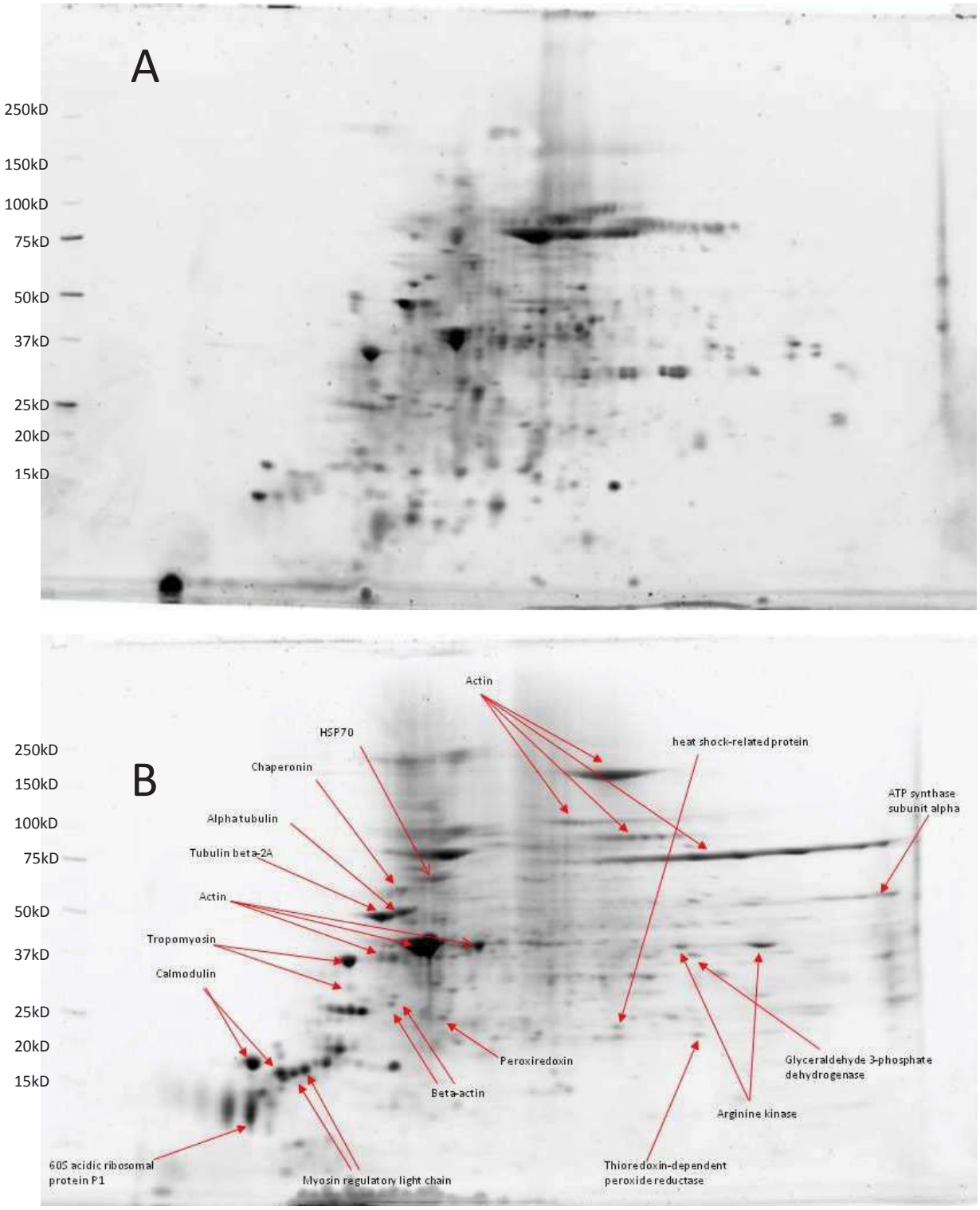

Figure 4.5: Protein isolated from whole, unengorged female *I. holocyclus* with PBS then subjected to equalisation using a Proteominer hexapeptide library. Gel A is loaded with the PBS extracted protein that has not been equalised. Gel B is loaded with protein eluted from the Proteominer beads by UTC7 or the equalised protein.

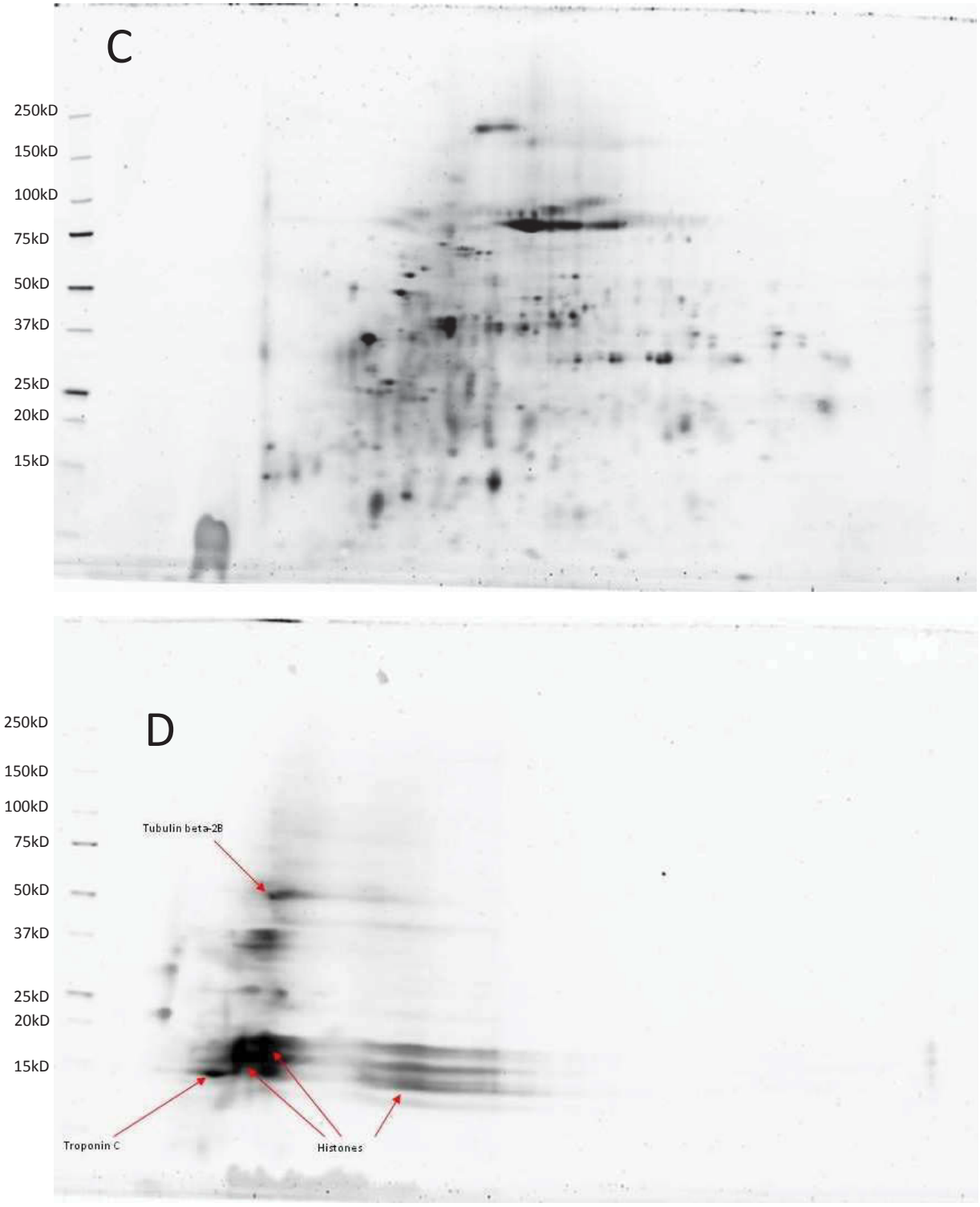

Figure 4.5: Protein isolated from whole, unengorged female *I. holocyclus* with PBS then subjected to equalisation using a Proteominer hexapeptide library. Gel C is loaded with protein that is not bound by the Proteominer beads and washed away before elution with UTC7. Gel D is loaded with protein eluted from the Proteominer beads by 6M Guanidine-HCl, which releases equalised protein not eluted by UTC7.

Spot #	Accession number/Identification/Species	Score	Matches	Coverage(%)	Database	pI	MW
2	tr|Q0PXZ8|Putative 60S acidic ribosomal protein P1.[*Diaphorina citri*]	64	1	14	NCBInr	4.24	11451
3	tr|Q1HQX3|Calmodulin.[*Aedes aegypti*]	523	9	54	NCBInr	4.09	16800
5	tr|Q32VZ5|Calmodulin (Fragment).[*Eucheilota bakeri*]	72	2	12	NCBInr	4.11	14855
6	tr|Q6B889|Myosin regulatory light chain.[*Ixodes pacificus*]	325	4	30	NCBInr	4.71	20002
7	tr|Q6B889|Myosin regulatory light chain [*Ixodes pacificus*]	150	2	13	NCBInr	4.71	20002
17	sp|Q8IT89|Tropomyosin.[*Haemaphysalis longicornis*]	807	30	36	NCBInr	4.69	32853
18	tr|Q7SZL6|Beta-actin.[*Monopterus albus*]	343	9	20	NCBInr	5.31	41768
	tr|Q6QNV2|Pupal-specific flight muscle actin (Actin).[*Aedes aegypti*]	340	8	27	NCBInr	5.29	41556
19	tr|Q7SZL6|Beta-actin.[*Monopterus albus*]	250	5	15	NCBInr	5.31	41768
20	tr|Q002L6|Actin (Fragment).[*Pterosperma cristatum*]	177	3	13	NCBInr	5.37	38313
21	tr|Q17C86|Actin.[*Aedes aegypti*]	178	3	11	NCBInr	5.3	41659
22	sp|Q4R5B3|Tubulin beta-2A chain.[*Macaca fascicularis*]	1170	31	53	NCBInr	4.78	49875
23	tr|A7RUT1|Predicted protein.[*Nematostella vectensis*]	783	17	40	NCBInr	4.91	50171
24	tr|Q7Q270|AGAP004002-PA.[*Anopheles gambiae str. PEST*]	542	9	16	NCBInr	5.55	60740
25	tr|Q86QM8|Hsp70 family member (Fragment).[*Locusta migratoria*]	596	11	14	NCBInr	5.39	71364
26	tr|Q5CAR2|Actin.[*Ixodes ricinus*]	121	3	10	NCBInr	5.3	41797
27	tr|Q4PKE5|Actin 5.[*Aedes aegypti*]	1179	73	63	NCBInr	5.3	41795
	tr|Q3LGW2|Actin.[*Ornithodoros moubata*]	1142	67	59	NCBInr	5.3	41811
28	tr|Q4PKE5|Actin 5.[*Aedes aegypti*]	1429	70	68	NCBInr	5.3	41795
	tr|Q6X4W2|Actin.[*Rhipicephalus appendiculatus*]	1387	67	64	NCBInr	5.3	41781
29	tr|Q4PKE5|Actin 5.[*Aedes aegypti*]	1194	69	64	NCBInr	5.3	41795
	tr|Q3LGW2|Actin.[*Ornithodoros moubata*]	1157	65	60	NCBInr	5.3	41811
30	tr|Q4PKE5|Actin 5.[*Aedes aegypti*]	800	17	52	NCBInr	5.3	41795
	tr|Q3LGW2|Actin.[*Ornithodoros moubata*]	757	16	47	NCBInr	5.3	41811
31	tr|Q4PKE5|Actin 5.[*Aedes aegypti*]	693	13	46	NCBInr	5.3	41795
32	tr|Q5CAR2|Actin.[*Ixodes ricinus*]	81	1	6	NCBInr	5.3	41797
33	tr|Q5CAR2|Actin.[*Ixodes ricinus*]	99	1	6	NCBInr	5.3	41797
34	tr|Q5CAR2|Actin.[*Ixodes ricinus*]	108	1	6	NCBInr	5.3	41797
35	tr|Q5CAR2|Actin.[*Ixodes ricinus*]	78	1	6	NCBInr	5.3	41797
36	tr|Q2M0D6|Catalase (EC 1.11.1.6) (Fragment).[*Drosophila pseudoobscura*]	84	1	3	NCBInr	8.41	57137
37	tr|Q6PFI4|Putative arginine kinase.[*Hamalodisca coagulata*]	279	4	13	NCBInr	5.8	39956

38	tr|Q6QWP0|Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (Fragment).[*Phormictopus sp.*]	111	2	7	NCBInr	6.42	31683
39	tr|Q6PP74|Putative arginine kinase.[*Homalodisca coagulata*]	344	7	17	NCBInr	5.8	39956
40	tr|Q2F5T3|ATP synthase subunit alpha (EC 3.6.3.14).[*Bombyx mori*]	386	8	14	NCBInr	9.08	59862
43	tr|Q5CAR2|Actin.[*Ixodes ricinus*]	99	1	6	NCBInr	5.3	41797
44	tr|Q5CAR2|Actin.[*Ixodes ricinus*]	76	1	6	NCBInr	5.3	41797
45	tr|Q5CAR2|Actin.[*Ixodes ricinus*]	95	1	6	NCBInr	5.3	41797
46	tr|Q5CAR2|Actin.[*Ixodes ricinus*]	90	1	6	NCBInr	5.3	41797
47	tr|Q5CAR2|Actin.[*Ixodes ricinus*]	85	1	6	NCBInr	5.3	41797
48	tr|A3DM56|Putative uncharacterized protein.[*Staphylothermus marinus*]	65	1	2	NCBInr	9.53	38362
49	sp|Q8IT89|Tropomyosin.[*Haemaphysalis longicornis*]	264	4	16	NCBInr	4.69	32893
50	tr|Q6QD58|Actin.[*Ornithodoros moubata*]	346	5	26	NCBInr	5.29	41767
51	tr|Q1PC37|Actin (Fragment).[*Teleopsis thaii*]	174	3	14	NCBInr	4.9	26923
52	tr|A6NA14|Truncated peroxiredoxin (Fragment).[*Ornithodoros parkeri*]	174	3	17	NCBInr	5.8	20956
54	tr|Q17C86|Actin.[*Aedes aegypti*]	75	2	7	NCBInr	5.3	41659
55	tr|Q17C86|Actin.[*Aedes aegypti*]	156	3	11	NCBInr	5.3	41659
59	tr|Q4PM52|Putative heat shock-related protein.[*Ixodes scapularis*]	60	1	5	NCBInr	7.07	19265
62	tr|Q6R127|Beta-actin (Fragment).[*Rhipicephalus (Boophilus) microplus*]	117	2	24	NCBInr	5.78	12302
63	tr|Q0Q0G9|Beta-actin (Fragment).[*Artemia sanfranciscana*]	127	2	19	NCBInr	6.14	15281
64	tr|Q4PM65|40S ribosomal protein S12.[*Ixodes scapularis*]	250	4	43	NCBInr	6.51	14375
65	tr|Q5T5U2|Rho GTPase activating protein 21.[*Homo sapiens*]	73	2	4	NCBInr	7.63	21273
69	tr|Q4PN07|Thioredoxin-dependent peroxide reductase.[*Ixodes scapularis*]	97	2	8	NCBInr	9.1	25674
	6M Guanidine-HCl pH 6.0 elution.						
6	tr|A0AP03|CG4299 protein.[*Drosophila simulans*]	76	1	3	NCBInr	4.26	30933
7	tr|A0AP03|CG4299 protein.[*Drosophila simulans*]	60	1	3	NCBInr	4.26	30933
8	Q3KRE8|Tubulin beta-2B chain - *Rattus norvegicus*	595	10	22	NCBInr	4.78	49921
9	P15159|Troponin C.[*Tachypleus tridentatus*]	176	4	11	NCBInr	4.11	17422
10	tr|A7TCL7|Predicted protein.[*Nematostella vectensis*]	206	4	21	NCBInr	11.09	20183
11	tr|Q4PM63|Histone H2B.[*Ixodes scapularis*]	205	5	20	NCBInr	10.52	13811
12	tr|Q4PM63|Histone H2B.[*Ixodes scapularis*]	190	5	26	NCBInr	10.52	13811
13	tr|A7TCL7|Predicted protein.[*Nematostella vectensis*]	446	9	29	NCBInr	11.09	20183
14	tr|A7TCL7|Predicted protein.[*Nematostella vectensis*]	90	4	9	NCBInr	11.09	20183
15	tr|Q4PM63|Histone H2B.[*Ixodes scapularis*]	164	3	26	NCBInr	10.52	13811

16	tr|Q4PM63|Histone H2B.[*Ixodes scapularis*]	273	10	28	NCBInr	10.52	13811
17	tr|A5Y4F5|Histone H3a (Fragment).[*Ostracoberyx dorygenys*]	207	4	21	NCBInr	10.83	12368
18	tr|A7TCL7|Predicted protein.[*Nematostella vectensis*]	156	6	15	NCBInr	11.09	20183
19	tr|Q4PM69|Histone H4.[*Ixodes scapularis*]	237	10	40	NCBInr	11.36	11360
20	tr|A7S6Z7|Predicted protein.[*Nematostella vectensis*]	378	9	25	NCBInr	10.67	24217
21	tr|Q9U7Q1|Beta tubulin.[*Trichuris trichiura*]	196	3	6	NCBInr	4.77	49790
23	ens|ENSXETP00000039264|ENSXETG00000018102 transcript:ENSXETT00000039264 [*Xenopus tropicalis*]	72	1	1	NCBInr	9.11	54293

Table 4.2: Identifications of proteins from unengorged female Ixodes holocyclus subjected to equalisation using a Proteominer hexapeptide library, eluted with UTC7, and separated by isoelectric focusing and SDS-PAGE. Protein spots were excised from the gel, reduced and alkylated, then digested overnight with trypsin. The resulting peptides were then separated and analysed by LC/MS/MS as described in chapter two. Data was then searched against the NCBI non redundant database using the Mascot search engine. Gel images showing the location of all spots are presented in appendix.

Other protein-protein interactions that would not normally occur, but may, due to the artificially high protein concentration are also possible. This can affect the equalisation in two ways. Firstly, if a protein binds to a high abundance protein whose population saturates its specific hexapeptide bead, the protein bound to the high abundance protein will be washed away with the unbound high abundance protein resulting in the loss of protein concentration. This could also work in reverse with low abundance proteins being bound by the beads also binding multiple high abundance proteins. Holocyclotoxin is an example of a low abundance protein that binds too many different host serum proteins, especially albumin (Thurn, 1994). Alternatively, proteins that bind to other ligand-bound proteins, possibly as part of a protein complex, as well as their own ligand will have their concentration increased effecting the equalisation. Possible strategies to deal with this issue will be discussed in section 4.3.5.

Treatment of the beads with 6M Guanidine-HCl shows that UTC7 is not a perfect elution buffer and it does not desorb proteins requiring high ionic strength (Boschetti *et al.*, 2007). Guanidine-HCl possesses both strong chaotropic effects and high ionic strength, but is not directly compatible with isoelectric focusing or other chromatographic separation techniques without desalting (Boschetti *et al.*, 2007). Figure 4.5D shows the protein present in the guanidine elution after removal of the guandine with a BioSpin centrifugal concentrator (Bio-Rad). A number of proteins are still present with either acidic or alkaline pIs, although the gel shows spots at the acidic end. Incomplete reduction and alkylation of this sample has caused histones, identified as such by mass spectrometry, to be bound to another protein or acidic compound and dragged through the focusing strip away from its correct pI. A lack of sample meant that the gel could not be repeated. The SDS sample was not isoelectrically focused due to the difficulty of removing the SDS from the sample. The presence of SDS in the sample would interfere with proper focusing of the proteins. Analysis of the protein in this sample could be achieved using a 1-D SDS-PAGE/1-DLC/MS experiment of the relevant SDS-PAGE lane from figure 4.4.

However, the apparent equalisation of the sample resulted in the identification of only one new protein not previously identified in chapter two. This protein was a catalase homologue identified with a single high scoring peptide. Only 48 of the 69 spots excised and subjected to trypsin digest/mass spectrometry were identified with at least one high scoring peptide. The other 21 spots returned no result even though peptides were selected and fragmented. This problem has been highlighted in previous chapters and needs no further discussion, as it is a recurrent theme throughout this work.

4.3.3: Fully engorged tick protein equalisation.

In almost all cases, the published literature of proteome equalisation using Proteominer uses the beads in a column format (Castagna *et al.*, 2005, Guerrier *et al.*, 2007, Sennels *et al.*, 2007, Simo *et al.*, 2008) rather than the rotating wheel method provided in the Proteominer kit for equalising plasma. A column format was also used for this particular experiment due to the volume of sample needing to be equalised. The volume of PBS needed (90mL) to solubilise the protein present in 36 fully engorged ticks was simply too large for a single tube to be used to incubate sample and beads. A key consideration underpinning the mechanism of equalisation using Proteominer to enrich for low abundance proteins is to not split the sample into smaller aliquots and use multiple columns for incubation on a rotating wheel. The splitting of a sample, use multiple columns and recombining of the eluted “equalised” protein would not result in a true equalisation. In addition, the expense of the Proteominer columns prevents this. Thus the column format was used for equalising this sample, but some washes and all elutions were carried out in the provided spin columns. This allowed equalised sample recovery in relatively small volumes.

The use of a rotating wheel is the suggested method for equalising plasma samples of 1mL. In spite of the mentioned problem of sample volume, the equalisation of larger samples (10-50mL at 10+mg/mL) on a rotating wheel may provide better conditions for

the enrichment of low abundance proteins because of the higher likelihood of a protein and its hexapeptide ligand encountering one another when both are free in solution. The Proteominer process relies on the system reaching equilibrium, which is essential if reproducible results are to be obtained. In contrast, a chromatographic approach where the raw sample is passed over a Proteominer column may not provide sufficient interaction time for equilibrium to be reached. In a column under moderate pressure, an individual bead would be very restricted in its movement within the column. If it is assumed that a single bead contains a single type of hexapeptide, this restriction of movement becomes crucial as the likelihood of a protein encountering that single bead is reduced and thus the protein does not contribute to the equalisation. Of course, more than one type of hexapeptide will bind a certain type of protein with varying affinity and this would minimise the effect of the above scenario. Recirculation of the sample over the beads a number of times would increase the likelihood of a protein binding to its hexapeptide but this extends the equalisation time, especially with large samples volumes such as in this experiment.

Figure 4.6 shows SDS-PAGE separations of the protein isolated during sequential extraction of fully engorged *I.holocyclus* and the protein eluted from the Proteominer beads with UTC7 and 6M guanidine-HCl. Once again, the gel of the Proteominer elutions shows that unbuffered UTC7 is not a good elution reagent when used in isolation. After 10 x 500µL elutions with UTC7, protein is still eluting from the beads although the majority was eluted with the first 500µL (lanes 18-27). In contrast 2 x 500µL elutions with 6M guanidine-HCl was sufficient to remove the protein able to be removed by this elution reagent. Closer inspection of the pattern of protein bands in the gel shows that UTC7 elutions 5-10 (lanes 22-27) contain the same pattern of proteins eluted by 6M guanidine-HCl (lanes 29 and 30), but at a reduced concentration as indicated by the band intensity. Boiling of the beads with SDS sample buffer was not performed in this case, but the previous results show that additional protein most likely would have been eluted.

At this point the continued use of unbuffered UTC7 to elute equalised protein from the beads needs to be explained. UTC7 was used because no other elution buffer that could be used was directly compatible with 2-D-PAGE. Although 6M guanidine-HCl is an effective elution buffer and more powerful than UTC7, its subsequent removal was problematic and resulted in loss of sample on some occasions (data not shown). Boiling the beads in SDS sample buffer is very effective, but the sample is then limited to separation by 1-D SDS-PAGE and analysis by a 1-DLC/MS experiment. This approach provides complimentary, but not the same, data to that obtained from 2-DGE. Buffering the UTC7 makes little difference to its elution power. Rather, buffering with either Tris-HCl or citric acid is useful for sequential elutions, further fractionating the equalised sample as demonstrated in section 2.3.1. The choice in these experiments was made to elute as many proteins as possible in a single elution step and unbuffered UTC7 was the buffer chosen to do this.

The two-dimensional gels depicting equalized and non-equalised engorged tick extracts (Figure 4.7A-F) are quite striking in their difference. Significantly more proteins are present on the gel of the equalized sample. Gels A and B show the unequalised PBS extract, separated by isoelectric focusing and SDS-PAGE. In gel A the proteins have not been reduced and alkylated while the proteins in gel B have undergone reduction with TBP and alkylation with acrylamide. As expected the spot patterns of the two samples differ. The most striking difference is the removal of haemoglobin multimers at the alkaline end of the gel where the sample is reduced and alkylated. Visual comparison of gels A and B with gels E and F that are loaded with the protein that does not bind to the Proteominer beads during equalisation would imply that the spot number and intensity are similar. Approximately 230 spots were determined by Proteomeweaver to be present in gel A and 177 in gel E. This demonstrates that protein has indeed been removed from the unbound fraction, and is bound to the beads.

Figure 4.7C and D are gels of Proteominer equalised, PBS extracted protein of engorged female *I.holocyclus*. Once again, the protein loaded on upper gel (C) has not been reduced and alkylated while the protein present on the lower gel (D) has been reduced with TBP and alkylated with acrylamide.

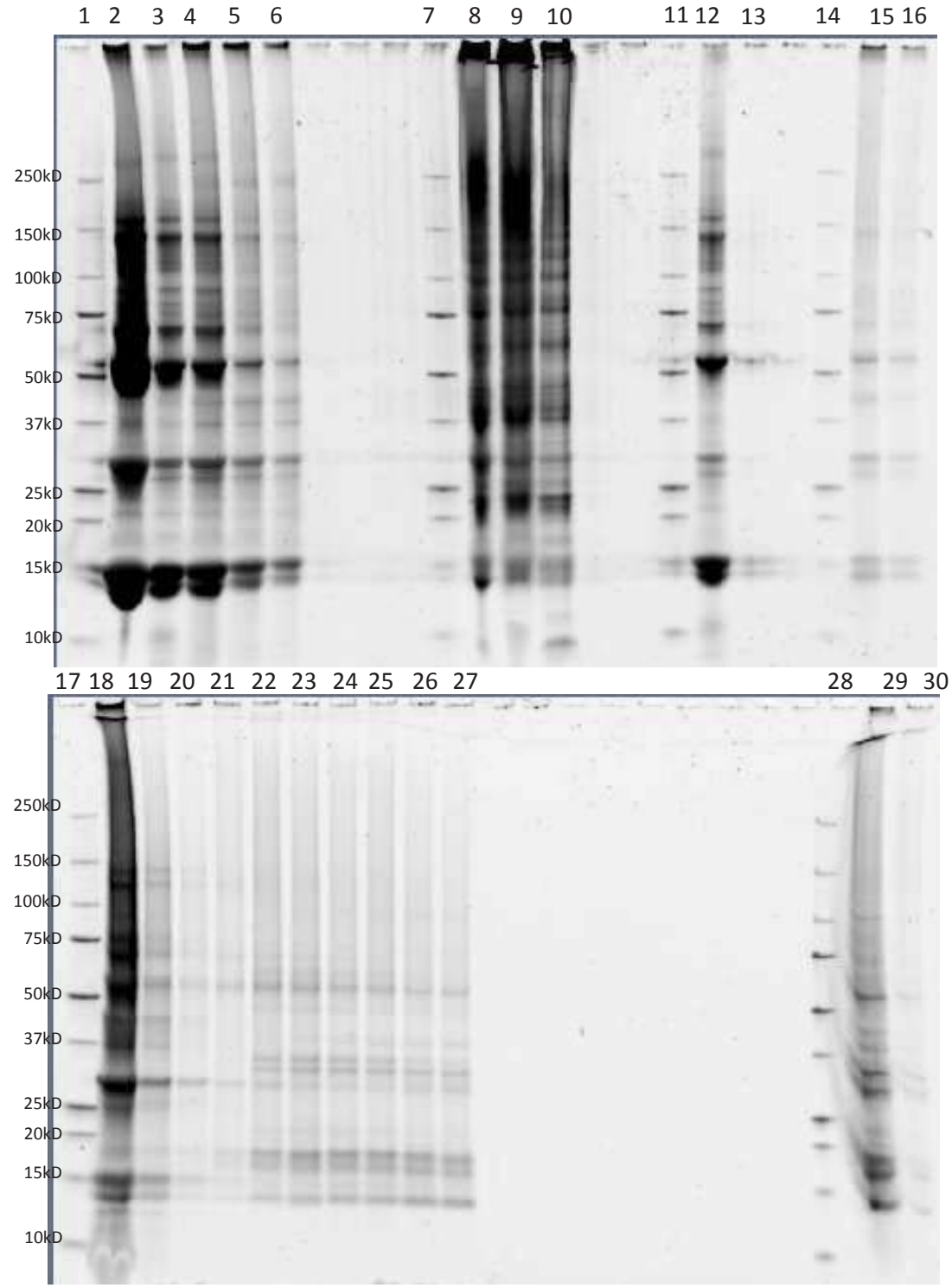

Figure 4.6: SDS-PAGE separation of sequentially extracted, fully engorged female *I.holocyclus* and Proteominer elutions of equalised PBS extracted, partially engorged female *I.holocyclus*.

Lanes 1, 7, 11, 14, 17 and 28: Molecular weight markers.
Lanes 2-6: Sequential PBS extractions of engorged ticks.
Lanes 8-10: UTC7Tris extractions of pellet after PBS extract.
Lanes 12-13: Unbound sample from Proteominer incubation.
Lanes 15-16: Proteominer PBS washes.
Lanes 18-27: UTC7 elutions Proteominer.
Lanes 29-30: Gu-HCl elutions Proteominer.

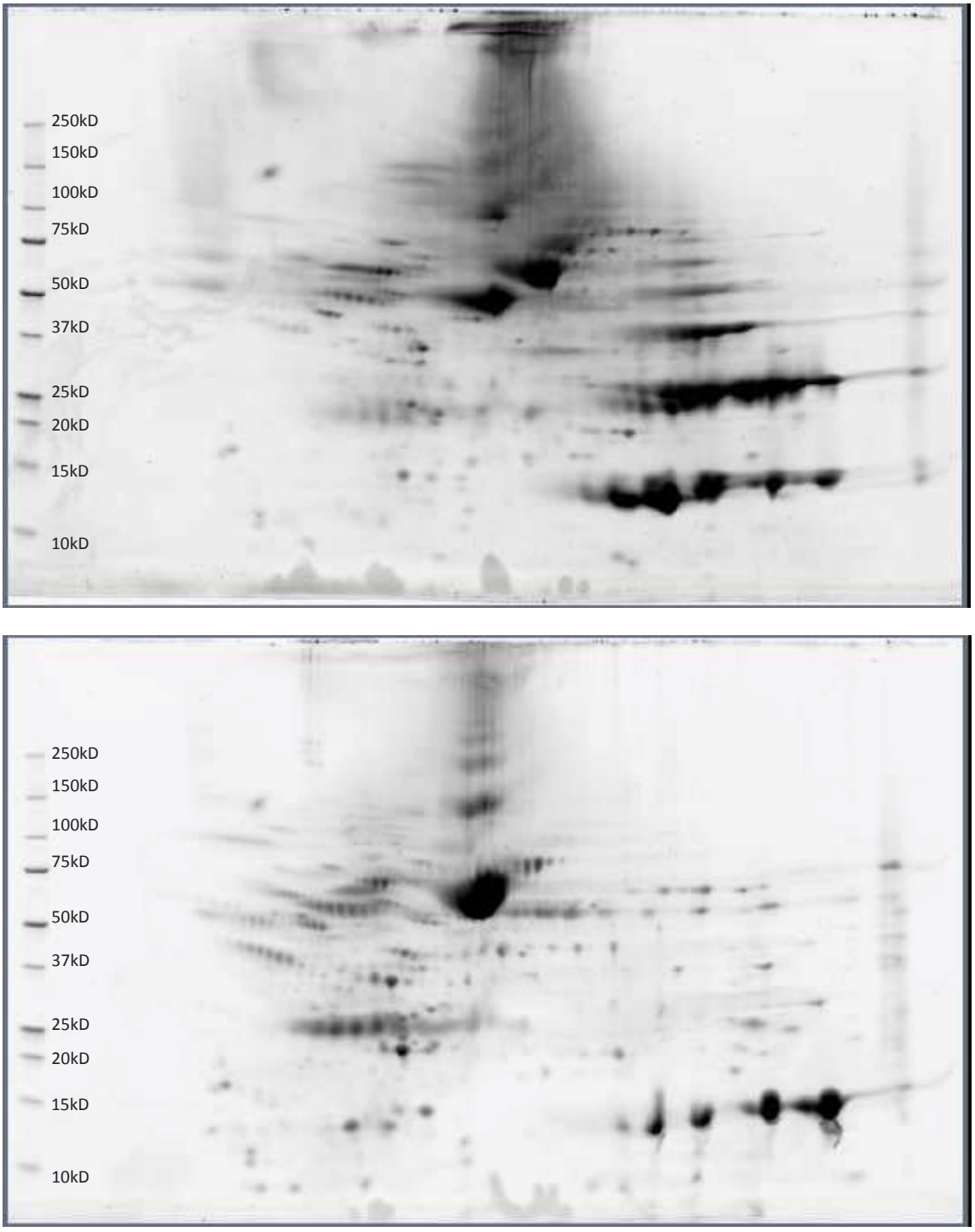

Figure 4.7: PBS extracted protein from fully engorged female *I.holocyclus* separated by isoelectric focusing and SDS-PAGE. The protein sample on gel A (upper) has not been reduced or alkylated while the proteins on gel B (lower) have been subjected to reduction with TBP and alkylation with acrylamide.

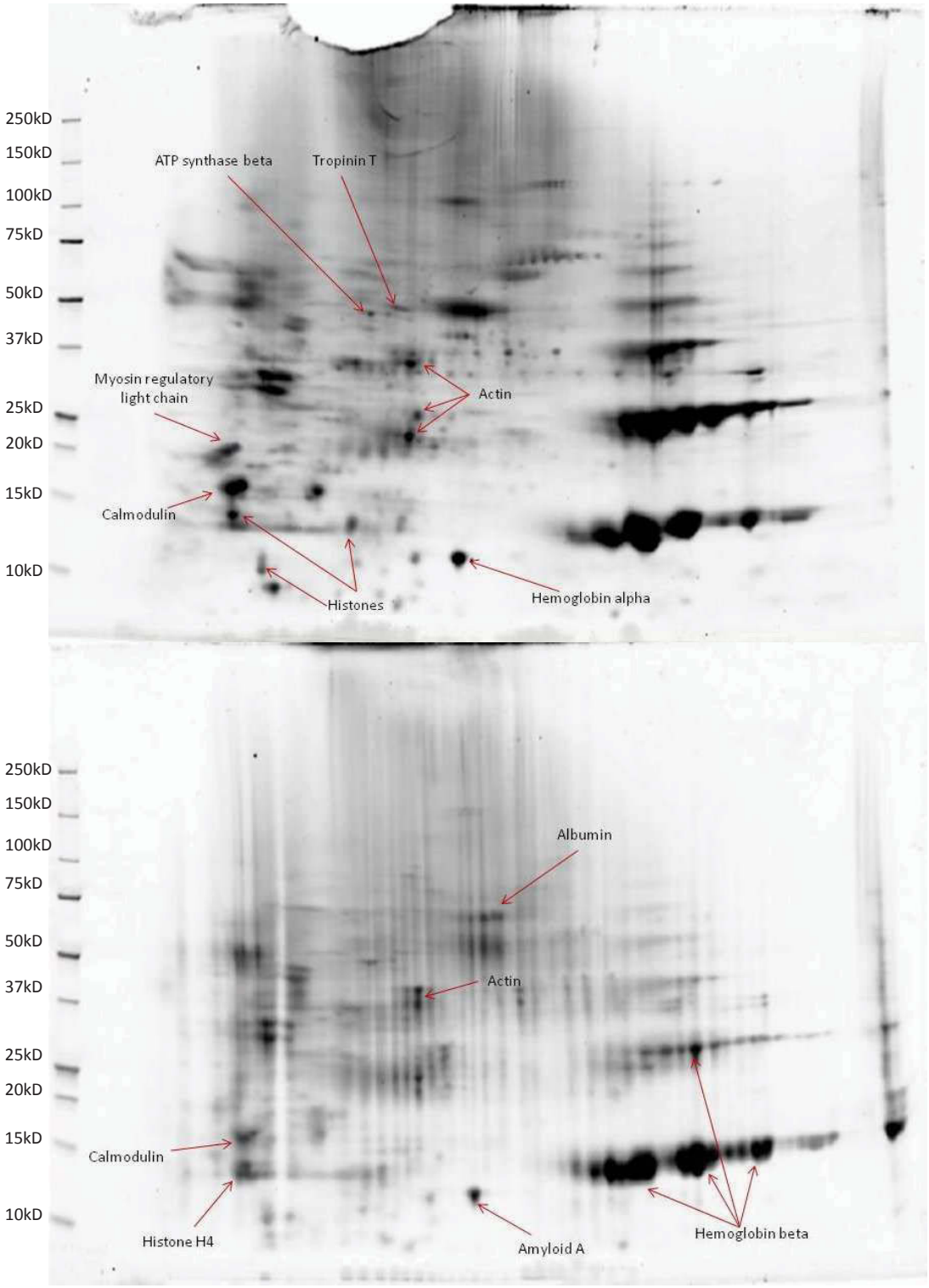

Figure 4.7: Proteominer equalised PBS extracted protein from fully engorged female *I.holocyclus* eluted from the beads with UTC7 then separated by isoelectric focusing and SDS-PAGE. The protein sample on gel C (upper) has not been reduced or alkylated while the proteins on gel D (lower) have been subjected to reduction with TBP and alkylation with acrylamide.

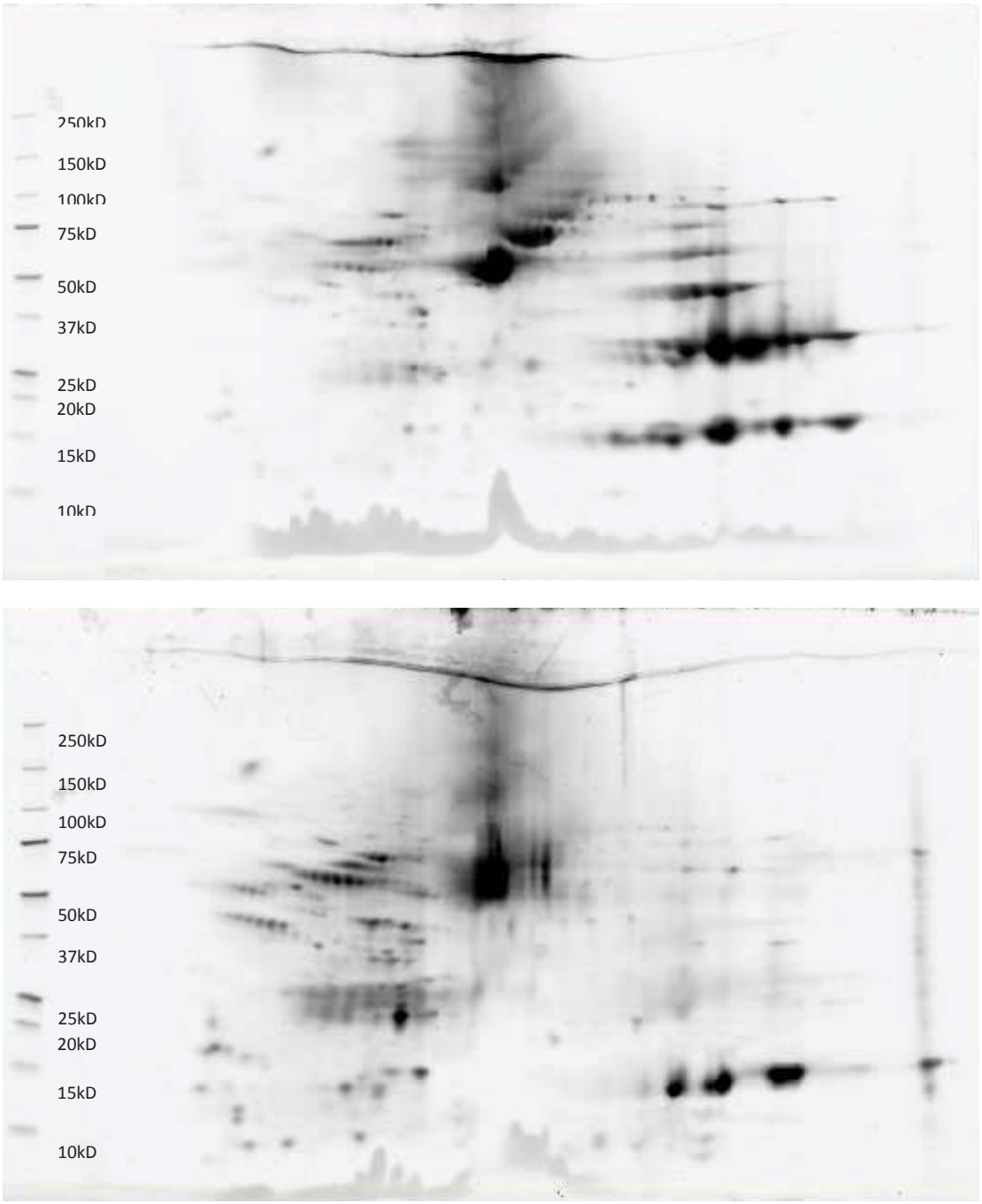

Figure 4.7: Unbound protein from PBS extracted protein from fully engorged female *I.holocyclus* equalised by Proteominer beads. This unbound protein fraction was then separated by isoelectric focusing and SDS-PAGE. The protein sample on gel E (upper) has not been reduced or alkylated while the proteins on gel F (lower) have been subjected to reduction with TBP and alkylation with acrylamide.

Spot #	Accession number/Description/Species	pI	MW	Score	Matches	Unique peptides	Coverage (%)	Database
	Reduced and alkylated (Gel 4.5D)							
1	ens transcript:ENSOANT00000001257[*Ornithorhynchus anatinus*]	9.23	14307	54	1	1	11	NCBInr
2	tr|A7SHX4|Predicted protein.[*Nematostella vectensis*] (Histone H2A)	10.82	24545	337	12	6	29	trEMBL
3	tr|Q98UH8|Calmodulin (Fragment).[*Oemmys japonica*]	4.05	15349	383	11	7	48	trEMBL
4	ens transcript:ENSMODT00000008866[*Monodelphis domestica*]	4.72	37987	427	10	9	25	NCBInr
5	ens transcript:ENSCAFT00000016273[*Canis familiaris*]	4.56	42286	206	4	4	11	NCBInr
6	ens transcript:ENSMLUT00000001219[*Myotis lucifugus*]	4.84	42286	159	2	2	8	NCBInr
7	ens transcript:ENSMODT00000008866[*Monodelphis domestica*]	4.72	37987	462	16	10	22	NCBInr
8	ens transcript:ENSMODT00000008866[*Monodelphis domestica*]	4.72	37987	465	14	11	25	NCBInr
9	sp|Q5H1U7|Hemoglobin subunit beta.[*Macropus rugenii*]	7.1	15122	145	3	3	23	SwissProt
10	sp|Q5H1U7|Hemoglobin subunit beta.[*Macropus rugenii*]	7.1	15122	212	2	2	14	SwissProt
11	tr|Q5CAR2|Actin.[*Ixodes ricinus*]	5.3	41797	186	4	4	12	trEMBL
12	tr|Q4PKE5|Actin 5.[*Aedes aegypti*]	5.3	41795	583	27	14	37	trEMBL
13	tr|A9QUS4|Beta-actin.[*Rachycentron canadum*]	5.31	41754	906	38	19	52	trEMBL
14	tr|Q9DD57|Beta tubulin [*Chionodraco rastrospinosus*]	4.74	49718	754	25	20	35	trEMBL
15	ens transcript:ENSMODT00000024341[*Monodelphis domestica*] (Albumin)	5.66	68048	242	7	6	13	NCBInr
16	sp|Q5H1U7|Hemoglobin subunit beta.[*Macropus rugenii*]	7.1	15122	252	7	4	24	SwissProt
17	sp|Q5H1U7|Hemoglobin subunit beta.[*Macropus rugenii*]	7.1	15122	389	57	7	50	SwissProt
18	sp|Q5H1U7|Hemoglobin subunit beta.[*Macropus rugenii*]	7.1	15122	223	12	5	35	SwissProt
19	sp|Q5H1U7|Hemoglobin subunit beta.[*Macropus rugenii*]	7.1	15122	344	24	6	35	SwissProt
20	sp|P01934|Hemoglobin subunit alpha-3.[Gorilla gorilla]	9.72	15228	204	28	3	14	SwissProt
22	sp|Q5H1U7|Hemoglobin subunit beta.[*Macropus rugenii*]	7.1	15122	121	3	3	23	SwissProt
23	tr|Q4PM69|Histone H4.[*Ixodes scapularis*]	11.36	11360	121	4	3	33	trEMBL
24	tr|Q29574|Histone H2B (Fragment).[*Sus scrofa*]	10.11	7626	197	2	2	34	trEMBL
	Not reduced and alkylated (Gel 4.5C)							
25	tr|Q98UH8|Calmodulin (Fragment).[*Oemmys japonica*]	4.05	15349	250	7	6	40	trEMBL

25	tr|Q98UH8|Calmodulin (Fragment).[*Clemmys japonica*]	4.05	15349	338	10	9	48	trEMBL
27	sp|Q3THE2|Myosin regulatory light chain 2-B, smooth muscle isoform.[*Mus musculus*]	4.71	19767	178	4	4	23	SwissProt
28	ens transcript:ENSMODT00000008866[*Monodelphis domestica*]	4.72	37987	252	5	5	21	NCBInr
29	ens transcript:ENSMODT00000008866[*Monodelphis domestica*]	4.72	37987	171	4	4	18	NCBInr
30	ens transcript:ENSCAFT00000016273[*Canis familiaris*]	4.56	43925	176	2	2	8	NCBInr
31	tr|A7SHX4|Predicted protein.[*Nematostella vectensis*] (Histone H2A)	10.82	24545	216	8	4	23	trEMBL
32	tr|A0SNU5|Beta-actin (Fragment)[*Sus scrofa*]	5.5	29393	197	7	6	31	trEMBL
33	tr|A0SNU5|Beta-actin (Fragment)[*Sus scrofa*]	5.5	29393	474	17	10	44	trEMBL
34	tr|A9CUS4|Beta-actin.[*Rachycentron canadum*]	5.31	41754	757	26	15	50	trEMBL
35	tr|A2BE93|SET translocation (Fragment).[*Mus musculus*]	5.43	24909	186	4	3	22	trEMBL
36	tr|Q6PTP3|ATP synthase subunit beta (EC 3.6.3.14) (Fragment).[*Encope michelini*]	4.96	45811	682	20	15	35	trEMBL
37	tr|A8E4K0|Troponin T.[*Haemaphysalis qinghaiensis*]	4.93	45442	268	10	7	12	trEMBL

Table 4.3: Identifications of proteins from engorged female *I. holocyclus* subjected to equalisation using a Proteominer hexapeptide library, elutec with UTC7, and separated by isoelectric focusing and SDS-PAGE (figure 4.5C & D). Protein spots were excised from the gel, reduced and alkylated, then digested overnight with trypsin. The resulting peptides were then separated and analysed by LC/MS/MS as described in chapter two. Data was then searched against the NCBI non redundant database using the Mascot search engine. Gel images showing the location of all spots are presented in appendix

As with the unequalised PBS extract, reducing and alkylating the sample prior to focusing removes haemoglobin multimers. However, protein loss has occurred in the reduced and alkylated sample for reasons unknown. As this gel was not of critical importance to this particular experiment, it was not repeated. Comparing gel A with gel C, it is apparent that a great many spot's intensity has increased, especially at the acidic end of the gel. Using Proteomeweaver, approximately 280 spots were found, an increase of 50. The intensity of the large protein spot in the centre of the gel has been greatly reduced. This protein is most likely highly abundant host albumin from the blood meal.

On examination of all the gels, it is apparent that the number and intensity of haemoglobin spots has not changed. It was expected that the intensity of these spots should be reduced in the equalised sample (figure 4.7C) when compared to the unequalised PBS extract (figure 4.7A). This does not seem to be the case. A possible explanation for this could be the extremely high concentration of the sample within the engorged tick itself prior to extraction. The process of tick feeding concentrates the blood meal by returning any water back to the host, raising the proteins concentration well beyond the level achievable in the host's serum. This is likely to lead to protein-protein interactions that would not normally occur and more importantly, may not be reversed when the solubilisation of the sample is attempted. These artificial protein complexes will then cause an unequal equalisation. Methodologies to deal with these problems are discussed in section 4.3.5.

4.3.4: Partially engorged tick protein equalisation.

The last tick feeding stage examined using equaliser technology was partially engorged ticks, that were removed from their hosts early in the engorgement process (1-3 days). It was hoped that comparison between partially engorged and fully engorged ticks would reveal different tick proteins associated with different stages of feeding.

Figure 4.8 shows the 2-D gel separated unequalised PBS extracted protein (A) and the equalised protein eluted from the beads with unbuffered UTC7 (B). As expected from the previously presented equalised samples, the gels are different. The increase in acidic protein spots is evident, but a number of spots have been greatly increased in intensity in the middle part of the IPG strip. Approximately 230 spots were found in gel A and approximately 260 in gel B when both gels were analysed using Proteomeweaver. This is similar to the number of spots found in the fully engorged sample.

Only five non-mammalian proteins were identified in the unequalised PBS extracts from the gel spots randomly excised. Only two, actin and cytochrome c oxidase were identified as tick proteins while the others identified closely with bee and mosquito homologues. The other spots were mainly identified as mammalian globins as expected. The equalised sample was quite different with approximately 50% of the proteins identifying to ticks or other insects. Protein spots not seen in the unequalised sample including the tick proteins calreticulin, tropomyosin, and protein disulphide isomerise. All of these were identified with 3 or more high-scoring peptides. In addition, some previously unseen proteins were identified, tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein epsilon polypeptide and Glucose-regulated protein 78kDa.

When compared to the gels of the equalised engorged tick extract (figure 4.7C and D), a number of tick protein homologues have been identified in the semi-engorged extract that have not been identified in the engorged extract. However, the difference is not significant and is more likely due to the poor focusing and SDS-PAGE of the equalised engorged sample in figure 4.6D. Many spots that are expected to be present,

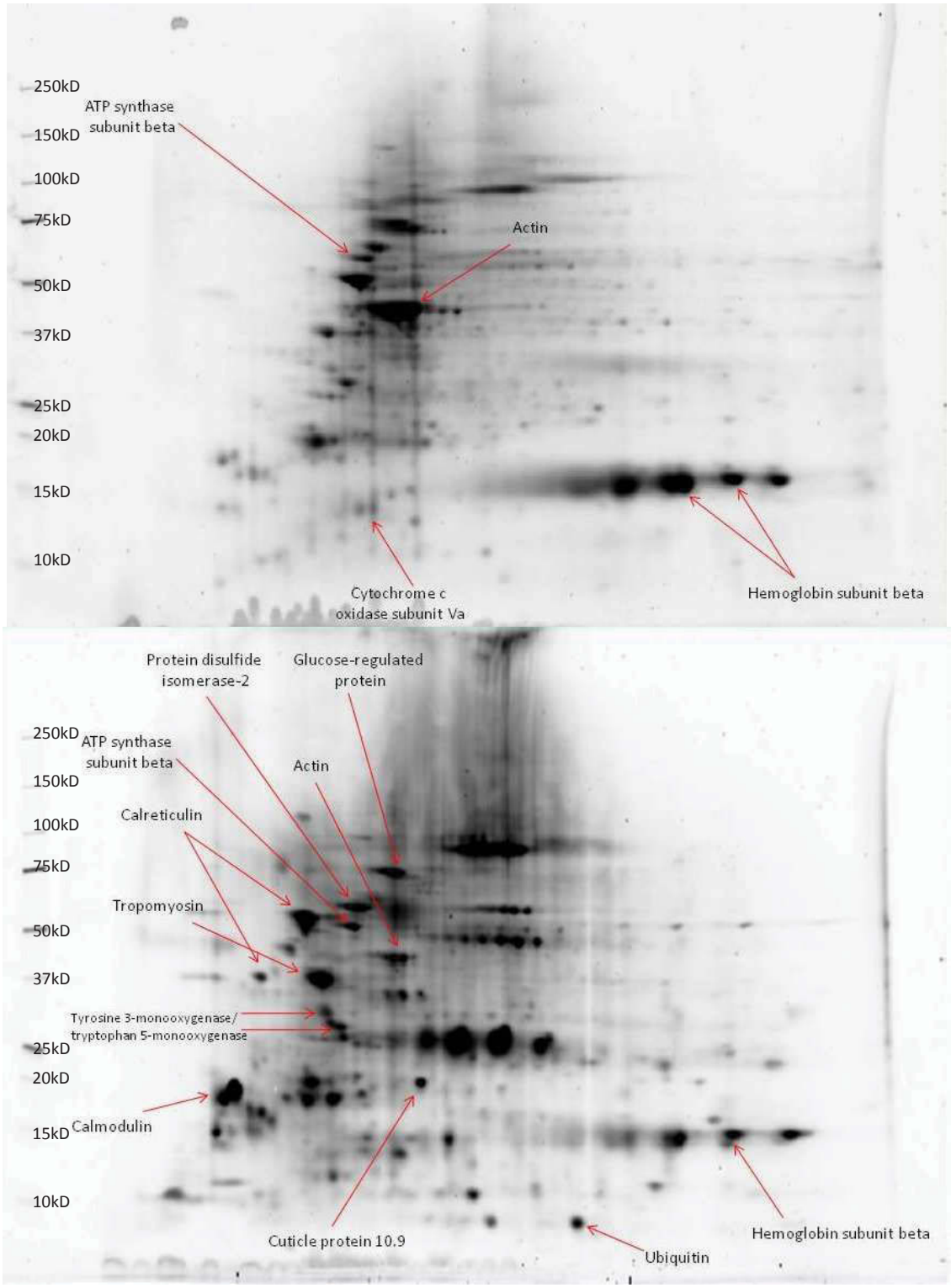

Figure 4.8: Gel A (upper) shows unequalised PBS extracted of partially engorged *I. holocylcus*. Gel B (lower) shows the UTC7 elution of equalised PBS extracted partially engorged ticks. Both samples have been reduced and alkylated prior to isoelectric focusing.

Spot #	Accession number/Description/Species	pI	MW	Score	Matches	Coverage (%)	Database
	PBS Extract (Unequalised)						
E1	tr|A8DUP5|Beta-globin.[Mus musculus]	7.1	15884	726	27	74	trEMBL
E2	tr|A8DUP5|Beta-globin.[Mus musculus]	7.1	15884	522	9	55	trEMBL
E3	tr|A8DUP5|Beta-globin.[Mus musculus]	7.94	15223	272	4	37	trEMBL
E5	tr|A2TK61|14-3-3zeta.[Bombyx mori]	4.9	28079	190	4	14	trEMBL
E6	transcript:ENSAPMT00000014867 [Apis mellifera]	4.78	23531	56	1	4	NCBInr
E8	transcript:ENSCAFT00000021138 [Canis familiaris]	5.28	30163	630	12	34	NCBInr
E9	tr|Q09JF2|Cytochrome c oxidase subunit Va.[Argas monolakensis]	6.32	17753	66	1	5	trEMBL
E10	tr|Q09JF2|Cytochrome c oxidase subunit Va.[Argas monolakensis]	6.32	17753	107	2	6	trEMBL
E11	tr|Q6PTP3|ATP synthase subunit beta (EC 3.6.3.14) (Fragment).[Encope michelini]	4.96	45811	792	14	34	trEMBL
E12	transcript:ENSGACT00000018465 [Gasterosteus aculeatus]	5.22	57348	463	6	14	NCBInr
F1	tr|Q7Q270|AGAP004002-PA.[Anopheles gambiae str. PEST]	5.55	60740	396	6	12	trEMBL
F3	tr|Q4FKE5|Actin 5.[Aedes aegypti]	5.3	41795	958	31	54	trEMBL
F4	transcript:ENSCAFT00000009239 [Canis familiaris]	5.24	41711	133	3	14	NCBInr
F5	tr|Q5CAR2|Actin.[Ixodes ricinus]	5.3	41797	425	9	28	trEMBL
F7	transcript:ENSCAFT00000022855 [Canis familiaris]	7.83	16117	638	18	65	NCBInr
F8	transcript:ENSCAFT00000022855 [Canis familiaris]	7.83	16117	556	12	64	NCBInr
F11	sp|Q5JAK2|Actin, cytoplasmic 2 (Gamma-actin) (Cytoplasmic actin type 5).[Rana lessonae]	5.3	41752	362	9	27	SwissProt
	Equalised protein sample (UTC7 elution)						
B1	tr|Q54K84|Calreticulin.[Ixodes jellisoni]	4.57	47612	339	7	16	trEMBL
B2	tr|Q54K84|Calreticulin.[Ixodes jellisoni]	4.57	47612	199	3	10	trEMBL
B3	sp|O97162|Tropomyosin.[Boophilus microplus]	4.7	32982	432	10	18	SwissProt
B4	tr|A5LHV9|Protein disulfide isomerase 2.[Haemaphysalis longicornis]	4.84	55871	91	3	6	trEMBL
B5	tr|Q6PTP3|ATP synthase subunit beta (EC 3.6.3.14) (Fragment).[Encope michelini]	4.96	45811	721	11	31	trEMBL
B6	sp|P02769|Serum albumin precursor (Allergen Bos d 6) (BSA).[Bos taurus]	5.82	69248	99	2	3	SwissProt
	tr|Q1HQQ5|Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein epsilon polypeptide.[Aedes aegypti]	9.6	25881	70	2	4	trEMBL

B7	tr|Q1HQQ5|Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein epsilon polypeptide.[Aedes aegypti]	9.6	25881	71	2	4	trEMBL
B8	tr|A2TK61|14-3-3zeta.[Bombyx mori]	4.9	28079	183	5	19	trEMBL
B9	sp|P02769|Serum albumin precursor (Allergen Bos d 6) (BSA).[Bos taurus]	5.82	69248	342	4	7	SwissProt
B10	sp|P04244|Hemoglobin subunit beta (Hemoglobin beta chain) (Beta-globin).[Panthera pardus orientalis]	7.1	16107	237	4	30	SwissProt
B11	sp|P84251|Cuticle protein 10.9 (Ir-ACP10.9).[Ixodes ricinus]	4.46	10914	77	1	13	SwissProt
C3	tr|Q4PKE5|Actin 5.[Aedes aegypti]	5.3	41795	378	10	26	trEMBL
C4	sp|Q5JAK2|Actin, cytoplasmic 2 (Gamma-actin) (Cytoplasmic actin type 5).[Rana lessonae]	5.3	41752	358	10	24	SwissProt
C5	tr|Q5DW64|Glucose-regulated protein 78kDa (Fragment).[Oncorhynchus mykiss]	5.02	69793	540	9	11	trEMBL
C6	transcript:ENSGACT00000022010 [Gasterosteus aculeatus]	5.04	73465	462	8	12	NCBInr
C8	tr|A3DM56|Putative uncharacterized protein.[Staphylothermus marinus]	9.53	38362	65	1	2	trEMBL
D8	sp|P60524|Hemoglobin subunit beta (Hemoglobin beta chain) (Beta-globin).[Canis familiaris]	7.83	15986	579	16	65	SwissProt
D10	Ubiquitin (mulitple organisms)		14587	146	2		NCBInr
A1-8	sp|P02595|Calmodulin (CaM).[Patinopecten sp.]	4.04	16802	284	5	29	SwissProt
A10	transcript:ENSCAFT00000021138 [Canis familiaris]	5.28	30163	565	10	39	NCBInr
A12	tr|Q8MWP3|Calreticulin.[Boophilus microplus]	4.49	47707	67	2	2	trEMBL

Table 4.4: Identifications of proteins from partially engorged female Ixodes holocyclus subjected to equalisation using a Proteominer hexapeptide library, eluted with unbuffered UTC7 then separated by isoelectric focusing and SDS-PAGE. Protein spots were excised from the gel, reduced and alkylated, then digested overnight with trypsin. The resulting peptides were then separated and analysed by LC/MS/MS as described in chapter three. Data was then searched against MSDB and NCBI non redundant databases using the Mascot search engine. Gel images showing the location of all spots are presented in appendix.

such as calreticulin and tropomyosin, which are present in the gel of engorged tick extract separated in the preliminary experiment (figure 4.3), are not in figure 4.7D.

The aim of this experiment and the experiment presented in the previous section (4.3.3) was to determine whether the tick and host proteome are different at different stages of engorgement. This has not been successful, even though equalisation has increased the number of tick protein homologues in both samples able to be identified. However these tick proteins are higher abundance proteins previously seen in experiments with unengorged ticks (chapter two and section 4.3.2). In addition, no host response proteins were identified following equalisation. The most likely explanation for this is that the amount of starting material is insufficient. 100 unengorged ticks were used for the equalisation performed in section 4.3.2 and, even though spot intensity was increased, only a single new protein homologue was identified. The equalisation of the protein from 100 engorged ticks would require the sample to be diluted to a very large volume (approximately one litre). This is necessary to ensure the protein concentration is low enough (<10mg/mL) to minimise protein-protein interactions and the possibility of a distorted equalisation as described in the previous section. Such a large volume would necessitate performing the equalisation in a column format. The potential problems with the column format are detailed in section 4.3.3.

4.3.5: Fully engorged tick protein equalisation after reduction and alkylation.

As stated in the introduction to this chapter, one of the main limitations of Proteominer is the need for the proteins to be equalised to be soluble in a physiological buffer, such as PBS or Tris-HCl. However, examination of the gel images of sequentially extracted tick proteins in chapter two shows that less than 50% of the proteome is soluble in PBS and Tris-HCl pH 8.8. Much of the remaining protein is able to be solubilised in UTC7Tris-HCl pH 8.8 + 150mM LiCl, a highly chaotropic and high ionic strength buffer. In addition to the necessity of being soluble in a physiological buffer, protein-protein interactions are still present. Thus some stable protein

complexes will remain but more importantly, unwanted protein-protein interactions may be present in the sample. This is especially a problem with samples containing proteins such as albumin that are able to bind a number of other proteins and in engorged ticks where blood meal protein concentration is artificially high. This effect can make the equalisation biased to particular proteins depending on their concentration and the dynamic range of proteins in the solution being equalised.

The traditional method for dealing with protein-protein interactions is to reduce and alkylate all proteins in the sample prior to equalisation. However this also has some serious limitations due to the reduced solubility of alkylated proteins. In this experiment, an attempt to deal with the solubility problem was made by resuspending precipitated protein in 30% trifluoroethanol (TFE) in PBS. TFE has been successfully used to solubilise other protein samples in our lab, specifically membrane proteins from brain (Michael Sivell - unpublished data), but it had not been used to solubilise alkylated proteins until this experiment. Following grinding of the ticks, the protein in the powder was precipitated with ice cold 10% trichloroacetic acid in acetone (TCA/acetone) and washed repeatedly with ice cold acetone until the acetone was clear rather than red from released heme. From 40 ground engorged ticks, approximately 20mL of wet protein pellet was produced. 90% of this protein was able to be solubilised with UTC7Tris-HCl pH 8.8 + 150mM LiCl and subsequently reduced and alkylated with double the amount of TBP and acrylamide to ensure all haemoglobin multimers were turned into monomers. This protein was precipitated with acetone and a pellet of approximately 20mL produced. Attempts were made to solubilise this pellet in 80mL of 30% TFE in PBS, but after repeated sonication and centrifugation, only a small amount of protein was able to be solubilised as indicated by a slight colour change in the supernatant.

A possible explanation for TFE not enhancing solubility is that, unlike surfactants that unfold proteins as they interact with hydrophobic parts of the protein, TFE is

enhancing solubility by ion pairing in much the same way that trifluoroacetic acid (TFA) ion pairs with proteins and peptides in reversed phase chromatography, improving peak resolution. In this way, TFE may stabilise folded protein structures that would otherwise be insoluble when a cell, organelle or membrane is disrupted. However when a protein's folded structure is completely disrupted by alkylation, the ion pairing interaction of TFE is not of sufficient strength to solubilise the protein. In contrast to surfactants which have strongly polar head groups to provide solubility when interacting with alkylated protein, TFE can only provide relatively weak hydrogen bonding between the alkylated protein and the aqueous solution being used for solubilisation.

To increase protein solubility, the proteins were alkylated with reagents containing a strongly polar functional group, such as quaternary ammonia. In this case, (3-acrylamidopropyl)-trimethylammonium chloride was used as an alkylating reagent with protein solubilised in UTC7Tris-HCl pH 8.8 + 150mM LiCl. Following acetone precipitation, 50mL 30% TFE in PBS was added and the sample alternately heated to 50^{o}C for 10 minutes and sonicated for 10 minutes to try and encourage the precipitated protein to return to solution. After centrifugation of the sample, visual comparison with unalkylated protein prepared in the same way suggested that more protein had been solubilised in the quaternary ammonium alkylated sample as a smaller pellet was present. The supernatants of both the alkylated and unalkylated samples were equalised with Proteominer beads on a rotating wheel as previously described. The equalised protein was eluted by boiling the beads in 125mM Tris-HCl + 2% SDS. 30µL of each elution was then analysed by 1-D SDS-PAGE and the gel is shown in figure 4.9. Interestingly, the protein band patterns are different, although it is likely that the quaternary alkylation has altered the migration in the gel of the proteins present in that sample. Table 4.5A and B lists the proteins identified from the 30 gel slices analysed from each lane. While α and β-haemoglobin are prominent in many gel slices, many tick and insect proteins are also present, indicating some equalisation has

occurred. In an unequalised sample, it is unlikely any tick proteins would be found. It is interesting to note that the highest abundance tick cytoskeletal protein found in chapter two, actin and its isoforms, was not found in these equalised samples. Actin was found in 2-DGE spots of equalised engorged ticks solubilised in PBS. It is possible that, after being precipitated, actin is not able to be solubilised by 30% TFE in PBS.

The results of the experiment seem to indicate that, in this case, alkylating with (3-acrylamidopropyl)-trimethylammonium chloride has not improved solubility. However the lack of protein solubility could be due to the TCA/acetone precipitation destroying the protein structure to such an extent that only highly chaotropic buffers would return the proteins to solution. In an effort to resolve this, the sample preparation was modified to maximally disrupt protein-protein interactions and maintain solubility. Ammonium chloride has been shown to burst red blood cells and it was thought it would disrupt haemoglobin sufficiently to reduce its interaction with other proteins (Maren and Wiley, 1970). The sample was not reduced or alkylated to increase resolubilisation of the protein after precipitation. TCA/acetone was used more for its ability to remove heme and other pigments than for its protein precipitation power. However the precipitated protein was not able to be returned to solution in PBS. Some protein was solubilised when 30% TFE in PBS was used, but not the amount expected.

The protein that was able to be solubilised was equalised and eluted by boiling the beads in a buffer containing Tris-HCl pH 8.8 and 2% SDS. The protein equalised is shown in figure 4.10. The proteins identified as homologous following sectioning of the gel lanes are presented in table 4.6. Once again, the list contains many homologues to tick and insect proteins indicating that the samples have been enriched for the lower abundance tick proteins. However the equalisation of the sample has not been as dramatic as those in the published literature.

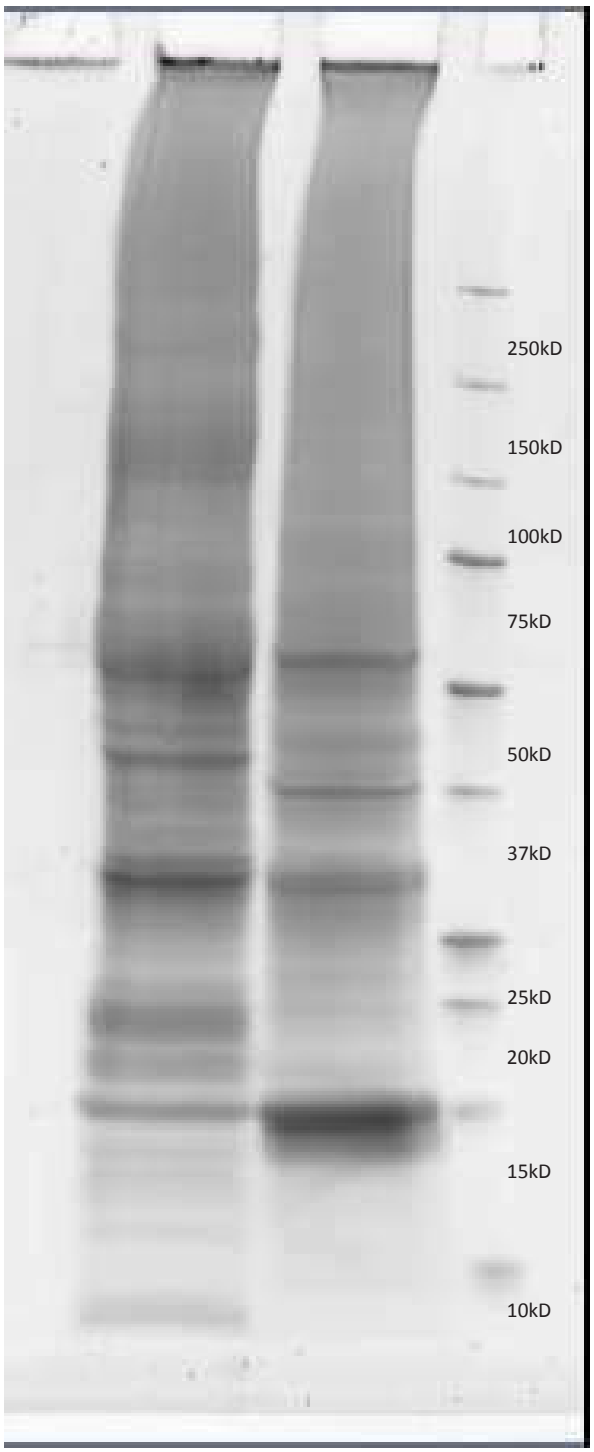

Figure 4.9: 1-D SDS-PAGE of Proteominer equalised engorged tick extract solubilised with 30% TFE in PBS.

1) Sample alkylated with (3-acrylamidopropyl)-trimethylammonium chloride following solubilisation in UTC7Tris-HCl pH8.8 + 150mM LiCl.
2) Unalkylated sample solubilised in UTC7Tris-HCl pH8.8 + 150mM LiCl.
3) Molecular weight markers.

Accession number/Description/Species	MW	Score	Matches	Unique peptides	Coverage (%)	Database	Engine	List position
sp|P02106| Hemoglobin beta chain [*Macropus giganteus*]	16005	1309	580	20	95	SwissProt	Mascot	1
tr|A7SHX4| Predicted protein [*Nematostella vectensis*] (Histone)	24545	805	79	14	42	NCBInr	Mascot	2
sp|O97162| Tropomyosin [*Boophilus microplus*]	32982	575	11	11	28	trEMBL	Mascot	6
sp|P01043| Hemoglobin alpha chain [*Macropus eugenii*]	15327	451	337	7	38	trEMBL	Mascot	10
tr|A6N9R2| Ribosomal protein S18 [*Ornithodoros parkeri*]	17815	437	41	8	36	SwissProt	Mascot	11
sp|Q6R520| Calmodulin [*Oreochromis massambicus*]	16835	364	5	5	44	NCBInr	Mascot	15
tr|A8E4K0| Troponin T [*Haemaphysalis qinghaiensis*]	46442	358	29	7	9	SwissProt	Mascot	17
tr|A5H1Q6| Ribosomal protein L23 [*Haemaphysalis longicornis*]	15450	335	11	4	35	SwissProt	Mascot	19
ens| transcript:ENSMODT00000010348 [*Monodelphis domestica*] (Histone H2A)	14117	331	49	7	35	NCBInr	Mascot	21
tr|Q4PM63| Histone H2B [*Ixodes scapularis*]	13811	320	39	7	28	trEMBL	Mascot	22
ens| transcript:ENSMODT00000004396 [*Monodelphis domestica*] (High-mobility group box-2 protein)	23616	230	4	4	19	trEMBL	Mascot	45
tr|Q4PM82| Ribosomal protein S25 [*Ixodes scapularis*]	12828	216	5	3	18	trEMBL	Mascot	48
ens| transcript:ENSMODT00000024341 [*Monodelphis domestica*] (Albumin)	68048	198	54	3	5	trEMBL	Mascot	57
tr|O97645| Fibrinogen A-alpha chain [*Macropus rufus*]	43865	198	12	2	8	NCBInr	Mascot	58
tr|Q4PM11| 40S ribosomal protein S13 [*Ixodes scapularis*]	17199	192	4	3	18	trEMBL	Mascot	62
tr|Q4PM30| Ribosomal protein S19 [*Ixodes scapularis*]	16509	190	6	3	14	trEMBL	Mascot	63
tr|Q4PMB7| Nonmuscle myosin essential light chain [*Ixodes scapularis*]	16507	180	4	2	16	trEMBL	Mascot	70
sp|O18783| Plasminogen; EC=3.4.21.7 [*Macropus eugenii*]	90922	179	27	3	3	SwissProt	Mascot	71
tr|Q4PM16| 60S ribosomal protein L23 [*Ixodes scapularis*]	14768	155	6	4	24	trEMBL	Mascot	89
tr|Q4PM00| 40S ribosomal protein S30 [*Ixodes scapularis*]	14589	124	8	2	8	trEMBL	Mascot	117
tr|Q4PMD3| 40Sribosomal protein S10 [*Ixodes scapularis*]	18073	116	3	2	14	trEMBL	Mascot	128
tr|Q9XT72| Serum iron transport protein transferrin [*Trichosurus vulpecula*]	77715	112	2	2	3	trEMBL	Mascot	131
ens| transcript:ENSMODT00000018754 [*Monodelphis domestica*] (Myeloperoxidase)	45989	101	2	2	5	trEMBL	Mascot	147
tr|Q692U7| Salivary gland protein [*Ixodes scapularis*]	31706	99	25	1	3	trEMBL	Mascot	148
tr|Q4PMB4| 60S acidic ribosomal protein P0 [*Ixodes scapularis*]	34713	94	1	1	3	trEMBL	Mascot	154
tr|Q09JJ3| Ribosomal protein L31 [*Argas monolakensis*]	14942	94	2	2	13	trEMBL	Mascot	155
tr|Q8H712| Peptidyl-prolyl cis-trans isomerase; EC=5.2.1.8 [*Phytophthora infestans*]	18308	93	2	2	7	SwissProt	Mascot	159
tr|Q4PMC0| Ribosomal protein LP2 [*Ixodes scapularis*]	11448	92	1	1	12	SwissProt	Mascot	162
sp|P84251| Cuticle protein 10.9 [*Ixodes ricinus*]	10914	90	6	1	13	SwissProt	Mascot	168
tr|Q4PM52| Putative heat shock-related protein [*Ixodes scapularis*]	19265	66	3	1	3	SwissProt	Mascot	264

Table 4.5A: Proteins identified following equalisation of engorged female *I.holocyclus* protein that have been alkylated with (3-acrylamidopropyl)-trimethylammonium chloride and solubilised in 30% trifluoroethanol in PBS. After equalisation, the proteins were eluted from the Proteominer beads by boiling in SDS sample buffer. The equalised proteins were then separated by 1D SDS-PAGE as shown in figure 4.7.

Accession number/Description/Species	MW	Score	Matches	Unique peptides	Coverage (%)	Database	Engine	List position
sp|P02107| Hemoglobin beta chain [*Macropus rufus*]	15991	2023	2411	27	100	SwissProt	Mascot	1
sp|Q8IT89| Tropomyosin [*Haemaphysalis longicornis*]	32853	1294	104	24	57	SwissProt	Mascot	2
ens| transcript:ENSOANT00000011920[*Ornithorhynchus anatinus*] (Myosin heavy chain)	227963	844	35	10	7	NCBInr	Mascot	4
ens| transcript:ENSMODT00000036182[*Monodelphis domestica*] (Spectrin)	280344	781	32	15	6	NCBInr	Mascot	5
tr|Q4PKE5| Actin 5 [*Aedes aegypti*]	41795	753	65	9	48	trEMBL	Mascot	7
ens| transcript:ENSMODT00000033884[*Monodelphis domestica*] (Spectrin Beta)	119151	727	36	9	11	NCBInr	Mascot	8
tr|Q8T8B2| Beta-tubulin [*Bombyx mori*]	50182	675	20	10	29	trEMBL	Mascot	10
sp|P81043| Hemoglobin alpha chain [*Macropus eugenii*]	15327	593	536	10	73	SwissProt	Mascot	27
sp|Q3THE2| Myosin regulatory light chain MRLC2 [*Mus musculus*]	19767	534	19	8	50	SwissProt	Mascot	36
tr|Q6B8B9| Myosin regulatory light chain [*Ixodes pacificus*]	20002	520	17	10	48	trEMBL	Mascot	38
tr|Q5DW64| Glucose-regulated protein 78kDa [*Oncorhynchus mykiss*]	69793	441	7	5	13	trEMBL	Mascot	45
tr|A8E4K0| Troponin T [*Haemaphysalis qinghaiensis*]	46442	438	23	9	12	trEMBL	Mascot	46
tr|Q4PMB4| 60S acidic ribosomal protein P0 [*Ixodes scapularis*]	34713	400	8	6	32	trEMBL	Mascot	56
ens| transcript:ENSGALT00000019196[*Gallus gallus*] (Histone H2A)	14727	370	80	6	43	NCBInr	Mascot	69
tr|Q64K84| Calreticulin [*Ixodes jellisoni*]	47612	348	17	5	16	trEMBL	Mascot	75
tr|Q6QAZ9| Annexin I [*Gallus gallus*]	38476	337	7	3	8	trEMBL	Mascot	80
tr|B0WAK0| Histone H2A;[*Culex quinquefasciatus*]	16708	332	62	6	29	trEMBL	Mascot	83
sp|P05661| Myosin heavy chain, muscle [*Drosophila melanogaster*]	224328	317	8	5	3	SwissProt	Mascot	87
ens| transcript:AGAP012871-RA[*Anopheles gambiae*] (Histone H2B)	7459	312	21	3	35	NCBInr	Mascot	91
tr|Q4PMD3| 40S ribosomal protein S10 [*Ixodes scapularis*]	18073	301	7	5	31	trEMBL	Mascot	97
ens| transcript:ENSMMUT00000003203[*Macaca mulatta*] (Myeloperoxidase)	87183	289	20	4	6	trEMBL	Mascot	102
sp|Q29616| Transthyretin [*Macropus giganteus*]	16480	283	14	4	31	SwissProt	Mascot	106
tr|Q4PMB9| 40S ribosomal protein S5;[*Ixodes scapularis*]	23271	282	3	3	22	trEMBL	Mascot	107
ens| transcript:ENSMODT00000009759[*Monodelphis domestica*] (Flavin reductase)	23287	234	9	5	21	NCBInr	Mascot	140
tr|B0JZS0| Putative uncharacterized protein;[*Xenopus tropicalis*]	22289	230	11	4	23	trEMBL	Mascot	143
tr|A6NA14| Truncated peroxiredoxin [*Ornithodoros parkeri*]	20956	197	10	4	31	trEMBL	Mascot	171
tr|B2L107| Visinin-like protein 1;[*Mus musculus*]	22110	194	3	3	18	trEMBL	Mascot	178
ens| transcript:ENSMODT00000003516[*Monodelphis domestica*] (Chromatin-modifying protein 4b)	24834	193	4	2	13	NCBInr	Mascot	181
ens| transcript:ENSMLUT00000001219[*Myotis lucifugus*] (nucleosome assembly protein 1-like 4)	42286	188	6	2	8	NCBInr	Mascot	187
tr|A6NA00| 40S ribosomal protein SA;[*Ornithodoros parkeri*]	33045	185	12	3	10	trEMBL	Mascot	196

tr|Q4PM87| Nonmuscle myosin essential light chain;[*Ixodes scapularis*]	16507	182	3	2	16	trEMBL	Mascot	203
tr|O97645| Fibrinogen A-alpha chain [*Macropus rufus*]	43865	175	17	3	8	trEMBL	Mascot	212
tr|Q3TE63| Peptidyl-prolyl cis-trans isomerase; EC=5.2.1.8;[*Mus musculus*]	18001	171	3	3	24	trEMBL	Mascot	219
tr|Q98UH8| Calmodulin [*Clemmys japonica*]	15349	168	2	2	24	trEMBL	Mascot	225
ens| transcript:ENSMODT00000003089[*Monodelphis domestica*] (Haptoglobin)	38478	167	8	5	10	NCBInr	Mascot	226
sp|Q09665| Troponin C, isoform 2;[*Caenorhabditis elegans*]	18216	166	4	2	10	SwissProt	Mascot	227
sp|P15159| Troponin C;[*Tachypleus tridentatus*]	17422	153	5	3	7	SwissProt	Mascot	255
ens| transcript:ENSMODT00000013318[*Monodelphis domestica*] (Ankyrin 1)	181691	151	9	2	2	NCBInr	Mascot	258
tr|Q4PM82| Ribosomal protein S25;[*Ixodes scapularis*]	12386	139	4	2	18	trEMBL	Mascot	280
tr|Q4PM81| Elongation factor 1-beta;[*Ixodes scapularis*]	24368	115	2	2	14	trEMBL	Mascot	346
tr|O97117| Glutathione S-transferase;[*Boophilus microplus*]	25558	102	7	2	13	trEMBL	Mascot	381
tr|Q4PM00| 40S ribosomal protein S30;[*Ixodes scapularis*]	14589	100	5	2	8	trEMBL	Mascot	390

Table 4.5B: Proteins identified following equalisation of engorged female *I.holocyclus* protein solubilised in 30% trifluoroethanol in PBS. After equalisation, the proteins were eluted from the Proteominer beads by boiling in SDS sample buffer. The equalised proteins were then separated by 1D SDS-PAGE as shown in figure 4.7.

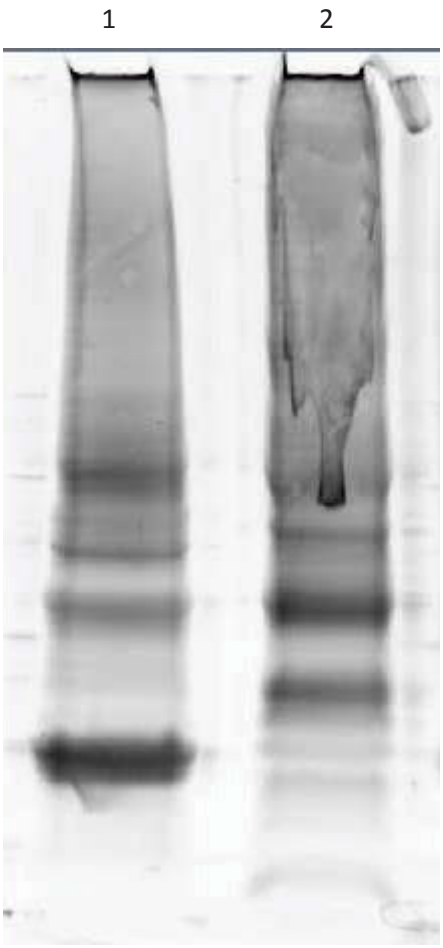

Figure 4.10: SDS-PAGE of Proteominer equalised protein from engorged *I. holocyclus* prepared by solubilising the sample in ammonium chloride.

1) Quaternary ammonium alkylated
2) Not alkylated.

Haemoglobin from the host is spread throughout the entire gel lane, indicating that these proteins are not only saturating their specific hexapeptide beads, but are also non-specifically binding to other beads and proteins. This is in spite of the extensive washing of the beads carried out before elution of the equalised protein by boiling the beads in the presence of SDS. Finally, alkylating with (3-acrylamidopropyl)-trimethylammonium chloride has made little difference to the proteins solubilised when compared to the unalkylated sample.

4.3.6: Fluorescent zymograms of equalised tick protein extracts.

The fluorescent zymogram method presented in chapter three was not successful in positively identifying serine proteases. Aside from the possibility of the assay being performed at a pH not optimum for the enzymes (Miyoshi *et al.*, 2008), a likely reason for not identifying the proteases is that they are present in too low abundance and not able to break down enough of the substrate to be detected. In addition, if the enzyme was in low abundance but had high activity, an apparently intense fluorescent spot may not produce enough peptides for identification using mass spectrometry. The application of the fluorescent zymogram method to engorged ticks presents the problem of overwhelming amounts of host enzymes which will be easily detected while relatively lower abundance tick proteins will not.

Equalisation of a PBS extract of engorged ticks using Proteominer prior to the fluorescent zymogram was used to overcome the above issues. Engorged tick extracts were prepared as described in section 4.2.4 and subjected to the fluorescent zymogram method presented in chapter three. The gels in figure 4.11 and 4.12 show both Flamingo-stained gels and corresponding fluorescent zymograms of PBS

Accession number|Description/Species	MW	Score	Matches	Unique peptides	Coverage (%)	Database	Engine	List position
sp|P02106| Hemoglobin beta chain [*Macropus giganteus*]	16005	1309	580	20	95	SwissProt	Mascot	1
tr|A7SHX4| Predicted protein [*Nematostella vectensis*] (Histone)	24545	865	79	14	42	NCBInr	Mascot	2
sp|O97162| Tropomyosin [*Boophilus microplus*]	32982	575	11	11	28	trEMB.	Mascot	6
sp|P81043| Hemoglobin alpha chain [*Macropus eugenii*]	15327	451	337	7	38	trEMB.	Mascot	10
tr|A6N9R2| Ribosomal protein S18 [*Ornithodoros parkeri*]	17815	437	41	8	36	SwissProt	Mascot	11
sp|Q6R520| Calmodulin [*Oreochromis mossambicus*]	16835	364	5	5	44	NCBInr	Mascot	15
tr|A8E4K0| Troponin T [*Haemaphysalis qinghaiensis*]	46442	358	29	7	9	SwissProt	Mascot	17
tr|A5H1Q6| Ribosomal protein L23 [*Haemaphysalis longicornis*]	15450	335	11	4	35	SwissProt	Mascot	19
ens| transcript:ENSMODT00000010348 [*Monodelphis domestica*] (Histone H2A)	14117	331	49	7	35	NCBInr	Mascot	21
tr|Q4PM63| Histone H2B;[*Ixodes scapularis*]	13811	320	39	7	28	trEMB.	Mascot	22
ens| transcript:ENSMODT00000004396 [*Monodelphis domestica*]	23616	230	4	4	19	trEMB.	Mascot	45
tr|Q4PMS2| Ribosomal protein S25 [*Ixodes scapularis*]	12829	216	5	3	10	trEMB.	Mascot	48
ens| transcript:ENSMODT00000024341 [*Monodelphis domestica*]	68048	188	54	3	5	trEMB.	Mascot	57
tr|O97645| Fibrinogen A-alpha chain [*Macropus rufus*]	43865	188	12	2	8	NCBInr	Mascot	58
tr|Q4PM11| 40S ribosomal protein S13 [*Ixodes scapularis*]	17199	182	4	3	18	trEMB.	Mascot	62
tr|Q4PM30| Ribosomal protein S19 [*Ixodes scapularis*]	16509	180	6	3	14	trEMB.	Mascot	63
tr|Q4PM87| Nonmuscle myosin essential light chain [*Ixodes scapularis*]	16507	180	4	2	16	trEMB.	Mascot	70
sp|O18783| Plasminogen; EC=3.4.21.7 [*Macropus eugenii*]	90922	179	27	3	3	SwissProt	Mascot	71
tr|Q4PM16| 60S ribosomal protein L23 [*Ixodes scapularis*]	14768	165	6	4	24	trEMB.	Mascot	89
tr|Q4PM00| 40S ribosomal protein S30 [*Ixodes scapularis*]	14589	124	8	2	8	trEMB.	Mascot	117
tr|Q4PMD3| 40S ribosomal protein S10 [*Ixodes scapularis*]	18073	116	3	2	14	trEMB.	Mascot	128
tr|Q9XT72| Serum iron transport protein transferrin [*Trichosurus vulpecula*]	77715	112	2	2	3	trEMB.	Mascot	131
ens| transcript:ENSMODT00000018754 [*Monodelphis domestica*] (Myeloperoxidase)	45989	101	2	2	5	trEMB.	Mascot	147
tr|Q592U7| Salivary gland protein [*Ixodes scapularis*]	31706	99	25	1	3	trEMB.	Mascot	148
tr|Q4PMB4| 60S acidic ribosomal protein P0 [*Ixodes scapularis*]	34713	94	1	1	3	trEMB.	Mascot	154
tr|Q09JJ3| Ribosomal protein L31 [*Argas monolakensis*]	14942	94	2	2	13	trEMB.	Mascot	155

tr|Q8H712| Peptidyl-prolyl cis-trans isomerase; EC=5.2.1.8[*Phytophthora infestans*]	18308	93	2	2	7	SwissProt	Mascot	159
tr|Q4PMK0| Ribosomalprotein LP2 [*Ixodes scapularis*]	11448	92	1	1	12	SwissProt	Mascot	152
sp|P84251| Cutide protein 10.9 [*Ixodes ricinus*]	10914	90	6	1	13	SwissProt	Mascot	158
tr|Q4PM52| Putative heat shock-related protein [*Ixodes scapularis*]	19265	66	3	1	3	SwissProt	Mascot	254

Table 4.6A: Proteins identified following equalisation of engorged female *I.holocyclus* protein, following lysis of the red blood cells with ammonium chloride. The protein was then alkylated with (3-acrylamidopropyl)-trimethylammonium chloride and solubilised in 30% trifluoroethanol in PBS. After equalisation, the proteins were eluted from the Proteominer beads by boiling in SDS sample buffer. The equalised proteins were then separated by 1D SDS-PAGE as shown in figure 4.8.

Accession number/Description/Species	MW	Score	Matches	Unique peptides	Coverage (%)	Database	Engine	List position
sp|P02107| Hemoglobin beta chain [Macropus rufus]	15991	1023	2411	27	100	SwissProt	Mascot	1
sp|Q8IT89| Tropomyosin [Haemaphysalis longicornis]	32853	1294	104	24	57	SwissProt	Mascot	2
ens| transcript:ENSOANT00000011910[Orrithorhynchus anatinus] (myosin heavy chain 9)	227963	844	35	10	7	NCBInr	Mascot	4
ens| transcript:ENSMODT00000036132[Monodephis domestica] (Spectrin alpha)	280344	761	32	15	6	NCBInr	Mascot	5
tr|Q4PKE5| Actin 5 [Aedes aegypti]	41795	753	65	9	48	trEMBL	Mascot	7
ens| transcript:ENSMODT00000033884[Monodephis domestica] (Spectrin beta)	119151	727	36	9	11	NCBInr	Mascot	8
tr|Q8T8B2| Beta-tubulin [Bombyx mori]	50182	575	20	10	29	trEMBL	Mascot	10
sp|P81043| Hemoglobin alpha chain [Macropus eugenii]	15327	593	536	10	73	SwissProt	Mascot	27
sp|Q3THE2| Myosin regulatory light chain MRLC2 [Mus musculus]	19767	534	19	8	50	SwissProt	Mascot	36
tr|Q6B889| Myosin regulatory light chain [xodes pacificus]	20002	520	17	10	48	trEMBL	Mascot	38
tr|Q5DW64| Glucose-regulated protein 78kDa [Oncorhynchus mykiss]	69793	441	7	5	13	trEMBL	Mascot	45
tr|A8E4K0| Tropon n T [Haemaphysa is qinghaiensis]	46442	438	23	9	12	trEMBL	Mascot	46
tr|Q4PMB4| 60S acidic ribosomal protein P0 [Ixodes scapularis]	34713	400	5	6	32	trEMBL	Mascot	56
ens| transcript:ENSGALT00000019195[Gallus gallus] (H stone H2A)	14727	370	80	6	43	NCBInr	Mascot	69
tr|Q64K84| Calreticulin [Ixodes jellisoni]	47612	348	17	5	16	trEMBL	Mascot	75
tr|Q5QAZ9| Annexin I [Gallus gallus]	38476	337	7	3	8	trEMBL	Mascot	80
tr|B0WAK0| Histone H2A;[Culex quinquefasciatus]	16708	332	62	6	29	trEMBL	Mascot	83
sp|P05661| Myosin heavy chain, muscle [Drosophila melanogaster]	224328	317	8	5	3	SwissProt	Mascot	87
ens| transcript:AGAP012871-RA[Anopheles gambiae] (Histone H2B)	7459	312	21	3	35	NCBInr	Mascot	91
tr|Q4PMD3| 40S ribosomal protein S10 [Ixodes scapularis]	18073	301	7	5	31	trEMBL	Mascot	97
ens| transcript:ENSMMUT00000003203[Macaca mulatta] (Myeloperoxidase)	87183	289	20	4	6	trEMBL	Mascot	102
sp|Q29616| Transthyretin [Macropus giganteus]	16480	283	14	4	31	SwissProt	Mascot	106
tr|Q4PMB9| 40S ribosomal protein S5;[Ixodes scapularis]	23271	282	8	3	22	trEMBL	Mascot	107
ens| transcript:ENSMODT00000009759[Monodephis domestica] (Flavin reductase)	23287	234	9	5	21	NCBInr	Mascot	140
tr|B0JZS0| Putative uncharacterized protein;[Xenopus tropicalis]	22289	230	11	4	23	trEMBL	Mascot	143
tr|A6NA14| Truncated peroxiredoxin [Ornithodoros parkeri]	20956	197	10	4	31	trEMBL	Mascot	171
tr|B2L107| Visinin-like protein 1;[Mus musculus]	22110	194	8	3	18	trEMBL	Mascot	178
ens| transcript:ENSMODT00000003516[Monodephis domestica] (Chromatin-modifying protein 4b)	24834	193	4	2	13	NCBInr	Mascot	181
ens| transcript:ENSMLUT00000001219[Myotis lucifugus] (nucleosome assembly protein 1-like 4)	42286	188	5	2	8	NCBInr	Mascot	187
tr|A6NA00| 40S ribosomal protein S4;[Ornithodoros parkeri]	33045	185	12	3	10	trEMBL	Mascot	196

tr|Q4PM87| Nonmuscle myosin essential light chain;[Ixodes scapularis]	16507	182	3	2	16	tr:EMBL	Mascot	203
tr|O97645| Fibrinogen A-alpha chain [Macropus rufus]	43865	175	17	3	8	tr:EMBL	Mascot	212
tr|Q3TE63| Peptidyl-prolyl cis-trans isomerase; EC=5.2.1.8;[Mus musculus]	18001	171	3	3	24	tr:EMBL	Mascot	219
tr|Q98UH8| Calmoculin [Clemmys japonica]	15349	163	2	2	24	tr:EMBL	Mascot	225
ens| transcript:ENSMODT00000003089[Monodelphis domestica] (Haptoglobin)	38478	167	8	5	10	NCBInr	Mascot	226
sp|Q09665| Troponin C, isoform 2;[Caenorhabditis elegans]	18216	165	4	2	10	SwissProt	Mascot	227
sp|P15159| Troponin C;[Tachypleus tridentatus]	17422	153	5	3	7	SwissProt	Mascot	255
ens| transcript:ENSMODT00000013313[Monodelphis domestica] (Ankyrin 1)	181691	151	9	2	2	NCBInr	Mascot	258
tr|Q4PM82| Ribosomal protein S25;[Ixodes scapularis]	12386	139	4	2	18	tr:EMBL	Mascot	280
tr|Q4PM81| Elongation factor 1-beta;[Ixodes scapularis]	24368	115	2	2	14	tr:EMBL	Mascot	346
tr|O97117| Glutathione S-transferase;[Boophilus microplus]	25558	102	7	2	13	tr:EMBL	Mascot	381
tr|Q4PM00| 40S ribosomal protein S30;[Ixodes scapularis]	14589	100	5	2	8	tr:EMBL	Mascot	390

Table 4.5B: Proteins identified following equalisation of engorged female I.holocyclus protein following lysis of the red blood cells with ammonium chloride. The protein was then solubilised in 30% trifluoroethanol in PBS. After ecualisation, the proteins were eluted from the Proteominer beads by boiling in SDS sample buffer. The equalised proteins were then separated by 1D SDS-PAGE as shown in figure 4.8.

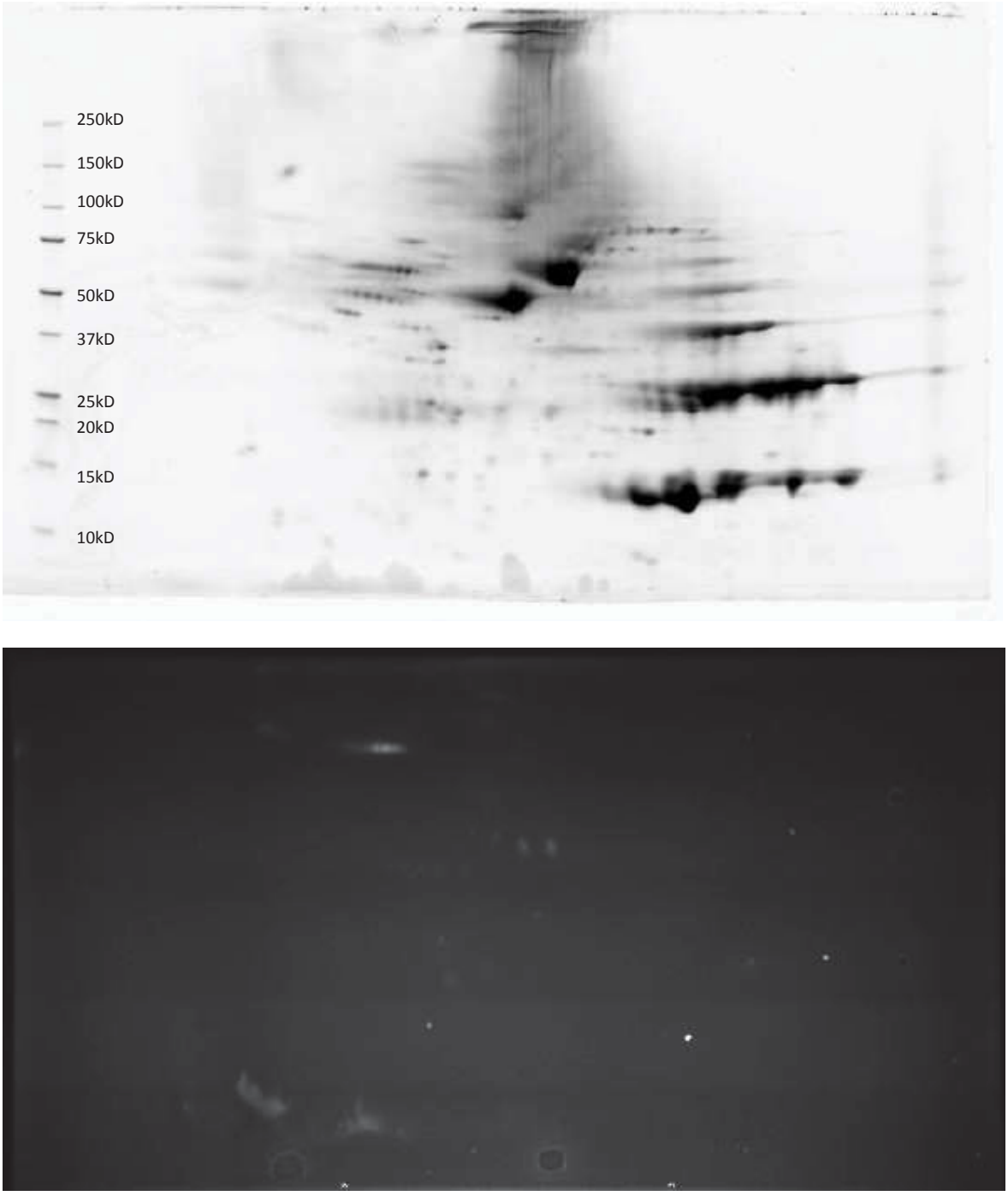

Figure 4.11: PBS extracted protein from fully engorged female *I.holocyclus* separated by isoelectric focusing and SDS-PAGE. Neither of these samples have been reduced or alkylated to allow protein renaturation in the gel. The upper gel (A) shows the sample separated on a 4-12% Criterion XT gel while the lower gel (B) shows the same sample separated on a 10% polyacrylamide gel copolymerised with Boc-Gln-Ala-Arg-MCA. Cleavage of the MCA group causes the MCA to fluoresce when illuminated with UV light.

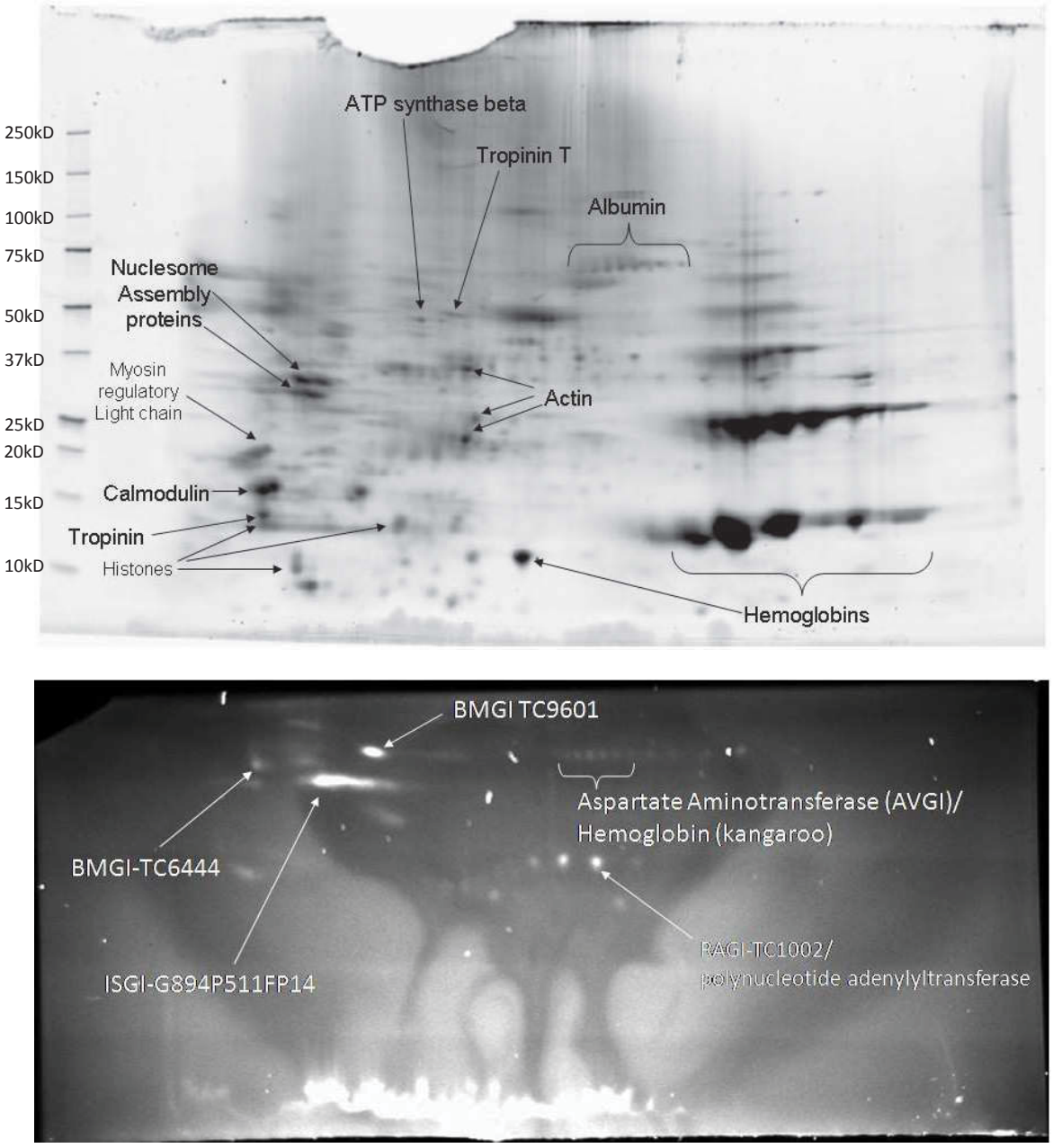

Figure 4.12: Proteominer equalised PBS extracted protein from fully engorged female *I.holocyclus* eluted from the beads with UTC7 then separated by isoelectric focusing and SDS-PAGE. Neither of these samples have been reduced or alkylated to allow protein renaturation in the gel. The upper gel (A) shows the sample separated on a 4-12% Criterion XT gel while the lower gel (B) shows the same sample separated on a 10% polyacrylamide gel copolymerised with Boc-Gln-Ala-Arg-MCA. Cleavage of the MCA group causes the MCA to fluoresce when illuminated with UV light. Proteins spots able to be identified are indicated.

extracted protein from engorged female *I. holocyclus* before (figure 4.11) and after equalisation with Proteominer (figure 4.12). On first inspection of the zymograms, it seems that the zymogram of the equalised sample has simply been exposed for a longer period of time. This is not the case. The ChemiDOC imaging system (Bio-Rad) is set to stop accumulating when a set level of pixel saturation occurs and this occurs within a few seconds when imaging the equalised sample. The exposure time for the unequalised sample is greater than 10 seconds. The equalisation of the sample is apparent when the Flamingo-stained gels are compared. Many protein spots at the acidic end of the gel are now clearly visible in the equalised sample.

Five homologous proteins were able to be identified from excised zymogram spots. The EST TC9601 matches to no homologous proteins when used to search NCBI nr with the BLASTX algorithm. Searching the NCBI genome databases with tBLASTx results in a match to *Chlamydomonas reinhardtii* uridine 5'- monophosphate (UMP) synthase (PYR5) mRNA. However, this match has only 37% identity (18/48. Score = 43.2. Expect = 0.44). UMP synthase catalyses the formation of uridine monophosphate, an energy carrying molecule that is important in many important biosynthetic pathways.

TC6444 when searched with BLASTX matches to NADH-ubiquinone oxidoreductase chain 4 from *R. sanguineus* with 64% identity (255/398. Score = 346. Expect = 4e-93). This protein is a core subunit of the mitochondrial membrane respiratory chain NADH dehydrogenase (Complex I) and is believed to belong to the minimal assembly of the complex required for catalysis. Complex I functions in the transfer of electrons from NADH to the respiratory chain. The SwissProt entry reports that the protein has many potential transmembrane domains, but in this experiment, the homologous spot was found in soluble fraction.

TC1002 matches to a predicted protein from *Nematostella vectensis*, the starlet sea anemone, but with only 29% identity (76/260. Score = 52. Expect = 6e-05). No other information about this protein was available.

The final zymogram spot was found to have homology to an EST in the *A. variegatum* gene indice that is homologous to asparate aminotransferase from *Aedes aegypti*. The function in the mosquito has been inferred by electronic annotation. This enzyme converts asparate and 2-oxogluterate to oxaloacetate and glutamate.

As in chapter three, none of these homologous proteins have the enzymatic activity expected. It is likely that these results have been affected by for the same reason discussed in chapter three, namely the possibility that the assay has been performed at a pH that is not optimal for tick serine proteinases or very low concentrations of enzymes are co-migrating in the gel with higher abundance proteins. A lack of sample stopped the fluorescent gel assays from being repeated at a more acidic pH.

4.4: Conclusions.

The results presented in this chapter show that proteome-wide affinity using hexapeptides can be used as an effective means of depleting highly abundant proteins. The system used, Proteominer, was used to equalise proteins from both unengorged and engorged ticks. However, the equalisation does not seem to be uniform. The work on unengorged ticks (section 4.3.2) shows the spot intensity of acidic proteins has been increased while the spot intensity of many spots has been reduced. Spot number was increased in the equalised sample, but not in the extreme manner reported in published work (Sennels *et al.*, 2007, Simo *et al.*, 2008). This is likely due to insufficient amounts of starting material and the lack of further fractionation by isoelectric means, such as the MCE, to focus the sample on narrow range strips which was shown in chapter two to more than double the number of spots detected. Without access to a laboratory tick colony, it is unlikely that enough ticks would be available to optimally exploit Proteominer for tick proteomics as it has been exploited for other samples, such as human sera (Sennels *et al.*, 2007) and red blood cells (Simo *et al.*, 2008), where availability of starting material is not an issue.

Equalisation of engorged ticks was more successful, with 50% of the spots selected from the 2-D gels of the equalised sample for analysis being tick proteins. In comparison, Simo *et al* (2008) reported 535 unique gene product proteins were found when the authors loaded 1417mg (22.5mL, 67mg/mL) of a red blood cell lysate on a single Proteominer column. One would expect a similar number as the host blood cells are taken up by the tick during feeding and are thus present in the sample. One reason for the discrepancy is the use of only 35-40 engorged ticks for sample preparation, one-third of the 100 unengorged ticks equalised in this chapter. While it is known that a significant number of tick genes are turned on during the rapid engorgement phase of tick feeding, the produced protein may be of too low concentration, and thus the dynamic range too great, to compensate for the reduction in tick number used. The reduced number of engorged ticks was used to minimise the volume of buffer required to solubilise the tick blood meal at a manageable level. The protein concentration

following solubilisation was routinely around 70mg/mL and volumes of 50-250mL were equalised. Thus it is unlikely that overall protein amount is causing the lack of protein identifications when compared to the research of Simo *et al* (2008), but the aforementioned issue of sample dynamic range.

Another issue that has not been considered is the effect on the host protein of being concentrated in the tick gut and partially processed prior to tick collection and freezing. As the tick takes up the blood during feeding, the water contained in the blood meal is returned to the host and the protein concentration rises dramatically. This could cause not only damage to the proteins by digestive processes, but produce non-specific protein-protein interactions which may not be broken when the protein is solubilised in a buffer such as PBS. This may then cause non specific effects during subsequent processing and the interactions with hexapeptides may not be uniform. Protein concentration does not seem to have had an adverse effect in the work of Simo *et al* (2008), and no special treatment of the RBC lysate is mentioned.

Modifications to the sample preparation technique are required to maximise the equalisation. Delipidisation of the liquid nitrogen ground engorged tick powder is necessary and TCA/acetone precipitation has been shown in this work to cause unfolding of some proteins to the point where they will not return to solution and are thus lost prior to the equalisation. This was confirmed by performing a modified sample preparation technique where ground engorged ticks were resuspended and sonicated in a solution containing ammonium chloride, which is used to lyse red blood cells (data not shown). After centrifugation to remove insoluble protein, the sample was precipitated with TCA/acetone and washed with acetone twice to remove the TCA. The result was a white pellet, similar to the pellet that resulted after 10 acetone washes necessary in section 4.3.5. However, this pellet could not be resuspended in PBS even though it had not undergone reduction and alkylation.

With this result, a complete rethinking of the sample preparation needs to be performed, not just for this sample, but for samples where poorly soluble low abundance proteins are those to be equalised. The maximal disruption of non-specific protein-protein interactions while maintaining protein solubility, and thus ability to bind to the hexapeptide library, is the aim. This may be difficult to achieve with the restrictions of the technique with regard to salt and chaotrope concentrations that can be used. One possible solution is fractionation to isolate protein subtypes that are able to bind to the hexapeptides in higher concentrations of reagents that disrupt the non-specific interactions. For example, an experimental workflow may start with resuspending the protein powder in 50mM Tris-HCl, pelleting the insoluble protein and removing the soluble protein for equalisation with one column of equaliser beads. The insoluble protein could then be resuspended in a moderate salt (0.2-0.5M) or chaotrope solution (<3M Urea) to solubilise a subset of proteins in the pellet. The solubilised proteins may then bind to the hexapeptides in a similar way that peptides are retained by interaction with alkyl chains in reversed phase chromatography, "preferring" to be associated with the hexapeptide than remain in the solution. This methodology was not tested in this work due to a lack of engorged ticks. The work of Simo *et al* (2008) on equalising the human red blood cell proteome included 0.5M NaCl in their buffers to reduce haemoglobin interaction with the beads and this did not seem to adversely affect equalisation.

The use of organic solvents which increase protein solubility and promote tertiary structure, such as trifluoroethanol, have not been completely explored in the literature. One methodology that has been explored for other samples in our laboratory, but not investigated for ticks, is the direct solubilisation of the ground engorged tick powder with 30% TFE in PBS. This solution may have enough chaotropic power to disrupt the suspected protein-protein interactions and produce better equalisation.

Chapter Five: Summary, conclusions and the future.

5.1: General conclusions.

The analysis of tick proteomes is not limited by available technology, but how tick researchers apply the available technology. The work of Rachinsky and colleagues (2007a & b) has demonstrated that well designed differential display experiments utilizing two-dimensional gel electrophoresis can be used to identify protein spots that are potential vaccine candidates for both vector and pathogen control. However, their work also shows the significant hurdle of protein identification using traditional MS/MS ion search algorithms such as Mascot and Sequest. This has been the experience in this project. Subjecting CAD generated ion series to *de novo* sequencing to determine peptide sequence has been used with little success. However, as stated numerous times throughout this work, the biggest hurdle to identifying potential vaccine candidates is the removal of high abundance protein that are masking the presence of low abundance proteins. In spite of this, the result of the work presented in chapter two is a list of approximately 200 redundant homologous proteins found in unengorged female *I. holocyclus* with some proteins such as actin and myosin being matched to tick homolog's as well as other insects and parasites. In addition, fractionation of the protein extracted from female unengorged *I. holocyclus* by MCE resolved approximately 700 spots when the MCE fractions were focused on narrow range IPG strips for 2-DGE. In contrast, the individual papers reviewed in section 1.4 report less than 100 homologous proteins identified and only the work of Untalan et al (2005) has a comparable number of resolved spots. The shotgun experiments, especially the 1D SDS-PAGE/1D-LC/MS/MS methodology, identified more than 500 homologous proteins (Table 2.8), although this number has not been corrected for redundancy. The use of targeted approaches, such as Western blots (section 2.3.6) and fluorescent enzyme gels (chapter three) were not as successful as similar studies.

In the work performed on unengorged ticks, by far the greatest contaminating high abundance proteins were cytoskeletal, specifically actin, paramyosin, tropomyosin, the tropinin complex and tubulins. The presence of these proteins in the 1-D SDS-PAGE/1-DLC/MS experiments made the results difficult to compile and interpret due to approximately 75% of the identified proteins listed in the Mascot results files being isoforms of actin from a variety of tick and insect species. This result further reinforces the need to fractionate samples to isolate and remove these proteins. Fractionation by solubility alone is not sufficient but its combination with liquid phase isoelectric fractionation such as the MCE could isolate most cytoskeletal proteins to the pH 5-6.5 chamber. Unfortunately there was insufficient sample to test this. The use of 24cm IPG strips to fractionate by pI is very useful when the amount of protein available is too low for the large volume MCE. Dissection of the strip into more manageable lengths enables 2-D gels to be performed on commercially available second dimension gels and opens the way for researchers without the extensive experience required to perform second dimension large format gels.

The need to remove high abundance tick cytoskeletal proteins is somewhat parallel to the need to remove high abundance proteins from human serum. In addition to the use of Proteominer beads (discussed in chapter four and further below), the removal of albumin and immunoglobulin can be achieved using either antibody based affinity columns, such as Agilent's MARS column, or a combination of sequential chromatographic steps with different sorbents (Tucholska et al., 2009). An affinity column specific for actin and other cytoskeletal proteins could be produced using protein isolated from 2-DGE to make polyclonal sera in rabbits. This would require 50ug of protein for the first injection followed by two boosters of 25ug each (Kevin Broady, personnel communication). This amount of protein could be obtained from 2D gels, either using numerous (10 or more) Criterion sized gel or a high load 24cm IPG and large format second dimension. The chromatographic sorbent approach is outlined in the work of Guerrier and Boschetti (2007 and 2008) and involves screening

different chromatographic resins to find firstly the one that best adsorbs the target protein with the least number of co-adsorbed species and secondly, a limited number of resins that do not interact with the target, but adsorb the co-adsorbed species found with the previous resin. A column can then be produced where sorbents to remove impurities are placed above the target capture resin. Screening of 48 different chromatographic sorbents is performed in a 96 well plate with 10-50ug of crude protein per sorbent, at pH 4 and 8. This approach is likely to be useful for a wide range of samples with a wide dynamic range for which specific antibodies and affinity columns are not available to remove high abundance proteins.

The amount of sample available is an issue that needs to be addressed in all proteomic studies and it can be a double-edged sword. Unlike genomic and transcriptomic projects, which make use of amplification via the polymerase chain reaction (PCR), proteomic projects are restricted by the amount of starting material which cannot be amplified. This can greatly restrict the level of fractionation and analysis that can be carried out and necessitates careful planning of experiments to make best use of valuable material. The double-edged sword appears when the researcher is confronted by a sample with a wide dynamic range of protein concentration and vast amounts of the sample. The temptation is to utilize a single technique in a single instance, loading the absolute maximum amount of protein the technique can tolerate. This can create more problems than it solves. In a device such as the MCE, overloading to enrich for low abundance proteins can cause precipitation of high abundance proteins on membranes or into chambers with a different isoelectric range to the protein's pI. Figure 2.6D depicting the protein precipitated on the pH 5 membrane of the MCE is reflective of this problem. The results produced when applying equalizer technology to highly concentrated tick blood meal (50+ mg/mL) are another example. The results of chapter four suggest that non-specific binding of proteins to other proteins due to the concentration produces a distorted equalization and that diluting the same amount sample to a much lower concentration (1mg/mL) in

a large volume for a longer contact time (up to 12 hours), in conjunction with improved sample preparation, may reduce unwanted protein-protein interactions. This may result in a more uniform equalization and enrichment of low abundance proteins. Once again, this could not be explored due to a lack of sample and technical issues that required addressing, such as the large sample volume (100-1000mL) necessary. These experiments, to be performed in parallel with fluorescent zymography at a range of pHs (4-10), will be carried out after thesis submission, when engorged ticks are obtained. The results will then be submitted to a peer-reviewed journal for publication.

The problem of too much sample also interferes with LC/MS based analysis techniques. The majority of the experiments presented in this thesis used 2-DGE as the separation technique, using 11cm IPG strips and midi sized gels which have a practical loading capacity of approximately 600µg of protein. This value assumes a moderate dynamic range of protein concentration. But, as exemplified by the gels of engorged tick extracts presented in chapter four, high abundance proteins cause problems with resolution due to their abundance and large, smeared spots they produce. However, even individual haemoglobin spots in gels of engorged ticks would not exceed 1µg of total protein in the gel. This is important because the use of capillary sized chromatography columns (typically 75µm ID x 100mm in length) and trapping cartridges restrict the maximum load of peptides to around 1µg (Michrom Bioresources CapTrap guide). The use of 2-DGE means that it is unlikely that this loading limit will be reached when analysing individual spots sequentially. As this work has shown, protein identification from 2-DGE is not dependant on the total protein load; rather the concentration of the least abundant proteins. Successful protein identification depends most on minimizing protein dynamic range in the sample by fractionation or equalization to maximize the intensity of all protein spots present on the gel.

The use of liquid digests and overloaded 1-D SDS-PAGE gels prior to 1-DLC/MS is a different story, but the result is the same. The practical loading capacity of a CapTrap trapping cartridge (Michrom Bioresources) as used on the QSTAR Elite mass spectrometer in this work is 2µg. Attempting to load more peptide than this would likely result in loss of peptide to waste, or in a worse scenario, precipitation which can potentially block the system, when saturation of the CapTrap is reached. Therefore, the loading limit of a complex peptide digest sample to be analysed by 1-DLC/MS is 2µg when using the autosampler/trap column set up employed on the QSTAR Elite system used in this work. This amount may be higher if the sample was directly injected onto the capillary analytical column, but resolution of individual peptides will be compromised as high abundance peptides (from high abundance proteins) elute over several minutes rather than 30 seconds. Sample carryover also becomes an issue and a recent publication by Wang and co-workers (2009) recommends avoiding sample loads greater than 1.5ug on a 75µm x 10cm column, which was the size used in this work. The use of cation exchange to fractionate complex peptide samples in a multi-dimensional chromatography experiment increases the amount of total sample that can be analysed. This is achieved by increasing the number of salt steps, but each salt elution can only contain approximately 2µg of peptides before loading onto reversed phase dimension of the chromatographic separation. To analyse the same proteins in the 600µg loaded onto the 11cm IPG strip, focused and subjected to SDS-PAGE presented in figure 2.4, 300 2µg peptide fractions are necessary. Once again this is best avoided by minimizing protein dynamic range in the sample by fractionation or equalization.

The same logic can be applied to a 1-D-SDS-PAGE gel. In this work, 100µg of protein was loaded onto a gel and, following electrophoresis, the gel sliced up into 10 pieces based on molecular weight of the proteins. Assuming the protein is evenly distributed across the molecular weight range (which it certainly was not), 10µg of protein is present in each slice. If 50% of this protein can be recovered from the gel piece

following trypsin digestion, 5µg of peptide is present and if all was loaded onto the CapTrap, three fifths are likely to be lost and not analysed. When this is combined with the presence of high abundance proteins/peptides, it is not surprising that the majority of the proteins identified from the 100µg loaded were isoforms of actin. Once again, this reinforces the need to fractionate samples, but also the need to understand the limits of the analytical techniques being used. In the work of Breci *et al* (2005), 65µg of a yeast lysate was loaded onto 1-D-SDS-PAGE and cut into 38 slices following electrophoresis and subjected to trypsin digest. 898 individual proteins were identified from the 38 slices with more than 100 proteins identified from individual gel slices. While it is important to note that yeast is a completely sequenced genome with all coding regions translated and ticks are not, the result in yeast would indicate that the number of slices used in the experiment with 100µg of tick protein was maybe too few. Too few slices would increase the concentration of already high abundance proteins and increasing the number of slices decreases the amount of protein analysed in a single LC/MS/MS experiment. Thus lower abundance proteins that may be co-migrating in the gel lane are more likely to be detected and analysed by the mass spectrometer. The experiment was repeated with 200µg of protein, but the gel lane was sliced into 35 2mm pieces, equating to approximately 5.7µg per piece if it is assumed that the proteins are of equal abundance. Analysis of each piece resulted in lists of 50-100 proteins when the data was searched with Mascot against NCBInr, however once redundancy was removed the complete list from the both gel lanes (Tris and UTC7Tris extracts) was approximately 150 proteins. This list may have been longer if the tick entries in the database were as comprehensive as those for yeast.

Once again, high abundance proteins, especially actin and paramyosin, were found in many slices with actin being found in nearly half the gel slices. This result has been seen with other samples in the laboratory and can be interpreted as ‘more is not necessarily better’. However, in this work it is likely that a great many proteins, which are present in the gel lane, are not being identified because their gene coding or

protein sequences are not present in the database. At the time of writing, NCBInr, a comprehensive, non-identical protein database compiled from GenBank CDS translations, Protein Information Resource (PIR), SWISS-PROT, Protein Research Foundation (PRF), and Protein Data Bank (PDB) databases (http://www.matrixscience.com/search_intro.html), contains 1 512 497 nucleotide sequences and 1146 protein sequences for all *Ixodes* ticks. Of the nucleotide sequences, 1 512 000 are from the *I.scapularis* gene sequencing project and are listed as either "SEQUENCING IN PROGRESS, x unordered-pieces" or "genomic scaffold, whole genome shotgun sequence". These sequence are not present in the NCBInr database as the coding regions (CDS) have not been determined, a prerequisite for a sequence's inclusion in the database. Thus, if peptides from a protein are successfully fragmented, but the protein's gene sequence falls into the "unordered pieces" category or the coding sequence remains undetermined, no match can be made. Matches have been made to trEMBL database entries such as "Chromosome undetermined, whole genome shotgun sequence" and unannotated transcripts of complete genes present in NCBInr. The majority of these can be identified by BLAST searching for homologous proteins. This would imply that if a complete gene sequence is present in NCBInr and good quality MS/MS data is generated from relevant peptides, a match would be made by using Mascot to search NCBInr.

Aside from these bioinformatic issues that will cause problems with all the methodologies presented in this work, 1-D SDS-PAGE is an extremely useful fractionation tool in global proteomic experiments for a number of reasons. The technique can be utilized with highly denaturing surfactants, such as SDS, allowing solubilisation and separation of proteins not soluble in isoelectric focusing. SDS has been used as the last step in the sequential extraction sample preparation workflow used in this work. 1-D SDS-PAGE of the SDS extract has shown that additional proteins are solubilised (figure 2.3), however the pattern was very similar to the UTC7Tris extracts and its analysis wasn't pursued further. The other great advantage of SDS-

PAGE is the locking of the protein in the gel itself, which allows multiple washes and chemical treatments to be performed prior to trypsin digestion. When performing sample preparation and digestion purely in liquid, the removal of chemical contaminants that would interfere with enzymatic digestion is difficult. This problem has encouraged the development of tube-gel digestion (Lu and Zhu, 2005), mentioned in chapter two where the protein sample in liquid is added to liquid acrylamide and polymerized resulting in the protein being locked in a polyacrylamide gel as in SDS-PAGE. The gel pieces can now be treated as a gel slice and washed, chemically treated, dehydrated and the protein enzymatically digested. The improvement in protein coverage can be seen by the example of myoglobin (Lu and Zhu, 2005). Solubilisation of myoglobin with 2% SDS followed by tube-gel digestion produces twice the number of peptides when compared to in solution digestion and tube-gel digestion without SDS. The main drawback of SDS-PAGE in its traditional vertical slab gel format is loading capacity. There is only a finite amount of physical space between the plates supporting a gel and thus overloading a gel with a high abundance protein causes it to occupy a longer and wider amount of space than a low abundance protein, which migrates in a tight band. This is another reason for complex samples to be fractionated or equalized prior to any analysis. The Prep Cell is less sensitive to this effect as the cross-sectional area of the gel tube is approximately 50-fold greater than a vertical slab gels area of 1 x 5mm, depending on well width.

As stated previously, all of the presented results demonstrate the need to simplify the proteome by either fractionation or equalization. The choice of whether to fractionate or equalize will be determined by the aim of the experiment. If quantitation of proteins between different samples is the aim, such as comparing up and down regulation of proteins in pathogen-infected ticks and non infected ticks (Rachinsky *et al.*, 2007a, Rachinsky *et al.*, 2007b), then fractionation is the only choice as equalization removes the ability to quantitate. Equalisation is useful to ascertain the presence of low abundance proteins. Once identified from equalized samples, low abundance proteins

may be quantitated in unequalised samples using the highly sensitive mass spectrometry technique of multiple reaction monitoring (MRM). This targets specific peptides at detection limits in the attomole range. When combined with a heavy isotope labeled version of the peptide, quantitative information can be obtained (Anderson and Hunter, 2005).

The issue of enriching for low abundance tick proteins was addressed using proteome-wide affinity with a combinatorial hexapeptide library, commercialized by Bio-Rad as Proteominer. The data presented in chapter four shows that Proteominer is able to both equalize the sample as judged by spot intensity and the number of spots present on the gel. However, this work did not manage to achieve the extreme increases in protein numbers seen in other published work, such as the 535 unique gene products in equalized human red blood cells in the work of Simo *et al* (2008) or the 1559 unique proteins found in human sera by Sennels *et al* (2007). The fewer protein identifications found in both unengorged and engorged ticks are not likely to be only related to the amount of protein subjected to equalization. Better results are likely when issues with preparation of the engorged tick sample are dealt with. The disruption of hemoglobin and other host blood proteins, judged by removal of the red colour of the sample, was never properly completed even with more than 20 washes with 50mL of ice-cold acetone accompanied by sonication. The repeated precipitation and prolonged time in acetone resulted in many proteins not returning to solution when attempting to solubilise the protein in PBS or other buffers compatible with Proteominer.

Engorged ticks represent an extreme form of a mixed proteome available where more than 75-80% of the sample is host blood that has been concentrated in the gut of the tick and almost all water returned to the host. Cells are also taken up during feeding, mainly red blood cells resulting in the majority of the sample being haemoglobin (Rechav *et al.*, 1994). With this in mind, the finding of many tick proteins in the equalised sample not visualised in the non-equalised sample is very encouraging.

Improved sample preparation with the use of trifluoroethanol as a solubilising reagent, combined with the use of sequential elution reagents such as high and low pH and the utilisation of 1-D-SDS-PAGE/1-DLC/MS as well as 2-DGE mean it is likely that more low abundance proteins from both the tick and the host response to feeding can be discovered. Interestingly, holocyclotoxin was not found in the selected spots analysed, but this could be due the protein being bound to high abundance proteins that were subsequently removed during washing of the Proteominer beads. This further supports the need to optimise sample preparation techniques that disrupt non-specific protein-protein interactions but maintain protein solubility and secondary/tertiary structure.

Equalisation of samples using Proteominer is a technique that requires a larger amount of sample than many other techniques. The amount of sample required will be most dependent on the dynamic range of the sample. To maximise the effectiveness of the equalisation, enough sample needs to be loaded so the lowest abundant protein saturates its hexapeptide bead. If this can be achieved, then intensity of the spots on a 2-D gel will be uniform. At the time of writing there is no published information on the amount of protein bound by each hexapeptide bead and thus the amount of starting material needs to be experimentally determined.

The final main constraint to successful tick proteomic projects is the lack of comprehensive databases to search the tandem mass spectrometry data generated as described earlier in this chapter. The completion of the *I. scapularis* genome-sequencing project will be welcomed, but it will be of little use without meaningful annotation. It was intended during this project to use the fully sequenced *Drosophila* genome to search data, but the lack of annotation of the entries made its use irrelevant. Consequently, all data was searched against the largest database possible, NCBInr. This resulted in many spots being identified with homologous proteins from related species, such as insects and crustaceans. Unmatched 2-DGE spots were then searched against the combined available tick EST databases which provided additional

identifications if the ESTs were annotated or matched to homologous genes or proteins with BLAST.

Even with the available genome, EST and protein sequences available, many intense spots focused on 2-D gels were not found to be homologous to any database entries following trypsin digestion, tandem mass spectrometry and MS/MS ion searching with Mascot or Peaks ProteinID algorithms. The only recourse is then to *de novo* sequence the MS/MS ion spectra to determine peptide sequence. Successful *de novo* sequencing is reliant on a number of factors, the first and foremost is good quality MS/MS spectra, higher than that required for database searching (Yergey *et al*., 2002). High quality MS/MS spectra is reliant on intense parent ions of at least a few hundred counts allowing the acquisition of useful amounts of fragment ions to accurately and confidently determine the peptide sequence. The LC/MS/MS strategy employed in this work and the protein available in some 2-DGE spots were not ideal for *de novo* sequence determination. The Quadrupole/Time of Flight (QTOF) instrument used has superior mass accuracy when compared to ion trap instruments, which allows more confident *de novo* sequence determination. However the use of nanoelectrospray and nanoflow chromatography imposes the constraint of a chromatographic timescale. This requires that chromatography needs to be optimised so that peptide peak widths are made as narrow as possible to increase the amount and therefore intensity of precursor ions selected for MS/MS. However, this increase in ion intensity can be negated if ion selection and fragmentation occurs on the shoulders of the chromatographic peaks (Chen et al., 2005). In addition, the more sample complexity increases, the chance of co-eluting peptides increases. If the number of co-eluting peptides is greater than MS/MS cycle time, that is the time it takes to select a precursor, fragment it and collect MS/MS spectra, low abundance peptides will not analysed (Chen *et al*, 2005). This scenario is of greater importance when using multidimensional chromatography on complex samples rather than for the analysis of 2-D gel spots or proteins purified to homogeneity prior to trypsin digestion. As

described at the beginning of chapter two, when working with the proteomes of organisms with little sequence data it is advisable to keep the proteins intact until they can be isolated with a high level of purity. This is important if *de novo* determination of peptide sequence is necessary to identify a protein of interest, such as one that reacts with immune sera on a Western blot. *De novo* sequencing of MS/MS spectra is difficult enough without adding the problem of which peptides are from which protein in a complex mixture.

MALDI (Matrix Assisted Laser Desorption Ionisation) is not constrained by a chromatographic timescale and maybe a more useful technique for generating data for *de novo* sequencing. MALDI has a higher tolerance for impurities and salt than ESI and the sample is reuseable, with as many laser shots as necessary can be performed to generate more intense MS/MS ions and higher quality spectra (Zhang *et al*, 2003). The coupling of MALDI to QTOF or TOF/TOF analysers means intense, high mass accuracy ions can be generated and confident peptide sequences determined. However, MALDI does have its own unique issues, the main one being the underrepresentation of lysine terminating peptides (Krause et al., 1999). This issue has been solved by chemical derivatisation of the peptides during desalting and concentration of the sample with ZipTips (Joss et al., 2006). The interfacing of nanoflow chromatography with accurate MALDI spotters means that peptide digests, whether derivatised or not, can be separated by a chromatographic timescale, but the acquisition of MS/MS spectra is decoupled from the chromatographic timescale (Andreev *et al.*, 2003, Pan *et al.*, 2005). This decoupling allows greater optimisation of collision energy, producing optimal fragmentation and thus better quality MS/MS spectra (Bodnar et al., 2003). A number of publications have shown the complementarity of LC-MALDI/MS/MS with nanoESI/MS/MS (Bienvenut *et al.*, 2002, Krutchinsky *et al.*, 2000, Wattenberg *et al.*, 2002) and samples analysed by both methods have improved proteome coverage at the protein and peptide level (Bodnar *et al*, 2003). However, the use of MALDI as an ionisation method still requires fractionation of the sample to reduce its complexity

and generate meaningful *de novo* peptide sequence. The laboratory is getting access to both a MALDI spotter and MALDI equipped QTOF, but not in time to acquire data for this project.

The efficacy of MALDI for the generation of good quality MS/MS spectra for de novo peptide sequencing has been shown by the work of Joss *et al* (2006 & 2007) characterising the proteome of Tammar wallaby milk. This marsupial has 34% amino acid sequence homology to humans and 45-70% sequence homology with other non primate eutherian mammals (Margulies *et al*., 2005), a similar bioinformatic situation to *I. holocyclus*. From a 2-D gel of separated milk protein, 143 spots were cut and only 31 spots were identified by database searching with Mascot. Eighty-one protein spots were identified by *de novo* sequencing of MALDI MS/MS spectra. The remaining 31 were unable to be matched to homologous proteins although good quality spectra and sequence were obtained. A number of these proteins may be unique to the wallaby and are thought to provide innate immunological protection to the vulnerable young.

Another factor determining the success of *de novo* sequencing is the data analysis itself. *De novo* sequencing by MS/MS spectra is likely to become the "gold standard" for database searching (Wielsch *et al.*, 2006, Zubarev *et al.*, 2008) but the software necessary is still being developed. Matthiesen (2007) has reported five software tools are available, but evaluation of them all was beyond the scope of this project. Thus it was decided to use one *de novo* sequencing package, Peaks Studio from Bioinformatics Solutions. In this project, Peaks was found to be of limited use in the determination of peptide sequence. Poor quality MS/MS spectra was one reason for this, due to the reasons outlined above. However, an additional reason was the transformation of the raw data from the mass spectrometer into peak lists before the *de novo* sequencing algorithm was applied. This transformation turns all of the peaks into a single mass, ignoring area under the peaks, but it can also turn noise spikes into "peaks" that are then included in the analysis. This finding meant that peptide sequence determined by

Peaks had to be manually validated within Analyst, the data acquisition software for the QSTAR Elite. This is a time consuming process and led to sequences being rejected due to “peaks” being noise spikes. This is not a problem restricted to Peaks, but to any peptide sequencing at low femtomole levels (Shevchenko *et al*, 2006).

Following the determination of peptide sequence from good quality MS/MS spectra, the peptide sequence is searched against databases with homology based search tools. However, the BLAST searching of individual sequences from a single spot against NCBInr often resulted in matches to completely different proteins. This result was perplexing as it was thought that individual peptides from the same protein spot would match to proteins with similar biological function or from the same species. Han *et al* (2005), one of the creators of the Peaks software, has reported that *de novo* sequencing software almost always produces peptide sequences with errors, mainly in parts of the sequence where a sequence of ions is missing from the MS/MS spectra. The algorithm then tries to guess the sequence and composition of the amino acids in the peptide and this is where the error can lie. In this work, no peptide sequences determined by *de novo* sequence determination were deemed to be of sufficient quality of confidence to be included in the results presented.

5.2: Vaccine candidates.

As stated earlier in this chapter, one of aims of this project, along with developing and applying techniques and technology to tick proteomics, was to discover proteins that are possible vaccine candidates. This part of the project was hampered by cross reactivity of anti-dog secondary antibodies with Western blotted, 2-DGE separated tick proteins. More success was obtained probing Western blots with hyperallergic human serum and the homologues of a number of previously described allergens in other organisms were found in *I. holocyclus* using this methodology. This tool was a crucial validation technique for proteins analysed and characterised by the other techniques employed in this work, as it would indicate whether or not a particular protein was capable of producing an immune response. This is the most important feature of a vaccine candidate.

Ten proteins were identified in the work presented in chapter two that have promise as either potential vaccine candidates or drug targets. A number, such as paramyosin, calreticulin, and tropomyosin, have been suggested as vaccine candidates in other organisms such as *Schistosoma*. Others proteins, such as the peroxiredoxins, have been proposed as drug targets while the protein disulphide isomerise transcript has been shown to be a very effective target for RNA interference. It must be noted that, in the list of all the proteins characterised from unengorged female ticks, no anticoagulants, anti-platelet factors or vasodilators were found. This is most likely due to the use of whole ticks rather than dissected tick salivary glands and other organs.

5.3: Future Directions.

There are a great number of technical obstacles to successful isolation and characterisation of proteins from a poorly characterised organism such as *I. holocyclus*. These obstacles and possible solutions have been discussed in the previous section. However, many of the solutions to these problems will come, not only from our laboratory and its continued research on the *I. holocyclus* proteome, but from other researchers and companies researching a broader range of proteomic and bioinformatic problems. Thus, this section will focus on the future experimental work necessary to further the knowledgebase for *I. holocyclus*.

The work presented in this project represents another instalment in the analysis of the Australian Paralysis Tick, *Ixodes holocyclus*. In our laboratory, the study of *I. holocyclus* was started in the early 1990's with the work of Thurn (1994) who isolated and characterised holocyclotoxin, the small protein responsible for paralysis and death in the host animal. The characterisation of the holocyclotoxin protein was followed by the isolation and characterisation of the holocyclotoxin gene sequence by Masina (1999). Determination of the gene sequence allowed the production of recombinant proteins that were subsequently used in animals to ascertain whether a protective immune response could be generated against the original protein. This is the classical workflow used to characterise proteins that have potential as vaccine candidates against important pathogens. However, after the 10 year cycle of that work and the production of many fusion protein variants, no recombinant proteins were produced that could provide a protective immune response to holocyclotoxin.

This work returns to the beginning of the cycle described in the previous paragraph, the isolation and characterisation of proteins that have the potential to be vaccines to prevent ticks attaching and feeding. To achieve this aim, a number of fractionation and separation techniques have been evaluated and optimised. The next step for this project is the refinement of those techniques to isolate proteins of interest as possible vaccine candidates. This will mainly consist of dealing with the issues of background in the Western blots and the stated sample preparation issues for Proteominer. The

biggest hurdle to identification is the bioinformatic resources available and the need to determine peptide sequences using *de novo* sequencing of MS/MS spectra. This is not a trivial exercise and requires a great amount of experience and patience due to the time consuming nature of the analysis.

Once sufficient peptide sequence is determined, the focus of the work turns away from proteomics and to molecular biological techniques, in same way the project moved historically from Thurn's work to Masina's work. In experienced hands, the isolation and determination of the gene sequences of interest should be a straightforward exercise. Once determined, the gene sequence can be used to search nucleotide sequence databases for homologous genes whose function has been annotated. It is then possible that the homologous gene may have been used as a vaccine candidate, which would then help decide if the possible candidate is worth pursuing further. The third stage of the process described in the first paragraph is the one that often requires the most effort, the production, purification and testing of recombinant proteins for their suitability as vaccine candidates. It was during this stage that the work on holocyclotoxin was halted after numerous fusion and non-fusion protein variants were unable to produce a protective immune response in susceptible laboratory animals. This was most likely due to incorrect folding of the recombinant protein. The successful production of an immune response by the recombinant protein is the end point of the project from a basic research viewpoint. The next stage is commercialisation of the vaccine. This would be best carried out by licensing the vaccine to a company with previous experience and success in this area.

Appendix

The following gel images have been annotated to show where protein spots were excised from the gel for further analysis. The figure numbers are the same as the relevant figures in the data chapters. The numbering of the spots is reflective of the spot numbers presented in the table of homologous protein identifications presented in the data chapters.

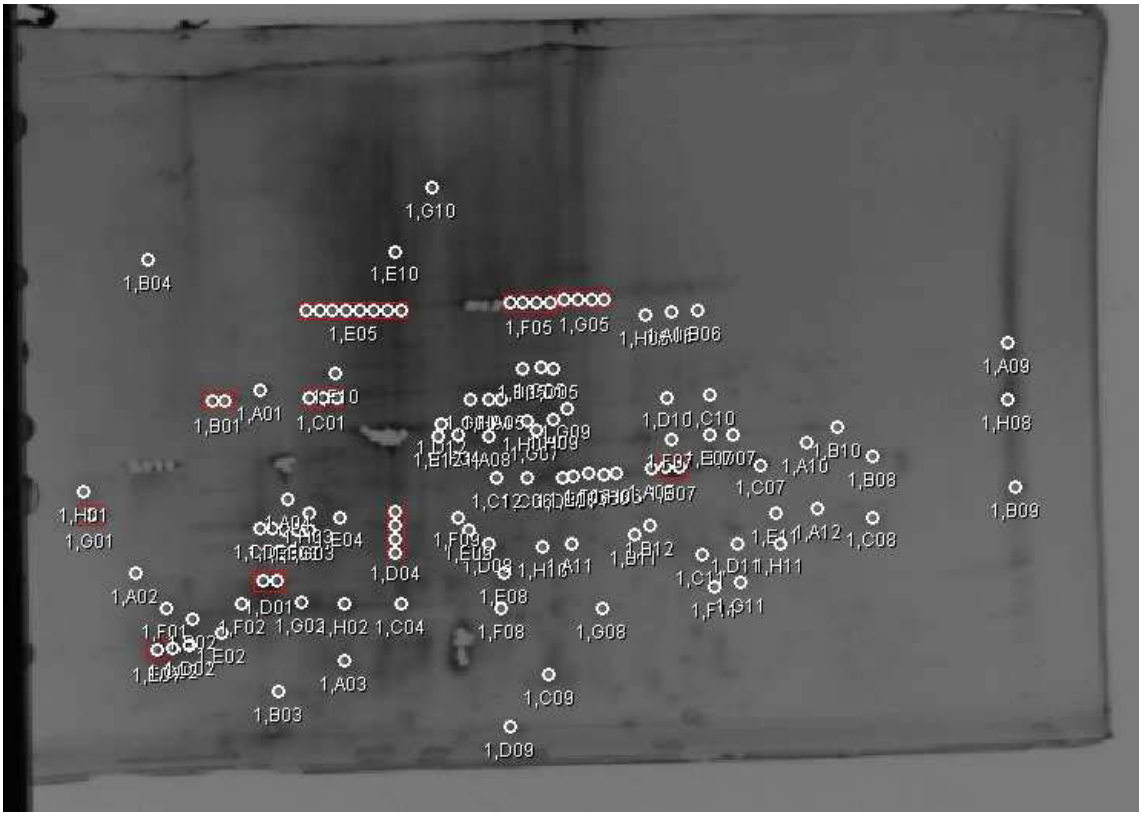

Figure 2.5A: 2-D gel of Tris-extracted protein from unengorged female *I. holocyclus*.

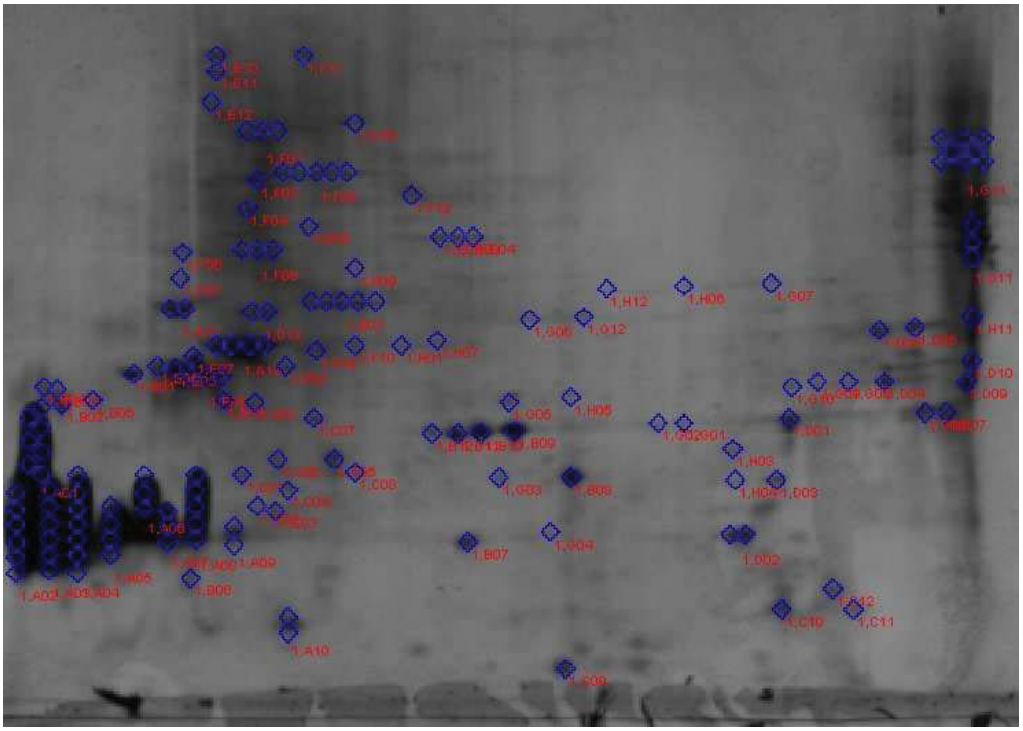

Figure 2.5B: 2-D gel of UTC7Tris-extracted protein from unengorged female *I. holocyclus*.

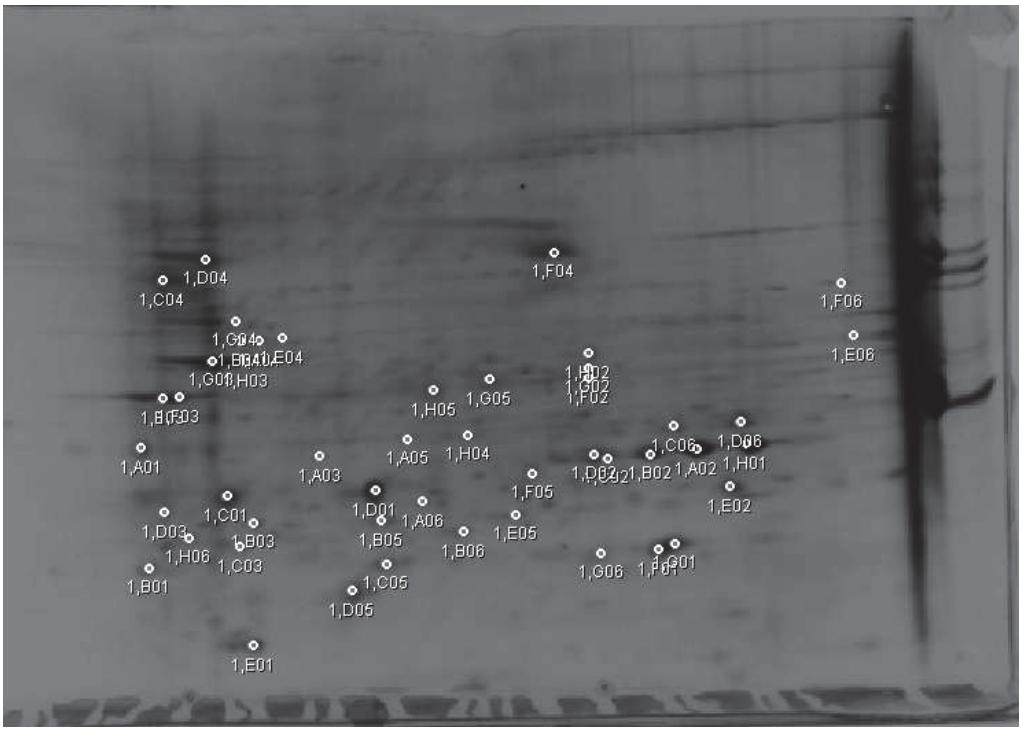

Figure 2.6B: 2-D gel of UTC7Tris-extracted protein from unengorged female *I. holocyclus* fractionated with an MCE. The protein in the pH 5-8 chamber has been loaded onto this gel.

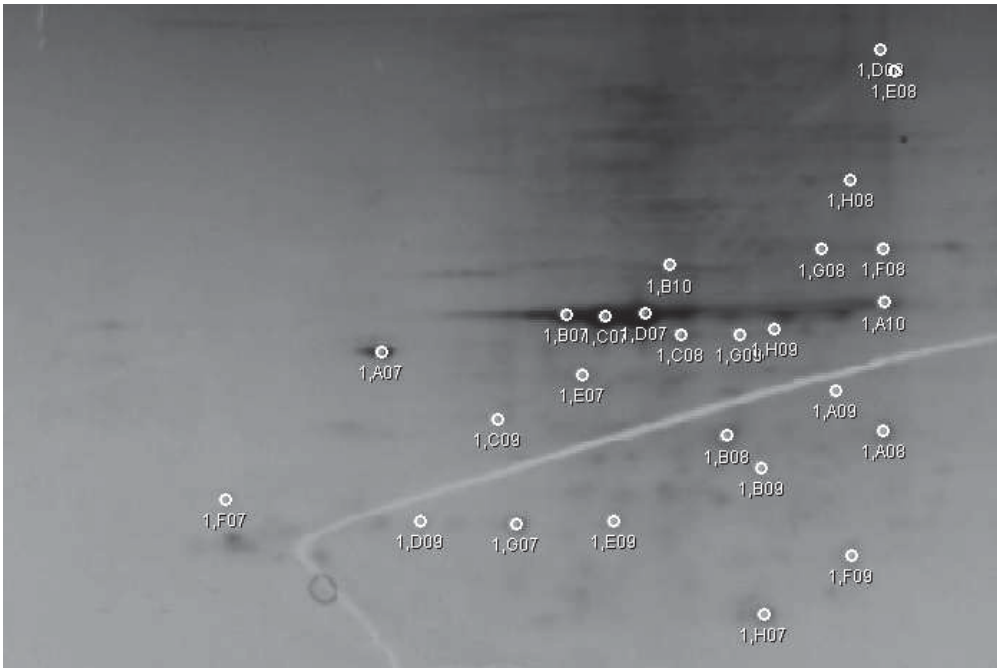

Figure 2.6D: 2-D gel of UTC7Tris-extracted protein from unengorged female *I. holocyclus* fractionated with an MCE. The protein precipitated on the pH 5 membrane has been loaded onto this gel.

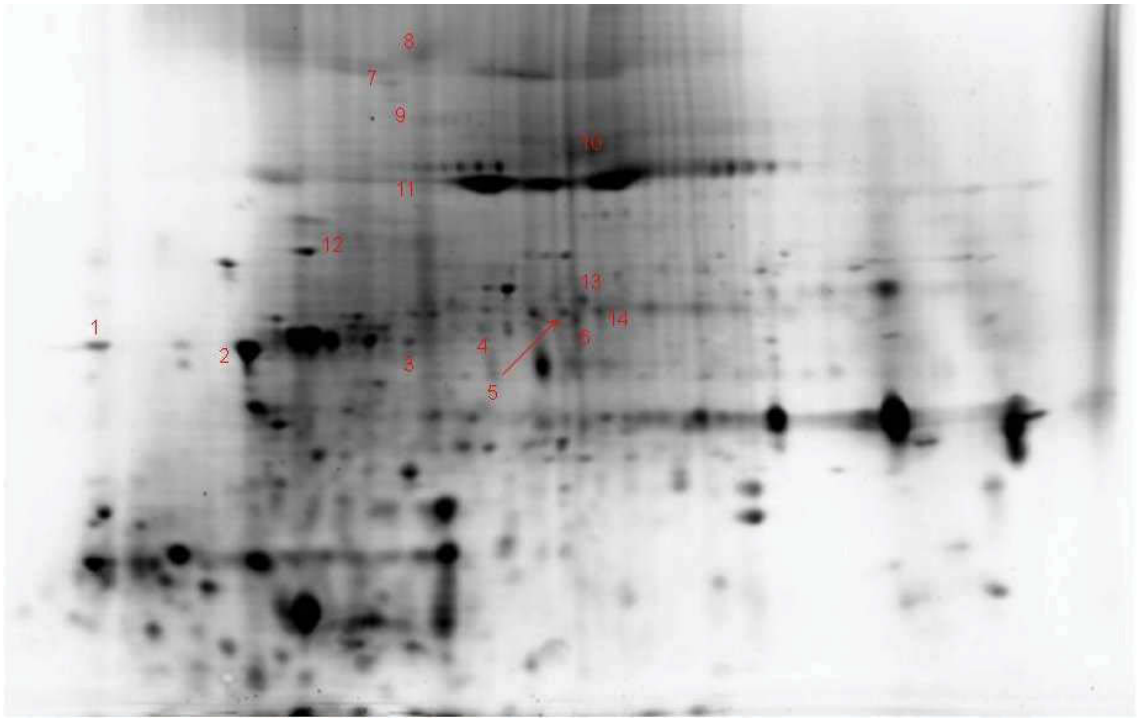

Figure 2.11D: 2-D SDS-PAGE of Tris-extracted protein from unengorged male *I.holocyclus*. Spots matching those visualised in Western blot probed with hyperallergic human serum.

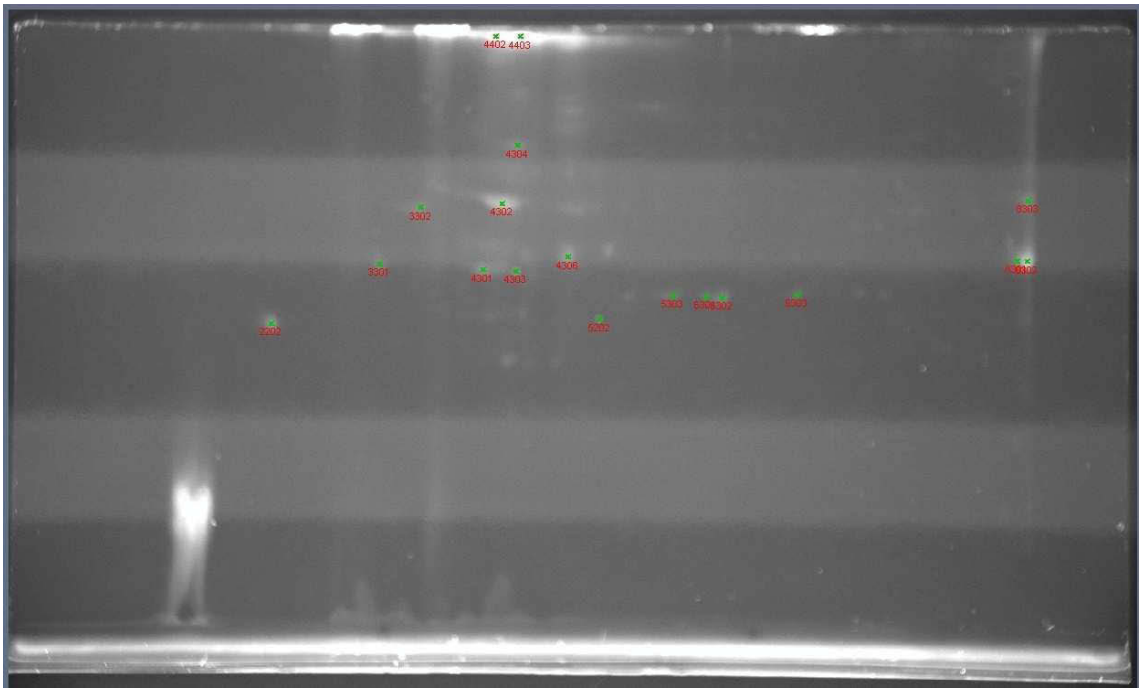

Figure 3.3: 2-D fluorescent zymogram gels of Tris extracted protein from *I. holocyclus*

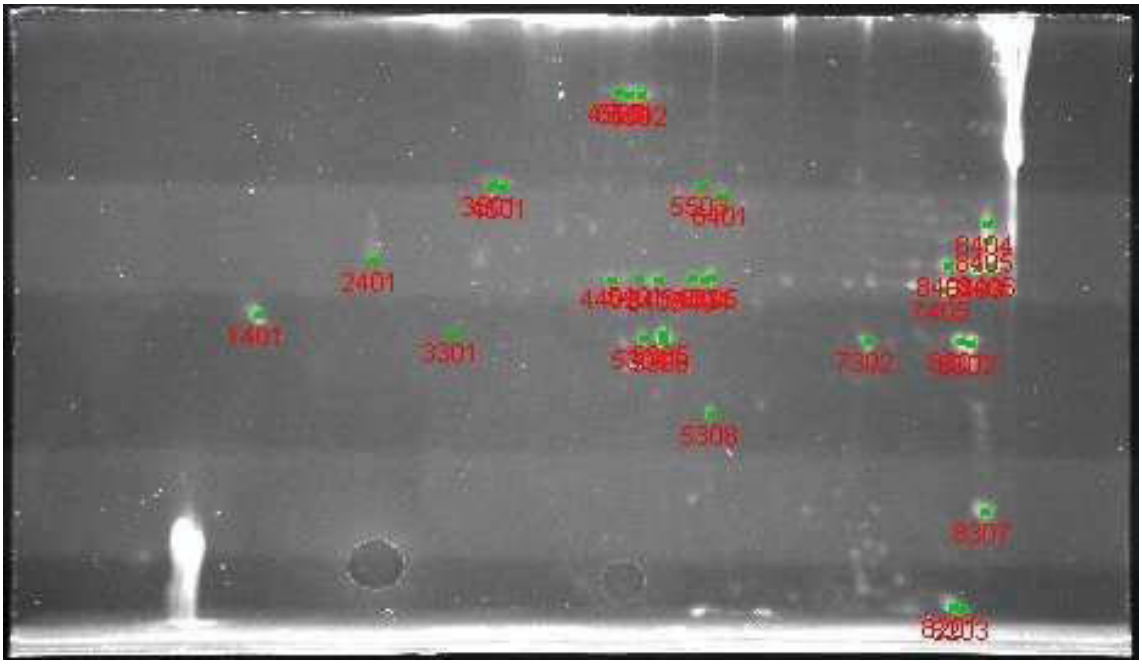

Figure 3.3: 2-D fluorescent zymogram gels of UTC7Tris extracted protein from *I. holocyclus*

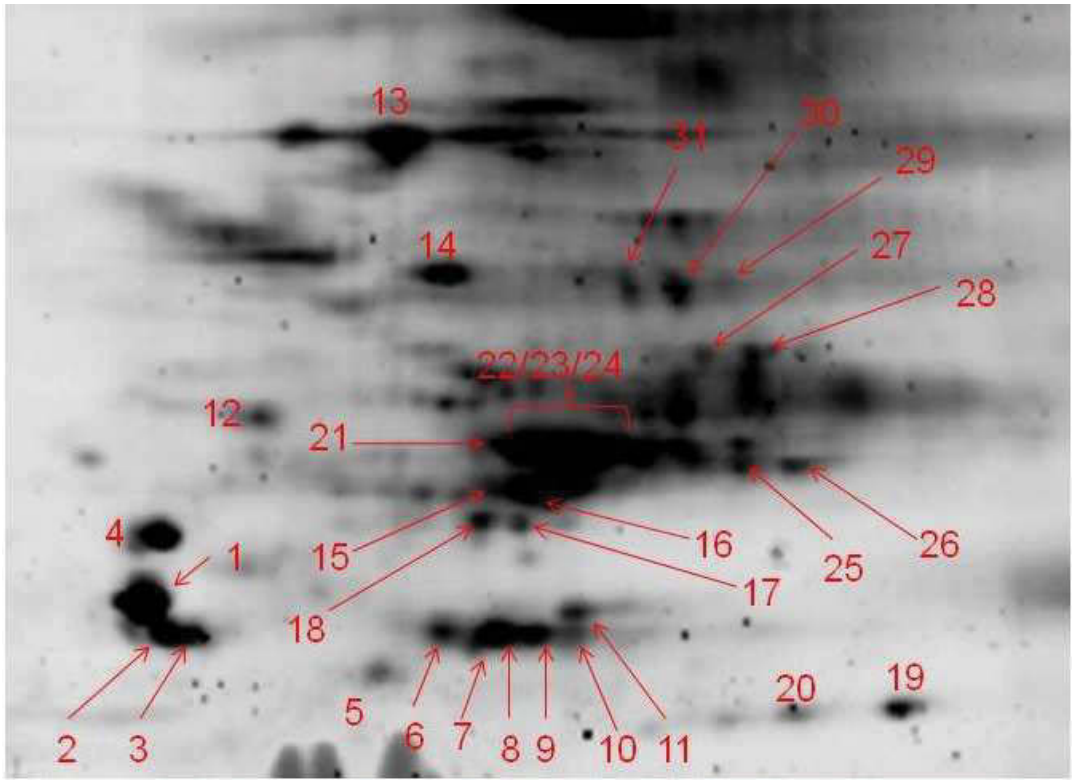

Figure 4.3A: PBS extracted protein from engorged *I. holocyclus* equalised with Proteominer and eluted with UTC7Tris-HCl pH 8.8

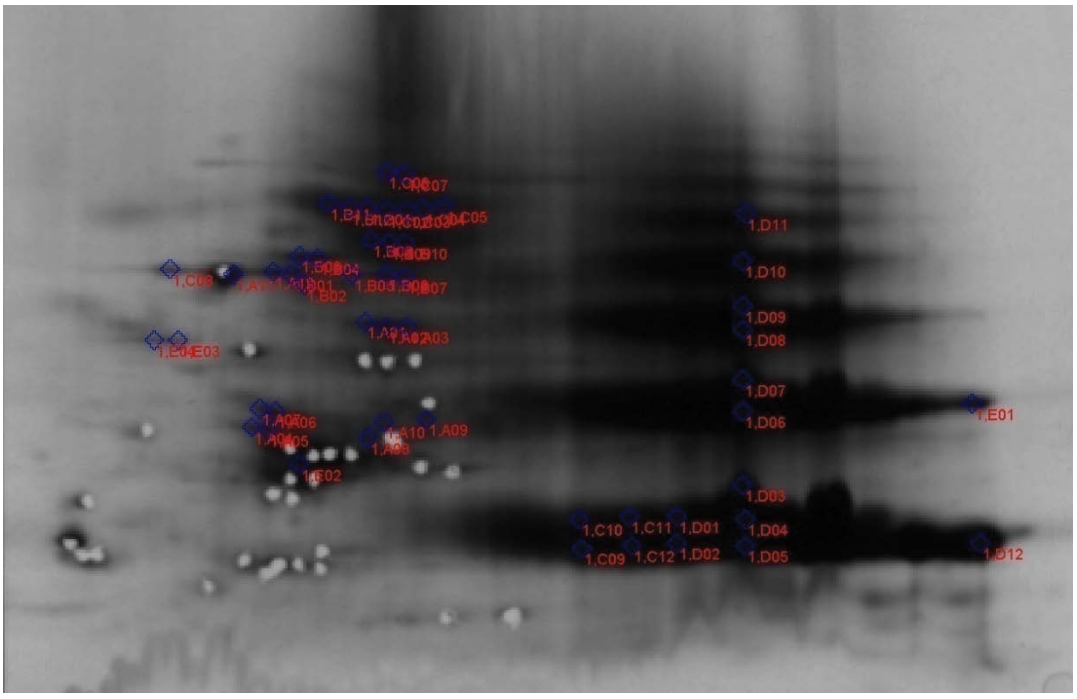

Figure 4.3A: PBS extracted protein from engorged *I. holocyclus* equalised with Proteominer and eluted with UTC7Tris-HCl pH 8.8

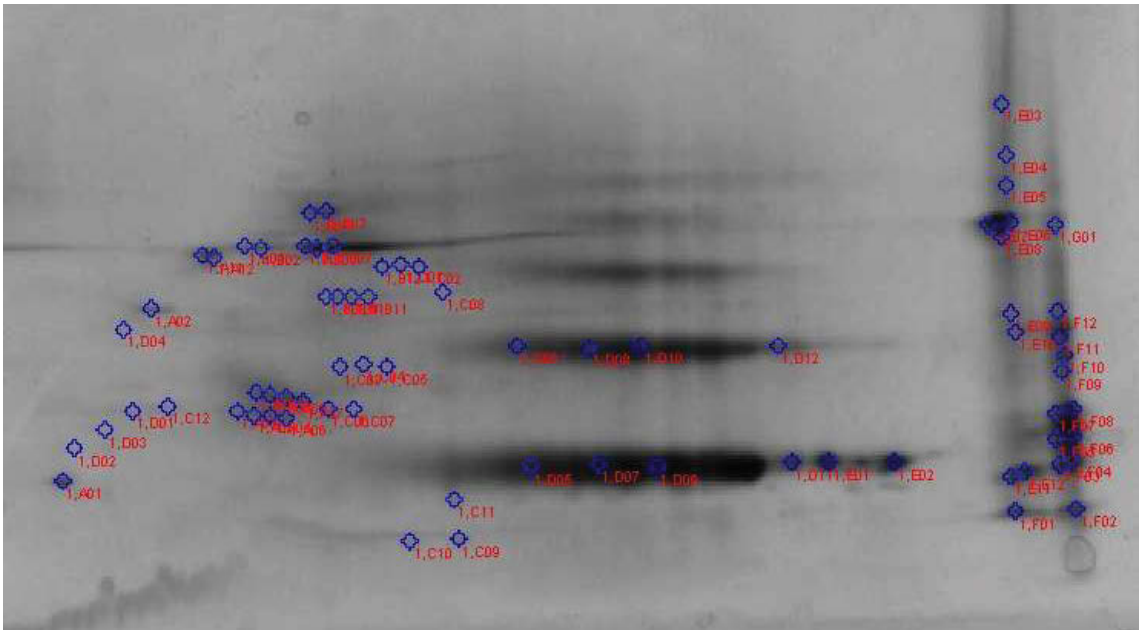

Figure 4.3B: PBS extracted protein from engorged *I.holocyclus* equalised with Proteominer and eluted with UTC7Citric acid pH 4.5

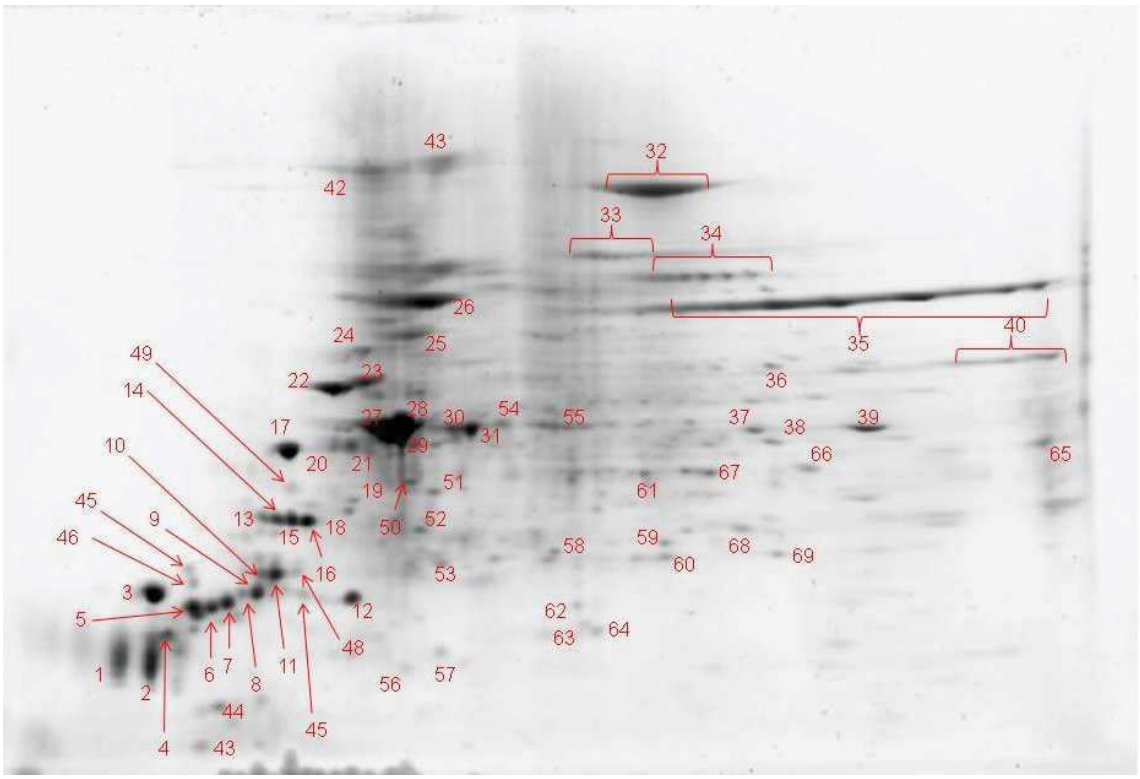

Figure 4.5B: Proteominer equalised, PBS extracted unengorged tick.

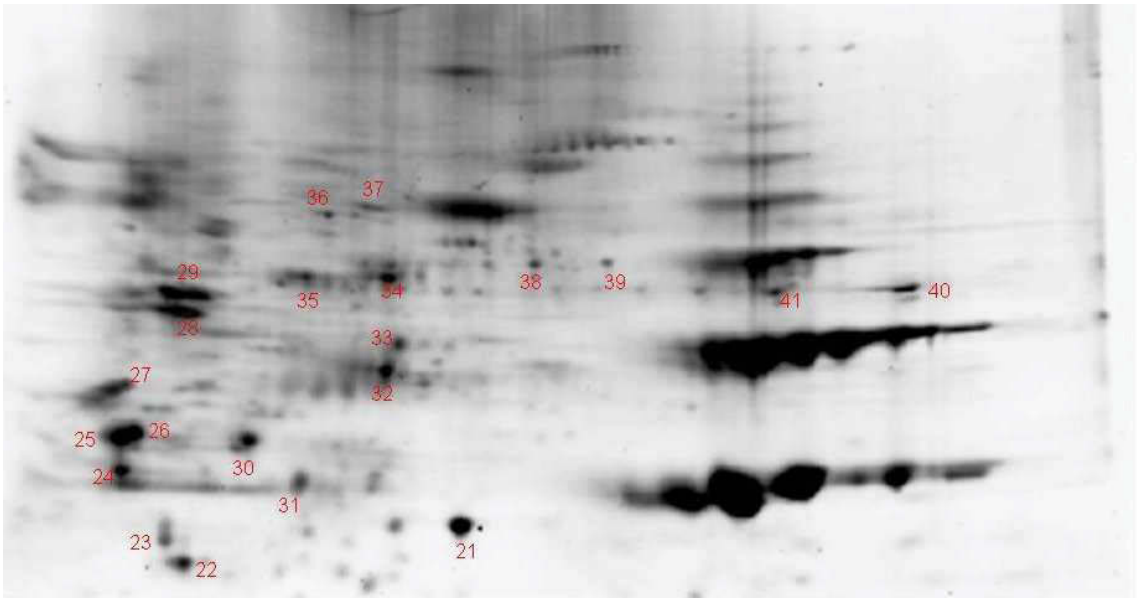

Figure 4.7C: PBS extracted protein from engorged female *I. holocyclus* equalised with Proteominer and eluted with UTC7. Not reduced and alkylated prior to IEF. This sample was also used in the gel presented in figure 4.11A.

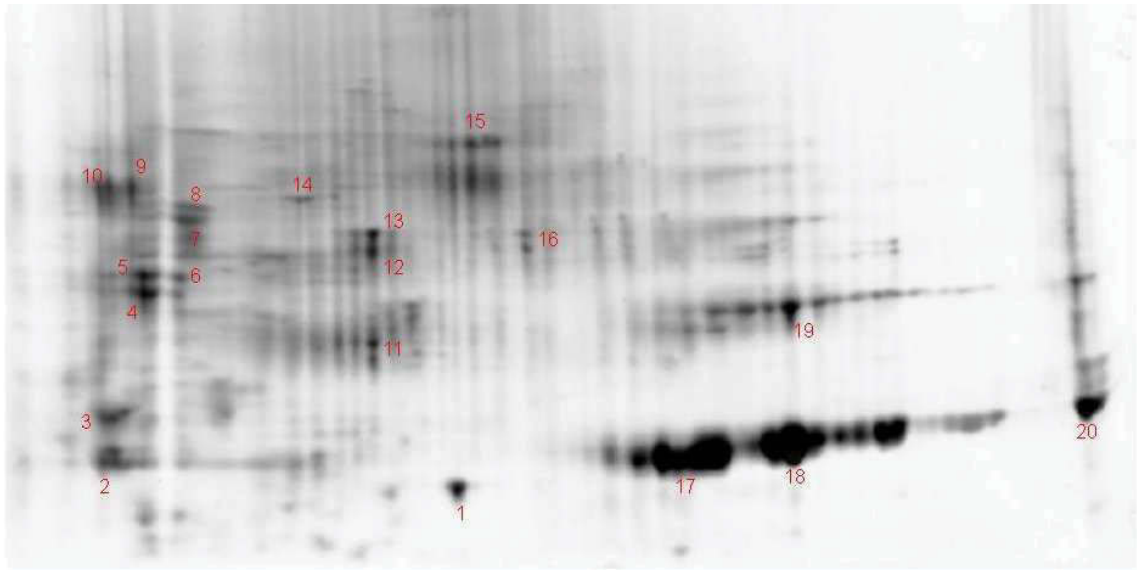

Figure 4.7C: PBS extracted protein from engorged female *I. holocylcus* equalised with Proteominer and eluted with UTC7. Reduced and alkylated prior to IEF.

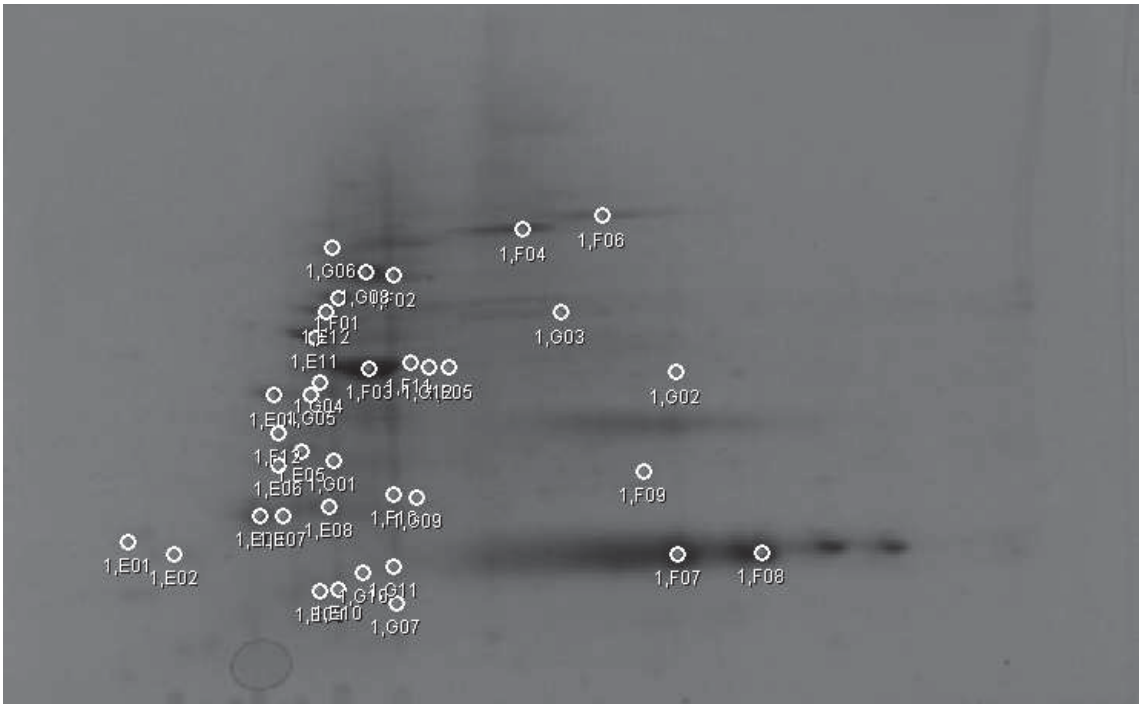

Figure 4.8A: PBS extracted protein from semi-engorged female *I. holocylcus*.

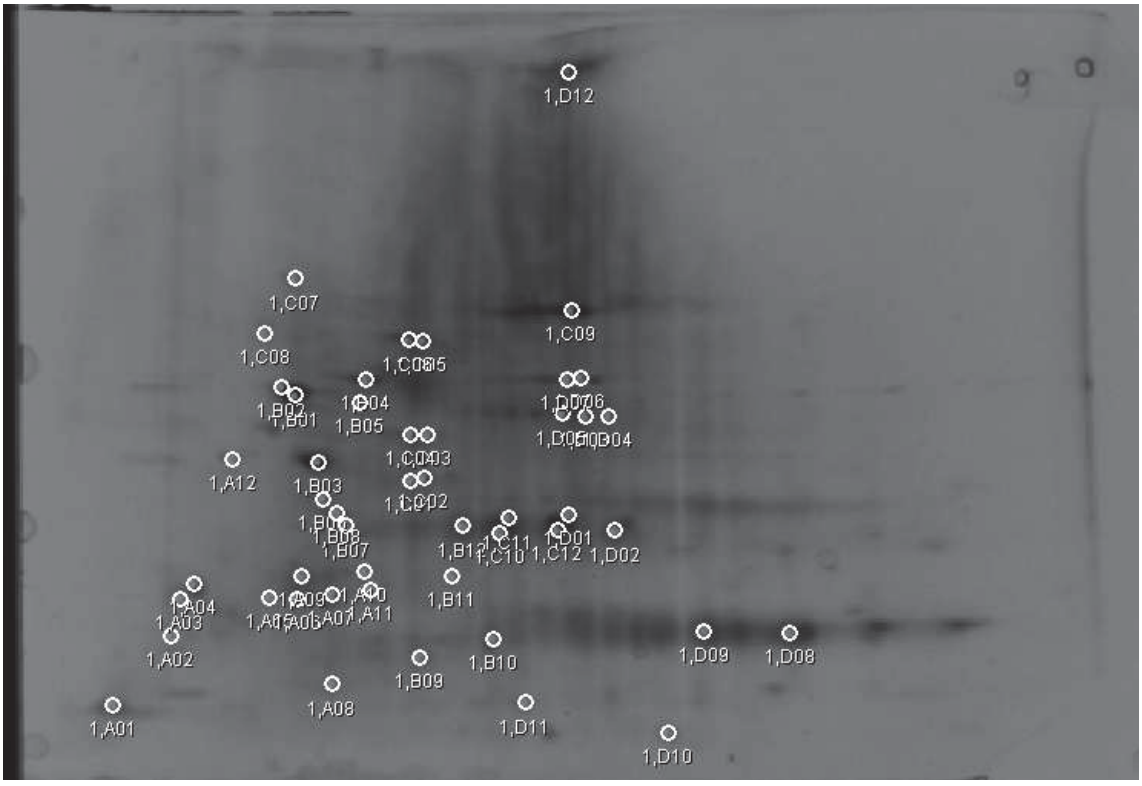

Figure 4.8B: PBS extracted protein from semi-engorged female *I. holocyclus* equalised with Proteominer.

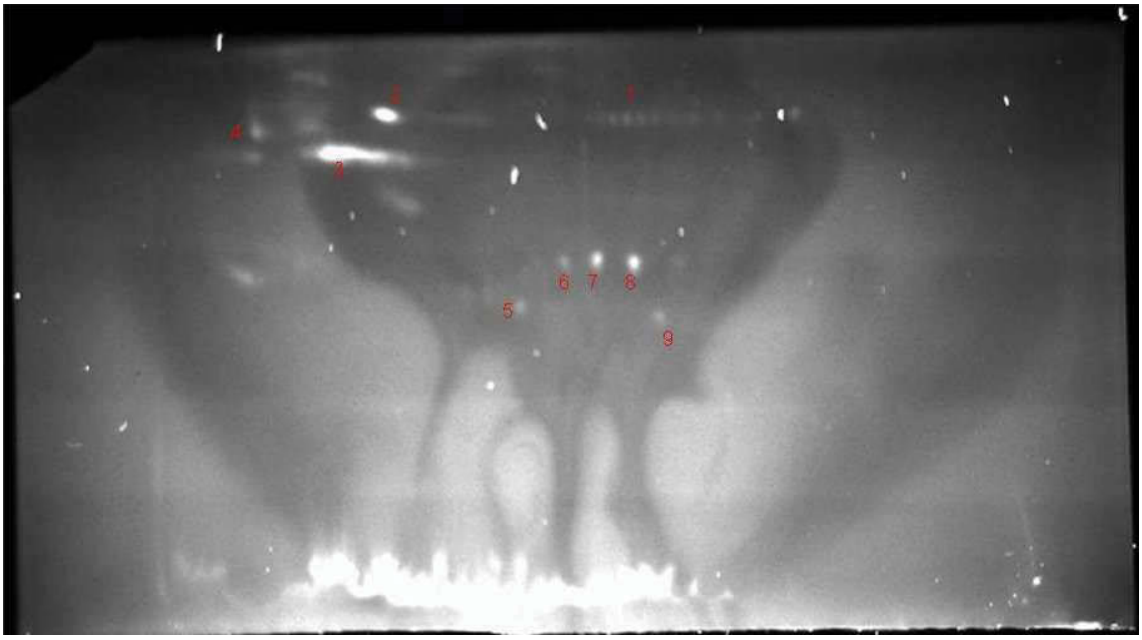

Figure 4.12B: Fluorescent zymogram of Proteominer equalised, PBS extracted protein from engorged *I. holocyclus*.

Bibliography.

AFT, R. L. & MUELLER, G. C. (1983) Hemin-mediated DNA strand scission. *J Biol Chem,* 258**,** 12069-72.

AFT, R. L. & MUELLER, G. C. (1984) Hemin-mediated oxidative degradation of proteins. *J Biol Chem,* 259**,** 301-5.

AKPEK, E. K., LIU, S. H., THOMPSON, R. & GOTTSCH, J. D. (2002) Identification of paramyosin as a binding protein for calgranulin C in experimental helminthic keratitis. *Invest Ophthalmol Vis Sci,* 43**,** 2677-84.

ALARCON-CHAIDEZ, F. J., SUN, J. & WIKEL, S. K. (2007) Transcriptome analysis of the salivary glands of Dermacentor andersoni Stiles (Acari: Ixodidae). *Insect Biochem Mol Biol,* 37**,** 48-71.

ALBISTON, H. E. (1968) Arthropod infestation - ticks and mites. *Diseases of domestic animals in Australia. Part 3.* Canberra, Australian Government Department of Health.

ALEXANDER, J. O. (1986) The physiology of itch. *Parasitol Today,* 2**,** 345-51.

ALLINGHAM, P., KERLIN, R., TELLAM, S., BRISCOE, S. & STANDFAST, H. (1992) Passage of host immunoglobulin across the mid-gut epithelium into the haemolymph of blood-fed buffalo flies *Haematobia irritans exigua*. *Journal of Insect Physiology,* 38**,** 9-17.

ALMAZAN, C., BLAS-MACHADO, U., KOCAN, K. M., YOSHIOKA, J. H., BLOUIN, E. F., MANGOLD, A. J. & DE LA FUENTE, J. (2005) Characterization of three Ixodes scapularis cDNAs protective against tick infestations. *Vaccine,* 23**,** 4403-16.

ANDERSON, N. L. & ANDERSON, N. G. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics,* 1**,** 845-67.

ANDERSON, N. L., POLANSKI, M., PIEPER, R., GATLIN, T., TIRUMALAI, R. S., CONRADS, T. P., VEENSTRA, T. D., ADKINS, J. N., POUNDS, J. G., FAGAN, R. & LOBLEY, A. (2004) The human plasma proteome: a nonredundant list developed by combination of four separate sources. *Mol Cell Proteomics,* 3**,** 311-26.

ANDRADE, B. B., TEIXEIRA, C. R., BARRAL, A. & BARRAL-NETTO, M. (2005) Haematophagous arthropod saliva and host defense system: a tale of tear and blood. *An Acad Bras Cienc,* 77**,** 665-93.

ANDREEV, V. P., REJTAR, T., CHEN, H. S., MOSKOVETS, E. V., IVANOV, A. R. & KARGER, B. L. (2003) A universal denoising and peak picking algorithm for LC-MS based on matched filtration in the chromatographic time domain. *Anal Chem,* 75**,** 6314-26.

ARMSTRONG, P. B. (2001) The contribution of proteinase inhibitors to immune defense. *Trends Immunol,* 22**,** 47-52.

ASTIGARRAGA, A., OLEAGA-PEREZ, A., PEREZ-SANCHEZ, R., BARANDA, J. A. & ENCINAS-GRANDES, A. (1997) Host immune response evasion strategies in Ornithodoros erraticus and O. moubata and their relationship to the development of an antiargasid vaccine. *Parasite Immunol,* 19**,** 401-10.

AU, J. S., CHO, W. C., YIP, T. T., YIP, C., ZHU, H., LEUNG, W. W., TSUI, P. Y., KWOK, D. L., KWAN, S. S., CHENG, W. W., TZANG, L. C., YANG, M. & LAW, S. C. (2007) Deep proteome profiling of sera from never-smoked lung cancer patients. *Biomed Pharmacother,* 61**,** 570-7.

AZARKAN, M., HUET, J., BAEYENS-VOLANT, D., LOOZE, Y. & VANDENBUSSCHE, G. (2007) Affinity chromatography: a useful tool in proteomics studies. *J Chromatogr B Analyt Technol Biomed Life Sci,* 849**,** 81-90.

BALLOUL, J. M., SONDERMEYER, P., DREYER, D., CAPRON, M., GRZYCH, J. M., PIERCE, R. J., CARVALLO, D., LECOCQ, J. P. & CAPRON, A. (1987) Molecular cloning of a protective antigen of schistosomes. *Nature,* 326**,** 149-53.

BEAUDOUIN, E., KANNY, G., GUERIN, B., GUERIN, L., PLENAT, F. & MONERET-VAUTRIN, D. A. (1997) Unusual manifestations of hypersensitivity after a tick bite: report of two cases. *Ann.Allergy Asthma Immunol.,* 79**,** 43-46.

BECKER, K., TILLEY, L., VENNERSTROM, J. L., ROBERTS, D., ROGERSON, S. & GINSBURG, H. (2004) Oxidative stress in malaria parasite-infected erythrocytes: host-parasite interactions. *Int J Parasitol,* 34**,** 163-89.

BECKMANN, R. (1989) When a tick takes hold. *Ecos,* 59**,** 13-17.

BEN-YAKIR, D. (1989) Quantitative studies of host immunoglobulin G in the hemolymph of ticks (Acari). *J Med Entomol,* 26**,** 243-6.

BERGMAN, D. K., RAMACHANDRA, R. N. & WIKEL, S. K. (1995) Dermacentor andersoni: salivary gland proteins suppressing T-lymphocyte responses to concanavalin A in vitro. *Exp Parasitol,* 81**,** 262-71.

BERGMANN, S., ROHDE, M., CHHATWAL, G. S. & HAMMERSCHMIDT, S. (2001) alpha-Enolase of Streptococcus pneumoniae is a plasmin(ogen)-binding protein displayed on the bacterial cell surface. *Mol Microbiol,* 40**,** 1273-87.

BIENVENUT, W. V., DEON, C., PASQUARELLO, C., CAMPBELL, J. M., SANCHEZ, J. C., VESTAL, M. L. & HOCHSTRASSER, D. F. (2002) Matrix-assisted laser desorption/ionization-tandem mass spectrometry with high resolution and sensitivity for identification and characterization of proteins. *Proteomics,* 2**,** 868-76.

BISHOP, R., LAMBSON, B., WELLS, C., PANDIT, P., OSASO, J., NKONGE, C., MORZARIA, S., MUSOKE, A. & NENE, V. (2002) A cement protein of the tick Rhipicephalus appendiculatus, located in the secretory e cell granules of the type III salivary gland acini, induces strong antibody responses in cattle. *Int J Parasitol,* 32**,** 833-42.

BLANTON, R. E., LICATE, L. S. & AMAN, R. A. (1994) Characterization of a native and recombinant Schistosoma haematobium serine protease inhibitor gene product. *Mol Biochem Parasitol,* 63**,** 1-11.

BODNAR, W. M., BLACKBURN, R. K., KRISE, J. M. & MOSELEY, M. A. (2003) Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. *J Am Soc Mass Spectrom,* 14**,** 971-9.

BOLDBAATAR, D., SIKALIZYO SIKASUNGE, C., BATTSETSEG, B., XUAN, X. & FUJISAKI, K. (2006) Molecular cloning and functional characterization of an aspartic protease from the hard tick Haemaphysalis longicornis. *Insect Biochem Mol Biol,* 36**,** 25-36.

BOSCHETTI, E., LOMAS, L., CITTERIO, A. & RIGHETTI, P. G. (2007) Romancing the "hidden proteome", Anno Domini two zero zero seven. *J Chromatogr A,* 1153**,** 277-90.

BOWMAN, A. S., COONS, L. B., NEEDHAM, G. R. & SAUER, J. R. (1997) Tick saliva: recent advances and implications for vector competence. *Med.Vet.Entomol.,* 11**,** 277-285.

BOWMAN, A. S., DILLWITH, J. W. & SAUER, J. R. (1996) Tick salivary prostaglandins:Presence, origin and significance. *Parasitology Today,* 12**,** 388-396.

BOWMAN, A. S., SAUER, J. R., NEESE, P. A. & DILLWITH, J. W. (1995) Origin of arachidonic acid in the salivary glands of the lone star tick, Amblyomma, americanum. *Insect Biochemistry and Molecular Biology.,* 25**,** 225-233.

BRAZ, G. R., COELHO, H. S., MASUDA, H. & OLIVEIRA, P. L. (1999) A missing metabolic pathway in the cattle tick Boophilus microplus. *Curr Biol,* 9**,** 703-6.

BRECI, L., HATTRUP, E., KEELER, M., LETARTE, J., JOHNSON, R. & HAYNES, P. A. (2005) Comprehensive proteomics in yeast using chromatographic fractionation, gas phase fractionation, protein gel electrophoresis, and isoelectric focusing. *Proteomics,* 5**,** 2018-28.

CASTAGNA, A., CECCONI, D., SENNELS, L., RAPPSILBER, J., GUERRIER, L., FORTIS, F., BOSCHETTI, E., LOMAS, L. & RIGHETTI, P. G. (2005) Exploring the hidden human urinary proteome via ligand library beads. *J Proteome Res,* 4**,** 1917-30.

CHAMPAGNE, D. E. (2005) Antihemostatic molecules from saliva of blood-feeding arthropods. *Pathophysiol Haemost Thromb,* 34**,** 221-7.

CHEN, H. S., REJTAR, T., ANDREEV, V., MOSKOVETS, E. & KARGER, B. L. (2005) Enhanced characterization of complex proteomic samples using LC-MALDI MS/MS: exclusion of redundant peptides from MS/MS analysis in replicate runs. *Anal Chem,* 77**,** 7816-25.

CHEVALLET, M., SANTONI, V., POINAS, A., ROUQUIE, D., FUCHS, A., KIEFFER, S., ROSSIGNOL, M., LUNARDI, J., GARIN, J. & RABILLOUD, T. (1998) New zwitterionic detergents improve the analysis of membrane proteins by two-dimensional electrophoresis. *Electrophoresis,* 19**,** 1901-9.

CHIN, D. & MEANS, A. R. (2000) Calmodulin: a prototypical calcium sensor. *Trends Cell Biol,* 10**,** 322-8.

CHMELAR, J., ANDERSON, J. M., MU, J., JOCHIM, R. C., VALENZUELA, J. G. & KOPECKY, J. (2008) Insight into the sialome of the castor bean tick, Ixodes ricinus. *BMC Genomics,* 9**,** 233.

CHOUDHARY, J. S., BLACKSTOCK, W. P., CREASY, D. M. & COTTRELL, J. S. (2001) Matching peptide mass spectra to EST and genomic DNA databases. *Trends Biotechnol,* 19**,** S17-22.

CITELLI, M., LARA, F. A., DA SILVA VAZ, I., JR. & OLIVEIRA, P. L. (2007) Oxidative stress impairs heme detoxification in the midgut of the cattle tick, Rhipicephalus (Boophilus) microplus. *Mol Biochem Parasitol,* 151**,** 81-8.

COLLIS, M. G. (1989) The vasodilator role of adenosine. *Pharmacol Ther,* 41**,** 143-62.

COON, J. J., UEBERHEIDE, B., SYKA, J. E., DRYHURST, D. D., AUSIO, J., SHABANOWITZ, J. & HUNT, D. F. (2005) Protein identification using sequential ion/ion reactions and tandem mass spectrometry. *Proc Natl Acad Sci U S A,* 102**,** 9463-8.

COONS, L. B., LAMOREAUX, W. J., ROSELL-DAVIS, R. & TARNOWSKI, B. I. (1989) Onset of vitellogenin production and vitellogenesis, and their relationship to changes in the midgut epithelium and oocytes in the tick Dermacentor variabilis. *Exp Appl Acarol,* 6**,** 291-305.

COPPOLINO, M. G. & DEDHAR, S. (1998) Calreticulin. *Int J Biochem Cell Biol,* 30**,** 553-8.

CUPP, E. W. (1991) Biology of ticks. *Vet.Clin.North Am.Small Anim Pract.,* 21**,** 1-26.

DAI, E., STEWART, M., RITCHIE, B., MESAELI, N., RAHA, S., KOLODZIEJCZYK, D., HOBMAN, M. L., LIU, L. Y., ETCHES, W., NATION, N., MICHALAK, M. & LUCAS, A. (1997) Calreticulin, a potential vascular regulatory protein, reduces intimal hyperplasia after arterial injury. *Arterioscler Thromb Vasc Biol,* 17**,** 2359-68.

DAVEY, R. B. & COOKSEY, L. M. (1989) Effects of prolonged exposure at low temperature on Boophilus microplus (Acari: Ixodidae). *J.Med.Entomol.,* 26**,** 407-410.

DE LA FUENTE, J., ALMAZAN, C., CANALES, M., PEREZ DE LA LASTRA, J. M., KOCAN, K. M. & WILLADSEN, P. (2007) A ten-year review of commercial vaccine performance for control of tick infestations on cattle. *Anim Health Res Rev,* 8**,** 23-8.

DE LA FUENTE, J., RODRIGUEZ, M. & GARCIA-GARCIA, J. C. (2000) Immunological control of ticks through vaccination with Boophilus microplus gut antigens. *Ann N Y Acad Sci,* 916**,** 617-21.

DE VOS, S., ZEINSTRA, L., TAOUFIK, A., WILLADSEN, P. & JONGEJAN, F. (2001) Evidence for the utility of the Bm86 antigen from Boophilus microplus in vaccination against other tick species. *Experimental & Applied Acarology,* 25**,** 245-261.

DENG, J., GOLD, D., LOVERDE, P. T. & FISHELSON, Z. (2003) Inhibition of the complement membrane attack complex by Schistosoma mansoni paramyosin. *Infect Immun,* 71**,** 6402-10.

DONOHUE, K. V., KHALIL, S. M., MITCHELL, R. D., SONENSHINE, D. E. & ROE, R. M. (2008) Molecular characterization of the major hemelipoglycoprotein in ixodid ticks. *Insect Mol Biol,* 17**,** 197-208.

DOREY, C. (1991) Allergens of the Australian Paralysis Tick, *Ixodes holocyclus*. Sydney, University of Technology, Sydney.

DOUCH, P. G. & BUCHANAN, L. L. (1978) Glutathione conjugation of some xenobiotics by Ascaris suum and Moniezia expansa. *Xenobiotica,* 8**,** 171-6.

ESCOUBAS, P., QUINTON, L. & NICHOLSON, G. M. (2008) Venomics: unravelling the complexity of animal venoms with mass spectrometry. *J Mass Spectrom,* 43**,** 279-95.

FENG, Q. L., DAVEY, K. G., PANG, A. S., PRIMAVERA, M., LADD, T. R., ZHENG, S. C., SOHI, S. S., RETNAKARAN, A. & PALLI, S. R. (1999) Glutathione S-transferase from the spruce budworm, Choristoneura fumiferana: identification, characterization, localization, cDNA cloning, and expression. *Insect Biochem Mol Biol,* 29**,** 779-93.

FERREIRA, B. R. & SILVA, J. S. (1999) Successive tick infestations selectively promote a T-helper 2 cytokine profile in mice. *Immunology,* 96**,** 434-439.

FERREIRA, C. A., BARBOSA, M. C., SILVEIRA, T. C., VALENZUELA, J. G., VAZ IDA, S., JR. & MASUDA, A. (2002a) cDNA cloning, expression and characterization of a Boophilus microplus paramyosin. *Parasitology,* 125**,** 265-74.

FERREIRA, C. A., DA SILVA VAZ, I., DA SILVA, S. S., HAAG, K. L., VALENZUELA, J. G. & MASUDA, A. (2002b) Cloning and partial characterization of a Boophilus microplus (Acari: Ixodidae) calreticulin. *Exp Parasitol,* 101**,** 25-34.

FIELDEN, L. J., JONES, R. M., GOLDBERG, M. & RECHAV, Y. (1999) Feeding and respiratory gas exchange in the American dog tick, Dermacentor variabilis. *J Insect Physiol,* 45**,** 297-304.

FRANCISCHETTI, I. M., MANS, B. J., MENG, Z., GUDDERRA, N., VEENSTRA, T. D., PHAM, V. M. & RIBEIRO, J. M. (2008) An insight into the sialome of the soft tick, Ornithodorus parkeri. *Insect Biochem Mol Biol,* 38**,** 1-21.

FRANCISCHETTI, I. M., MATHER, T. N. & RIBEIRO, J. M. (2003) Cloning of a salivary gland metalloprotease and characterization of gelatinase and fibrin(ogen)lytic activities in the saliva of the Lyme disease tick vector Ixodes scapularis. *Biochem Biophys Res Commun,* 305**,** 869-75.

FRANCISCHETTI, I. M., MY PHAM, V., MANS, B. J., ANDERSEN, J. F., MATHER, T. N., LANE, R. S. & RIBEIRO, J. M. (2005) The transcriptome of the salivary glands of the female western black-legged tick Ixodes pacificus (Acari: Ixodidae). *Insect Biochem Mol Biol,* 35**,** 1142-61.

FRANCISCHETTI, I. M., VALENZUELA, J. G., ANDERSEN, J. F., MATHER, T. N. & RIBEIRO, J. M. (2002) Ixolaris, a novel recombinant tissue factor pathway inhibitor (TFPI) from the salivary gland of the tick, Ixodes scapularis: identification of factor X and factor Xa as scaffolds for the inhibition of factor VIIa/tissue factor complex. *Blood,* 99**,** 3602-12.

FREITAS, D. R., ROSA, R. M., MORAES, J., CAMPOS, E., LOGULLO, C., DA SILVA VAZ, I., JR. & MASUDA, A. (2007) Relationship between glutathione S-transferase, catalase, oxygen consumption, lipid peroxidation and oxidative stress in eggs and larvae of Boophilus microplus (Acarina: Ixodidae). *Comp Biochem Physiol A Mol Integr Physiol,* 146**,** 688-94.

FUCHSBERGER, N., KITA, M., HAJNICKA, V., IMANISHI, J., LABUDA, M. & NUTTALL, P. A. (1995) Ixodid tick salivary gland extracts inhibit production of lipopolysaccharide-induced mRNA of several different human cytokines. *Exp Appl Acarol,* 19**,** 671-6.

FUJIKI, Y., HUBBARD, A. L., FOWLER, S. & LAZAROW, P. B. (1982) Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. *J Cell Biol,* 93**,** 97-102.

GASPAR, A. R., JOUBERT, A. M., CRAUSE, J. C. & NEITZ, A. W. (1996) Isolation and characterization of an anticoagulant from the salivary glands of the tick, Ornithodoros savignyi (Acari: Argasidae). *Exp.Appl.Acarol.,* 20**,** 583-598.

GAUCI, M., STONE, B. F. & THONG, Y. H. (1988a) Detection in allergic individuals of IgE specific for the Australian paralysis tick, Ixodes holocyclus. *Int.Arch.Allergy Appl.Immunol.,* 85**,** 190-193.

GAUCI, M., STONE, B. F. & THONG, Y. H. (1988b) Isolation and immunological characterisation of allergens from salivary glands of the Australian paralysis tick Ixodes holocyclus. *Int.Arch.Allergy Appl.Immunol.,* 87**,** 208-212.

GHAEMMAGHAMI, S., HUH, W. K., BOWER, K., HOWSON, R. W., BELLE, A., DEPHOURE, N., O'SHEA, E. K. & WEISSMAN, J. S. (2003) Global analysis of protein expression in yeast. *Nature,* 425**,** 737-41.

GILLESPIE, R. D., DOLAN, M. C., PIESMAN, J. & TITUS, R. G. (2001) Identification of an IL-2 binding protein in the saliva of the Lyme disease vector tick, Ixodes scapularis. *J Immunol,* 166**,** 4319-26.

GILLESPIE, R. D., MBOW, M. L. & TITUS, R. G. (2000) The immunomodulatory factors of bloodfeeding arthropod saliva. *Parasite Immunol,* 22**,** 319-31.

GOBERT, G. N. & MCMANUS, D. P. (2005) Update on paramyosin in parasitic worms. *Parasitol Int,* 54**,** 101-7.

GOLDBERG, D. E., SLATER, A. F., CERAMI, A. & HENDERSON, G. B. (1990) Hemoglobin degradation in the malaria parasite Plasmodium falciparum: an ordered process in a unique organelle. *Proc Natl Acad Sci U S A,* 87**,** 2931-5.

GORDON, J. R. & ALLEN, J. R. (1991) Factors V and VII anticoagulant activities in the salivary glands of feeding Dermacentor andersoni ticks. *J Parasitol,* 77**,** 167-70.

GORG, A., WEISS, W. & DUNN, M. J. (2004) Current two-dimensional electrophoresis technology for proteomics. *Proteomics,* 4**,** 3665-85.

GRACA-SOUZA, A. V., MAYA-MONTEIRO, C., PAIVA-SILVA, G. O., BRAZ, G. R., PAES, M. C., SORGINE, M. H., OLIVEIRA, M. F. & OLIVEIRA, P. L. (2006) Adaptations against heme toxicity in blood-feeding arthropods. *Insect Biochem Mol Biol,* 36**,** 322-35.

GRANVOGL, B., GRUBER, P. & EICHACKER, L. A. (2007a) Standardisation of rapid in-gel digestion by mass spectrometry. *Proteomics,* 7**,** 642-54.

GRANVOGL, B., PLOSCHER, M. & EICHACKER, L. A. (2007b) Sample preparation by in-gel digestion for mass spectrometry-based proteomics. *Anal Bioanal Chem,* 389**,** 991-1002.

GUDDERRA, N. P., NEESE, P. A., SONENSHINE, D. E., APPERSON, C. S. & ROE, R. M. (2001) Developmental profile, isolation, and biochemical characterization of a novel

lipoglycoheme-carrier protein from the American dog tick, Dermacentor variabilis (Acari: Ixodidae) and observations on a similar protein in the soft tick, Ornithodoros parkeri (Acari: Argasidae). *Insect Biochem Mol Biol,* 31**,** 299-311.

GUDDERRA, N. P., SONENSHINE, D. E., APPERSON, C. S. & ROE, R. M. (2002) Hemolymph proteins in ticks. *J Insect Physiol,* 48**,** 269-278.

GUERRERO, F. D., MILLER, R. J., ROUSSEAU, M. E., SUNKARA, S., QUACKENBUSH, J., LEE, Y. & NENE, V. (2005) BmiGI: a database of cDNAs expressed in Boophilus microplus, the tropical/southern cattle tick. *Insect Biochem Mol Biol,* 35**,** 585-95.

GUERRIER, L. & BOSCHETTI, E. (2007) Protocol for the purification of proteins from biological extracts for identification by mass spectrometry. *Nat Protoc,* 2**,** 831-7.

GUERRIER, L., CLAVEROL, S., FORTIS, F., RINALDUCCI, S., TIMPERIO, A. M., ANTONIOLI, P., JANDROT-PERRUS, M., BOSCHETTI, E. & RIGHETTI, P. G. (2007) Exploring the platelet proteome via combinatorial, hexapeptide ligand libraries. *J Proteome Res,* 6**,** 4290-303.

GUERRIER, L., D'AUTREAUX, B., ATANASSOV, C., KHODER, G. & BOSCHETTI, E. (2008) Evaluation of a standardized method of protein purification and identification after discovery by mass spectrometry. *J Proteomics,* 71**,** 368-78.

GUILFOILE, P. G. & PACKILA, M. (2004) Identification of four genes expressed by feeding female Ixodes scapularis, including three with sequence similarity to previously recognized genes. *Exp Appl Acarol,* 32**,** 103-10.

GUTTERIDGE, J. M. & SMITH, A. (1988) Antioxidant protection by haemopexin of haem-stimulated lipid peroxidation. *Biochem J,* 256**,** 861-5.

HAN, Y., MA, B. & ZHANG, K. (2005) SPIDER: software for protein identification from sequence tags with de novo sequencing error. *J Bioinform Comput Biol,* 3**,** 697-716.

HE, H., CHEN, A. C., DAVEY, R. B., IVIE, G. W. & GEORGE, J. E. (1999) Characterization and molecular cloning of a glutathione S-transferase gene from the tick, Boophilus microplus (Acari: Ixodidae) [In Process Citation]. *Insect Biochemistry and Molecular Biology.,* 29**,** 737-743.

HERBERT, B. (1999) Advances in protein solubilisation for two-dimensional electrophoresis. *Electrophoresis,* 20**,** 660-3.

HERBERT, B., GALVANI, M., HAMDAN, M., OLIVIERI, E., MACCARTHY, J., PEDERSEN, S. & RIGHETTI, P. G. (2001) Reduction and alkylation of proteins in preparation of two-dimensional map analysis: why, when, and how? *Electrophoresis,* 22**,** 2046-57.

HERBERT, B. & RIGHETTI, P. G. (2000) A turning point in proteome analysis: sample prefractionation via multicompartment electrolyzers with isoelectric membranes. *Electrophoresis,* 21**,** 3639-48.

HERBERT, B. R. & HARRY, E. (2008) Difficult Proteins. IN SHEEHAN, D. (Ed.) *Two Dimensional Electrophoresis Protocols.*, Humana Press.

HERBERT, B. R., RIGHETTI, P. G., CITTERIO, A. & BOSCHETTI, E. (2007) Sample Preparation and Prefractionation Techniques for Electrophoresis-Based Proteomics. IN WILKINS, M. R., APPEL, R. D., WILLIAMS, K. L. & HOCHSTRASSER, D. F. (Eds.) *Proteome Research: Concepts, Technology and Application. .* Berlin Heidelberg, Springer-Verlag

HILL, C. A. & WIKEL, S. K. (2005) The Ixodes scapularis Genome Project: an opportunity for advancing tick research. *Trends Parasitol,* 21**,** 151-3.

HOCHACHKA, P. W. & LUTZ, P. L. (2001) Mechanism, origin, and evolution of anoxia tolerance in animals. *Comp Biochem Physiol B Biochem Mol Biol,* 130**,** 435-59.

HOFMEISTER, F. (1888) Zur Lehre von der Wirkung der Salze. *Archive of Experimental Pathology and Pharmacology.,* 24**,** 247-260.

HUANG, H. S., HU, N. T., YAO, Y. E., WU, C. Y., CHIANG, S. W. & SUN, C. N. (1998) Molecular cloning and heterologous expression of a glutathione S-transferase involved in insecticide resistance from the diamondback moth, Plutella xylostella. *Insect Biochem Mol Biol,* 28**,** 651-8.

HUNTINGTON, J. A., READ, R. J. & CARRELL, R. W. (2000) Structure of a serpin-protease complex shows inhibition by deformation. *Nature,* 407**,** 923-6.

IMAMURA, S., DA SILVA VAZ JUNIOR, I., SUGINO, M., OHASHI, K. & ONUMA, M. (2005) A serine protease inhibitor (serpin) from Haemaphysalis longicornis as an anti-tick vaccine. *Vaccine,* 23**,** 1301-11.

INOKUMA, H., KEMP, D. H. & WILLADSEN, P. (1994) Prostaglandin E2 production by the cattle tick (Boophilus microplus) into feeding sites and its effect on the response of bovine mononuclear cells to mitogen. *Vet Parasitol,* 53**,** 293-9.

JAEGER, T. & FLOHE, L. (2006) The thiol-based redox networks of pathogens: unexploited targets in the search for new drugs. *Biofactors,* 27**,** 109-20.

JAWORSKI, D., SIMMEN, F., LAMOREAUX, W., COONS, L., MULLER, M. & NEEDHAM, G. (1995) A secreted calreticulin protein in ixodid tick (Amblyomma americanum) saliva. *Journal of Insect Physiology,* 41**,** 369-375.

JEONG, K. Y., HONG, C. S. & YONG, T. S. (2006) Allergenic tropomyosins and their cross-reactivities. *Protein Pept Lett,* 13**,** 835-45.

JONG, A. Y., CHEN, S. H., STINS, M. F., KIM, K. S., TUAN, T. L. & HUANG, S. H. (2003) Binding of Candida albicans enolase to plasmin(ogen) results in enhanced invasion of human brain microvascular endothelial cells. *J Med Microbiol,* 52**,** 615-22.

JOSS, J. L., MOLLOY, M. P., HINDS, L. & DEANE, E. (2007) Proteomic analysis of early lactation milk of the tammar wallaby (*Macropus eugenii*). *Comparative Biochemistry and Physiology, Part D.,* 2**,** 150-164.

JOSS, J. L., MOLLOY, M. P., HINDS, L. A. & DEANE, E. M. (2006) Evaluation of chemical derivatisation methods for protein identification using MALDI MS/MS. *International Journal of Peptide Research and Therapeutics,* 12**,** 225-235.

JUAN, H. F., CHANG, S. C., HUANG, H. C. & CHEN, S. T. (2005) A new application of microwave technology to proteomics. *Proteomics,* 5**,** 840-2.

KASHINO, S. S., RESENDE, J., SACCO, A. M., ROCHA, C., PROENCA, L., CARVALHO, W. A., FIRMINO, A. A., QUEIROZ, R., BENAVIDES, M., GERSHWIN, L. J. & DE MIRANDA SANTOS, I. K. (2005) Boophilus microplus: the pattern of bovine immunoglobulin isotype responses to high and low tick infestations. *Exp Parasitol,* 110**,** 12-21.

KETTERMAN, A. J., PROMMEENATE, P., BOONCHAUY, C., CHANAMA, U., LEETACHEWA, S., PROMTET, N. & PRAPANTHADARA, L. (2001) Single amino acid changes outside the active site significantly affect activity of glutathione S-transferases. *Insect Biochem Mol Biol,* 31**,** 65-74.

KHYSE-ANDERSON, J. (1984) Electroblotting of multiple gels. A simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide gels to nitrocellulose. *Journal of Biochemical and Biophysical Methods.,* 10**,** 203-209.

KIEL, M., JOSH, P., JONES, A., WINDON, R., HUNT, P. & KONGSUWAN, K. (2007) Identification of immuno-reactive proteins from a sheep gastrointestinal nematode, Trichostrongylus colubriformis, using two-dimensional electrophoresis and mass spectrometry. *Int J Parasitol,* 37**,** 1419-29.

KIM, K., KIM, I. H., LEE, K. Y., RHEE, S. G. & STADTMAN, E. R. (1988) The isolation and purification of a specific "protector" protein which inhibits enzyme inactivation by a thiol/Fe(III)/O2 mixed-function oxidation system. *J Biol Chem,* 263**,** 4704-11.

KNIZETOVA, P., VANCOVA, I., KOCAKOVA, P., SLOVAK, M., PROOST, P. & KOPACEK, J. (2006) New member of the protein disulfide isomerase (PDI) family identified in Amblyomma variegatum tick. *Insect Biochem Mol Biol,* 36**,** 943-53.

KOTLYAR, S., WEIHRAUCH, D., PAULSEN, R. S. & TOWLE, D. W. (2000) Expression of arginine kinase enzymatic activity and mRNA in gills of the euryhaline crabs Carcinus maenas and Callinectes sapidus. *J Exp Biol,* 203**,** 2395-404.

KRAMER, C., NAHMIAS, Z., NORMAN, D. D., MULVIHILL, T. A., COONS, L. B. & COLE, J. A. (2008) Dermacentor variabilis: Regulation of fibroblast migration by tick salivary gland extract and saliva. *Exp Parasitol,* 119**,** 391-7.

KRAUSE, E., WENSCHUH, H. & JUNGBLUT, P. R. (1999) The dominance of arginine-containing peptides in MALDI-derived tryptic mass fingerprints of proteins. *Anal Chem,* 71**,** 4160-5.

KREM, M. M. & DI CERA, E. (2001) Molecular markers of serine protease evolution. *EMBO J,* 20**,** 3036-45.

KRNAJSKI, Z., GILBERGER, T. W., WALTER, R. D. & MULLER, S. (2001) The malaria parasite Plasmodium falciparum possesses a functional thioredoxin system. *Mol Biochem Parasitol,* 112**,** 219-28.

KROBER, T. & GUERIN, P. M. (2007) In vitro feeding assays for hard ticks. *Trends Parasitol,* 23**,** 445-9.

KRUTCHINSKY, A. N., ZHANG, W. & CHAIT, B. T. (2000) Rapidly switchable matrix-assisted laser desorption/ionization and electrospray quadrupole-time-of-flight mass spectrometry for protein identification. *J Am Soc Mass Spectrom,* 11**,** 493-504.

KUBES, M., KOCAKOVA, P., SLOVAK, M., SLAVIKOVA, M., FUCHSBERGER, N. & NUTTALL, P. A. (2002) Heterogeneity in the effect of different ixodid tick species on human natural killer cell activity. *Parasite Immunol,* 24**,** 23-8.

LABRUNA, M. B., AMAKU, M., METZNER, J. A., PINTER, A. & FERREIRA, F. (2003) Larval behavioral diapause regulates life cycle of Amblyomma cajennense (Acari: Ixodidae) in Southeast Brazil. *J Med Entomol,* 40**,** 170-8.

LACLETTE, J. P., SHOEMAKER, C. B., RICHTER, D., ARCOS, L., PANTE, N., COHEN, C., BING, D. & NICHOLSON-WELLER, A. (1992) Paramyosin inhibits complement C1. *J Immunol,* 148**,** 124-8.

LAI, R., LOMAS, L. O., JONCZY, J., TURNER, P. C. & REES, H. H. (2004) Two novel non-cationic defensin-like antimicrobial peptides from haemolymph of the female tick, Amblyomma hebraeum. *Biochem J,* 379**,** 681-5.

LANG, A. B., WYSS, C. & EPPENBERGER, H. M. (1980) Localization of arginine kinase in muscles fibres of Drosophila melanogaster. *J Muscle Res Cell Motil,* 1**,** 147-61.

LANGE, V., MALMSTROM, J. A., DIDION, J., KING, N. L., JOHANSSON, B. P., SCHAFER, J., RAMESEDER, J., WONG, C. H., DEUTSCH, E. W., BRUSNIAK, M. Y., BUHLMANN, P., BJORCK, L., DOMON, B. & AEBERSOLD, R. (2008) Targeted quantitative analysis of Streptococcus pyogenes virulence factors by multiple reaction monitoring. *Mol Cell Proteomics*.

LARA, F. A., LINS, U., BECHARA, G. H. & OLIVEIRA, P. L. (2005) Tracing heme in a living cell: hemoglobin degradation and heme traffic in digest cells of the cattle tick Boophilus microplus. *J Exp Biol,* 208**,** 3093-101.

LARA, F. A., LINS, U., PAIVA-SILVA, G., ALMEIDA, I. C., BRAGA, C. M., MIGUENS, F. C., OLIVEIRA, P. L. & DANSA-PETRETSKI, M. (2003) A new intracellular pathway of haem detoxification in the midgut of the cattle tick Boophilus microplus: aggregation inside a specialized organelle, the hemosome. *J Exp Biol,* 206**,** 1707-15.

LEAL, A. T., SEIXAS, A., POHL, P. C., FERREIRA, C. A., LOGULLO, C., OLIVEIRA, P. L., FARIAS, S. E., TERMIGNONI, C., DA SILVA VAZ, I., JR. & MASUDA, A. (2006) Vaccination of bovines with recombinant Boophilus Yolk pro-Cathepsin. *Vet Immunol Immunopathol*.

LEVASHINA, E. A., LANGLEY, E., GREEN, C., GUBB, D., ASHBURNER, M., HOFFMANN, J. A. & REICHHART, J. M. (1999) Constitutive activation of toll-mediated antifungal defense in serpin-deficient Drosophila. *Science,* 285**,** 1917-9.

LIAO, M., BOLDBAATAR, D., GONG, H., HUANG, P., UMEMIYA, R., HARNNOI, T., ZHOU, J., TANAKA, T., SUZUKI, H., XUAN, X. & FUJISAKI, K. (2008) Functional analysis of protein disulfide isomerases in blood feeding, viability and oocyte development in Haemaphysalis longicornis ticks. *Insect Biochem Mol Biol,* 38**,** 285-95.

LIAO, M., HATTA, T., UMEMIYA, R., HUANG, P., JIA, H., GONG, H., ZHOU, J., NISHIKAWA, Y., XUAN, X. & FUJISAKI, K. (2007) Identification of three protein disulfide isomerase members from Haemaphysalis longicornis tick. *Insect Biochem Mol Biol,* 37**,** 641-54.

LIMO, M. K., VOIGT, W. P., TUMBO-OERI, A. G., NJOGU, R. M. & OLE-MOIYOI, O. K. (1991) Purification and characterization of an anticoagulant from the salivary glands of the ixodid tick Rhipicephalus appendiculatus. *Exp Parasitol,* 72**,** 418-29.

LOPEZ-FERRER, D., CAPELO, J. L. & VAZQUEZ, J. (2005) Ultra fast trypsin digestion of proteins by high intensity focused ultrasound. *J Proteome Res,* 4**,** 1569-74.

LOPEZ-FERRER, D., PETRITIS, K., HIXSON, K. K., HEIBECK, T. H., MOORE, R. J., BELOV, M. E., CAMP, D. G., 2ND & SMITH, R. D. (2008) Application of Pressurized Solvents for Ultrafast Trypsin Hydrolysis in Proteomics: Proteomics on the Fly. *J Proteome Res*.

LOUKAS, A., JONES, M. K., KING, L. T., BRINDLEY, P. J. & MCMANUS, D. P. (2001) Receptor for Fc on the surfaces of schistosomes. *Infect Immun,* 69**,** 3646-51.

LU, X. & ZHU, H. (2005) Tube-gel digestion: a novel proteomic approach for high throughput analysis of membrane proteins. *Mol Cell Proteomics,* 4**,** 1948-58.

LYLES, M. M. & GILBERT, H. F. (1991) Catalysis of the oxidative folding of ribonuclease A by protein disulfide isomerase: dependence of the rate on the composition of the redox buffer. *Biochemistry,* 30**,** 613-9.

MACALUSO, K. R. & WIKEL, S. K. (2001) Dermacentor andersoni: effects of repeated infestations on lymphocyte proliferation, cytokine production, and adhesion-molecule expression by BALB/c mice. *Ann Trop Med Parasitol,* 95**,** 413-27.

MACEDO-RIBEIRO, S., ALMEIDA, C., CALISTO, B. M., FRIEDRICH, T., MENTELE, R., STURZEBECHER, J., FUENTES-PRIOR, P. & PEREIRA, P. J. (2008) Isolation, cloning and structural characterisation of boophilin, a multifunctional kunitz-type proteinase inhibitor from the cattle tick. *PLoS ONE,* 3**,** e1624.

MADDEN, R. D., SAUER, J. R. & DILLWITH, J. W. (2004) A proteomics approach to characterizing tick salivary secretions. *Exp Appl Acarol,* 32**,** 77-87.

MADDEN, R. D., SAUER, J. R., DILLWITH, J. W. & BOWMAN, A. S. (1996) Alteration of arachidonate levels in tick salivary glands by dietary modification of host blood lipids. *Archives.of Insect Biochemistry and Physiology.,* 31**,** 53-72.

MANS, B. J., ANDERSEN, J. F., FRANCISCHETTI, I. M., VALENZUELA, J. G., SCHWAN, T. G., PHAM, V. M., GARFIELD, M. K., HAMMER, C. H. & RIBEIRO, J. M. (2008) Comparative sialomics between hard and soft ticks: implications for the evolution of blood-feeding behavior. *Insect Biochem Mol Biol,* 38**,** 42-58.

MANS, B. J., LOUW, A. I. & NEITZ, A. W. (2002a) Amino acid sequence and structure modeling of savignin, a thrombin inhibitor from the tick, Ornithodoros savignyi. *Insect Biochemistry and Molecular Biology.,* 32**,** 821-828.

MANS, B. J., LOUW, A. I. & NEITZ, A. W. (2002b) Evolution of hematophagy in ticks: common origins for blood coagulation and platelet aggregation inhibitors from soft ticks of the genus Ornithodoros. *Mol Biol Evol,* 19**,** 1695-705.

MANS, B. J., LOUW, A. I. & NEITZ, A. W. (2002c) Savignygrin, a platelet aggregation inhibitor from the soft tick Ornithodoros savignyi, presents the RGD integrin recognition motif on the Kunitz-BPTI fold. *J Biol Chem,* 277**,** 21371-8.

MANS, B. J. & NEITZ, A. W. (2004) Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. *Insect Biochem Mol Biol,* 34**,** 1-17.

MANS, B. J., STEINMANN, C. M., VENTER, J. D., LOUW, A. I. & NEITZ, A. W. (2002d) Pathogenic mechanisms of sand tampan toxicoses induced by the tick, Ornithodoros savignyi. *Toxicon,* 40**,** 1007-1016.

MAREN, T. H. & WILEY, C. W. (1970) Kinetics of carbonic anhydrase in whole red cells as measured by transfer of carbon dioxide and ammonia. *Mol Pharmacol,* 6**,** 430-40.

MARGULIES, E. H., MADURO, V. V., THOMAS, P. J., TOMKINS, J. P., AMEMIYA, C. T., LUO, M. & GREEN, E. D. (2005) Comparative sequencing provides insights about the structure and conservation of marsupial and monotreme genomes. *Proc Natl Acad Sci U S A,* 102**,** 3354-9.

MARITZ-OLIVIER, C., STUTZER, C., JONGEJAN, F., NEITZ, A. W. & GASPAR, A. R. (2007) Tick anti-hemostatics: targets for future vaccines and therapeutics. *Trends Parasitol,* 23**,** 397-407.

MASINA, S. (1999) The sequence, expression and immunological characterisation of the HT-1 neurotoxin from the Australian Paralysis Tick, *Ixodes holocyclus*. Sydney, University of Technology, Sydney.

MASINA, S. & BROADY, K. W. (1999) Tick paralysis: development of a vaccine. *Int J Parasitol,* 29**,** 535-41.

MATTHIESEN, R. (2007) Methods, algorithms and tools in computational proteomics: a practical point of view. *Proteomics,* 7**,** 2815-32.

MAYA-MONTEIRO, C. M., DAFFRE, S., LOGULLO, C., LARA, F. A., ALVES, E. W., CAPURRO, M. L., ZINGALI, R., ALMEIDA, I. C. & OLIVEIRA, P. L. (2000) HeLp, a heme lipoprotein from the hemolymph of the cattle tick, Boophilus microplus. *J Biol Chem,* 275**,** 36584-9.

MCLAFFERTY, F. W., BREUKER, K., JIN, M., HAN, X., INFUSINI, G., JIANG, H., KONG, X. & BEGLEY, T. P. (2007) Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *FEBS J,* 274**,** 6256-68.

MEJRI, N., RUTTI, B. & BROSSARD, M. (2002) Immunosuppressive effects of ixodes ricinus tick saliva or salivary gland extracts on innate and acquired immune response of BALB/c mice. *Parasitol.Res.,* 88**,** 192-197.

MENDIOLA, J., ALONSO, M., MARQUETTI, M. C. & FINLAY, C. (1996) Boophilus microplus: multiple proteolytic activities in the midgut. *Exp Parasitol,* 82**,** 27-33.

MEYER, D. J. & THOMAS, M. (1995) Characterization of rat spleen prostaglandin H D-isomerase as a sigma-class GSH transferase. *Biochem J,* 311 (Pt 3)**,** 739-42.

MIKESH, L. M., UEBERHEIDE, B., CHI, A., COON, J. J., SYKA, J. E., SHABANOWITZ, J. & HUNT, D. F. (2006) The utility of ETD mass spectrometry in proteomic analysis. *Biochim Biophys Acta,* 1764**,** 1811-22.

MIYOSHI, T., TSUJI, N., ISLAM, M. K., ALIM, M. A., HATTA, T., HUANG, X. & FUJISAKI, K. (2008) A set of serine proteinase paralogs are required for blood-digestion in the ixodid tick Haemaphysalis longicornis. *Parasitol Int*.

MODHA, J., ROBERTS, M. C. & KUSEL, J. R. (1996) Schistosomes and serpins: a complex business. *Parasitol Today,* 12**,** 119-21.

MOLLOY, M. P., HERBERT, B. R., WILLIAMS, K. L. & GOOLEY, A. A. (1999) Extraction of Escherichia coli proteins with organic solvents prior to two-dimensional electrophoresis. *Electrophoresis,* 20**,** 701-4.

MOTAMENI, A. R., JUNCADELLA, I. J., ANANTHANARAYANAN, S. K., HEDRICK, M. N., HUET-HUDSON, Y. & ANGUITA, J. (2004) Delivery of the immunosuppressive antigen Salp15 to antigen-presenting cells by Salmonella enterica serovar Typhimurium aroA mutants. *Infect Immun,* 72**,** 3638-42.

MULENGA, A., MISAO, O. & SUGIMOTO, C. (2003) Three serine proteinases from midguts of the hard tick Rhipicephalus appendiculatus; cDNA cloning and preliminary characterization. *Exp Appl Acarol,* 29**,** 151-64.

MULENGA, A., SUGIMOTO, C., INGRAM, G., OHASHI, K. & ONUMA, M. (1999a) Molecular cloning of two Haemaphysalis longicornis cathepsin L-like cysteine proteinase genes. *J Vet Med Sci,* 61**,** 497-502.

MULENGA, A., SUGIMOTO, C., SAKO, Y., OHASHI, K., MUSOKE, A., SHUBASH, M. & ONUMA, M. (1999b) Molecular characterization of a Haemaphysalis longicornis tick salivary gland-associated 29-kilodalton protein and its effect as a vaccine against tick infestation in rabbits. *Infect Immun,* 67**,** 1652-8.

MYUNG-JO, Y. (2005) Immunization of mice with recombinant P27/30 protein confers protection against hard tick Haemaphysalis longicornis (Acari: Ixodidae) infestation. *Journal of Veterinary Science,* 6**,** 47-51.

NASCIMENTO-SILVA, M. C., LEAL, A. T., DAFFRE, S., JULIANO, L., DA SILVA VAZ, I., JR., PAIVA-SILVA GDE, O., OLIVEIRA, P. L. & SORGINE, M. H. (2008) BYC, an atypical aspartic

endopeptidase from Rhipicephalus (Boophilus) microplus eggs. *Comp Biochem Physiol B Biochem Mol Biol,* 149**,** 599-607.

NAZARIO, S., DAS, S., DE SILVA, A. M., DEPONTE, K., MARCANTONIO, N., ANDERSON, J. F., FISH, D., FIKRIG, E. & KANTOR, F. S. (1998) Prevention of Borrelia burgdorferi transmission in guinea pigs by tick immunity. *Am.J.Trop.Med.Hyg.,* 58**,** 780-785.

NENE, V., LEE, D., QUACKENBUSH, J., SKILTON, R., MWAURA, S., GARDNER, M. J. & BISHOP, R. (2002) AvGI, an index of genes transcribed in the salivary glands of the ixodid tick Amblyomma variegatum. *Int.J.Parasitol.,* 32**,** 1447.

NEURATH, H. (1984) Evolution of proteolytic enzymes. *Science,* 224**,** 350-7.

NEURATH, H. (1986) The versatility of proteolytic enzymes. *J Cell Biochem,* 32**,** 35-49.

NUNN, M. A., SHARMA, A., PAESEN, G. C., ADAMSON, S., LISSINA, O., WILLIS, A. C. & NUTTALL, P. A. (2005) Complement inhibitor of C5 activation from the soft tick Ornithodoros moubata. *J Immunol,* 174**,** 2084-91.

NUTTALL, P. A., TRIMNELL, A. R., KAZIMIROVA, M. & LABUDA, M. (2006) Exposed and concealed antigens as vaccine targets for controlling ticks and tick-borne diseases. *Parasite Immunol,* 28**,** 155-63.

OLEAGA, A., ESCUDERO-POBLACION, A., CAMAFEITA, E. & PEREZ-SANCHEZ, R. (2007) A proteomic approach to the identification of salivary proteins from the argasid ticks Ornithodoros moubata and Ornithodoros erraticus. *Insect Biochem Mol Biol,* 37**,** 1149-59.

OXER, D. T. & RICARDO, C. L. (1942) Notes on the biology, toxicity and breeding of Ixodes holocyclus (Neumann). . *Aust Vet J* 18**,** 194-9.

PACKER, N. H., VON DER LIETH, C. W., AOKI-KINOSHITA, K. F., LEBRILLA, C. B., PAULSON, J. C., RAMAN, R., RUDD, P., SASISEKHARAN, R., TANIGUCHI, N. & YORK, W. S. (2008) Frontiers in glycomics: bioinformatics and biomarkers in disease. An NIH white paper prepared from discussions by the focus groups at a workshop on the NIH campus, Bethesda MD (September 11-13, 2006). *Proteomics,* 8**,** 8-20.

PAESEN, G. C., ADAMS, P. L., HARLOS, K., NUTTALL, P. A. & STUART, D. I. (1999) Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure. *Mol Cell,* 3**,** 661-71.

PAESEN, G. C., ADAMS, P. L., NUTTALL, P. A. & STUART, D. L. (2000) Tick histamine-binding proteins: lipocalins with a second binding cavity. *Biochim Biophys Acta,* 1482**,** 92-101.

PAGEL VAN ZEE, J., GERACI, N. S., GUERRERO, F. D., WIKEL, S. K., STUART, J. J., NENE, V. M. & HILL, C. A. (2007) Tick genomics: the Ixodes genome project and beyond. *Int J Parasitol,* 37**,** 1297-305.

PAN, S., ZHANG, H., RUSH, J., ENG, J., ZHANG, N., PATTERSON, D., COMB, M. J. & AEBERSOLD, R. (2005) High throughput proteome screening for biomarker detection. *Mol Cell Proteomics,* 4**,** 182-90.

PANCHOLI, V. (2001) Multifunctional alpha-enolase: its role in diseases. *Cell Mol Life Sci,* 58**,** 902-20.

PAVEGLIO, S. A., ALLARD, J., MAYETTE, J., WHITTAKER, L. A., JUNCADELLA, I., ANGUITA, J. & POYNTER, M. E. (2007) The tick salivary protein, Salp15, inhibits the development of experimental asthma. *J Immunol,* 178**,** 7064-71.

PITTERI, S. J., CHRISMAN, P. A., HOGAN, J. M. & MCLUCKEY, S. A. (2005) Electron transfer ion/ion reactions in a three-dimensional quadrupole ion trap: reactions of doubly and triply protonated peptides with SO2*. *Anal Chem,* 77**,** 1831-9.

PLOW, E. F., HERREN, T., REDLITZ, A., MILES, L. A. & HOOVER-PLOW, J. L. (1995) The cell biology of the plasminogen system. *FASEB J,* 9**,** 939-45.

POHL, P. C., SORGINE, M. H., LEAL, A. T., LOGULLO, C., OLIVEIRA, P. L., DA SILVA VAZ, I., JR. & MASUDA, A. (2008) An extraovarian aspartic protease accumulated in tick oocytes with vitellin-degradation activity. *Comp Biochem Physiol B Biochem Mol Biol*.

PRITCHARD, D. I., BROWN, A., KASPER, G., MCELROY, P., LOUKAS, A., HEWITT, C., BERRY, C., FULLKRUG, R. & BECK, E. (1999) A hookworm allergen which strongly resembles calreticulin. *Parasite Immunol,* 21**,** 439-50.

PRITCHARD, D. I., HEWITT, C. & MOQBEL, R. (1997) The relationship between immunological responsiveness controlled by T-helper 2 lymphocytes and infections with parasitic helminths. *Parasitology,* 115 Suppl**,** S33-44.

RABILLOUD, T., BLISNICK, T., HELLER, M., LUCHE, S., AEBERSOLD, R., LUNARDI, J. & BRAUN-BRETON, C. (1999) Analysis of membrane proteins by two-dimensional

electrophoresis: comparison of the proteins extracted from normal or Plasmodium falciparum-infected erythrocyte ghosts. *Electrophoresis,* 20**,** 3603-10.

RABILLOUD, T., LUCHE, S., SANTONI, V. & CHEVALLET, M. (2007) Detergents and chaotropes for protein solubilization before two-dimensional electrophoresis. *Methods Mol Biol,* 355**,** 111-9.

RACHINSKY, A., GUERRERO, F. D. & SCOLES, G. A. (2007a) Differential protein expression in ovaries of uninfected and Babesia-infected southern cattle ticks, Rhipicephalus (Boophilus) microplus. *Insect Biochem Mol Biol,* 37**,** 1291-308.

RACHINSKY, A., GUERRERO, F. D. & SCOLES, G. A. (2007b) Proteomic profiling of Rhipicephalus (Boophilus) microplus midgut responses to infection with Babesia bovis. *Vet Parasitol*.

RADOMSKI, M. W., PALMER, R. M. & MONCADA, S. (1987) Comparative pharmacology of endothelium-derived relaxing factor, nitric oxide and prostacyclin in platelets. *Br J Pharmacol,* 92**,** 181-7.

RAMACHANDRA, R. N. & WIKEL, S. K. (1992) Modulation of host-immune responses by ticks (Acari: Ixodidae): effect of salivary gland extracts on host macrophages and lymphocyte cytokine production. *J Med Entomol,* 29**,** 818-26.

RAMAMOORTHI, N., NARASIMHAN, S., PAL, U., BAO, F., YANG, X. F., FISH, D., ANGUITA, J., NORGARD, M. V., KANTOR, F. S., ANDERSON, J. F., KOSKI, R. A. & FIKRIG, E. (2005) The Lyme disease agent exploits a tick protein to infect the mammalian host. *Nature,* 436**,** 573-7.

RANSON, H., PRAPANTHADARA, L. & HEMINGWAY, J. (1997) Cloning and characterization of two glutathione S-transferases from a DDT-resistant strain of Anopheles gambiae. *Biochem J,* 324 (Pt 1)**,** 97-102.

RAO, M. B., TANKSALE, A. M., GHATGE, M. S. & DESHPANDE, V. V. (1998) Molecular and biotechnological aspects of microbial proteases. *Microbiol Mol Biol Rev,* 62**,** 597-635.

RECHAV, Y., STRYDOM, W. J., CLARKE, F. C., BURGER, L. B., MACKIE, A. J. & FIELDEN, L. J. (1994) Isotopes as host blood markers to measure blood intake by feeding ticks (Acari: Ixodidae). *J.Med.Entomol.,* 31**,** 511-515.

REDLITZ, A., FOWLER, B. J., PLOW, E. F. & MILES, L. A. (1995) The role of an enolase-related molecule in plasminogen binding to cells. *Eur J Biochem,* 227**,** 407-15.

REGO, R. O., KOVAR, V., KOPACEK, P., WEISE, C., MAN, P., SAUMAN, I. & GRUBHOFFER, L. (2006) The tick plasma lectin, Dorin M, is a fibrinogen-related molecule. *Insect Biochem Mol Biol,* 36**,** 291-9.

REICH, C. I. & ZORZOPULOS, J. (1978) Boophilus microplus: characterization of larval proteases. *Exp Parasitol,* 44**,** 1-6.

RIBEIRO, J. M. (1987) Ixodes dammini: salivary anti-complement activity. *Exp Parasitol,* 64**,** 347-53.

RIBEIRO, J. M. (1995) Blood-feeding arthropods: live syringes or invertebrate pharmacologists? *Infect Agents Dis,* 4**,** 143-52.

RIBEIRO, J. M., ALARCON-CHAIDEZ, F., FRANCISCHETTI, I. M., MANS, B. J., MATHER, T. N., VALENZUELA, J. G. & WIKEL, S. K. (2006) An annotated catalog of salivary gland transcripts from Ixodes scapularis ticks. *Insect Biochem Mol Biol,* 36**,** 111-29.

RIBEIRO, J. M., MAKOUL, G. T., LEVINE, J., ROBINSON, D. R. & SPIELMAN, A. (1985) Antihemostatic, antiinflammatory, and immunosuppressive properties of the saliva of a tick, Ixodes dammini. *J Exp Med,* 161**,** 332-44.

RIBEIRO, J. M., MAKOUL, G. T. & ROBINSON, D. R. (1988) Ixodes dammini: evidence for salivary prostacyclin secretion. *J Parasitol,* 74**,** 1068-9.

RIBEIRO, J. M. & MATHER, T. N. (1998) Ixodes scapularis: salivary kininase activity is a metallo dipeptidyl carboxypeptidase. *Exp Parasitol,* 89**,** 213-21.

RIBEIRO, J. M. & SPIELMAN, A. (1986) Ixodes dammini: salivary anaphylatoxin inactivating activity. *Exp Parasitol,* 62**,** 292-7.

RIBEIRO, J. M., WEIS, J. J. & TELFORD, S. R., III (1990) Saliva of the tick Ixodes dammini inhibits neutrophil function. *Exp.Parasitol.,* 70**,** 382-388.

RIBERIO, J. (1987) Role of saliva in blood-feeding by arthropods. . *Annu Rev Entomol* 32**,** 463-478.

RIGHETTI, P. G. & BOSCHETTI, E. (2007) Sherlock Holmes and the proteome--a detective story. *FEBS J,* 274**,** 897-905.

RIGHETTI, P. G., CASTAGNA, A., HERBERT, B. & CANDIANO, G. (2005) How to bring the "unseen" proteome to the limelight via electrophoretic pre-fractionation techniques. *Biosci Rep,* 25**,** 3-17.

RIGHETTI, P. G., CASTAGNA, A., HERBERT, B., REYMOND, F. & ROSSIER, J. S. (2003) Prefractionation techniques in proteome analysis. *Proteomics,* 3**,** 1397-407.

ROBERTS, F. H. S. (1960) A systematic study of the australian species of the genus i xodes (Acarina:Ixodidae). *Aust.J.Zoology,* 8**,** 392-435.

ROSA DE LIMA, M. F., SANCHEZ FERREIRA, C. A., JOAQUIM DE FREITAS, D. R., VALENZUELA, J. G. & MASUDA, A. (2002) Cloning and partial characterization of a Boophilus microplus (Acari: Ixodidae) glutathione S-transferase. *Insect Biochem Mol Biol,* 32**,** 747-54.

ROSS, P. L., HUANG, Y. N., MARCHESE, J. N., WILLIAMSON, B., PARKER, K., HATTAN, S., KHAINOVSKI, N., PILLAI, S., DEY, S., DANIELS, S., PURKAYASTHA, S., JUHASZ, P., MARTIN, S., BARTLET-JONES, M., HE, F., JACOBSON, A. & PAPPIN, D. J. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics,* 3**,** 1154-69.

RUBIN, H. (1996) Serine protease inhibitors (SERPINS): where mechanism meets medicine. *Nat Med,* 2**,** 632-3.

SANDERS, H. R., EVANS, A. M., ROSS, L. S. & GILL, S. S. (2003) Blood meal induces global changes in midgut gene expression in the disease vector, Aedes aegypti. *Insect Biochem Mol Biol,* 33**,** 1105-22.

SANDERS, M. L., GLASS, G. E., NADELMAN, R. B., WORMSER, G. P., SCOTT, A. L., RAHA, S., RITCHIE, B. C., JAWORSKI, D. C. & SCHWARTZ, B. S. (1999) Antibody levels to recombinant tick calreticulin increase in humans after exposure to Ixodes scapularis (Say) and are correlated with tick engorgement indices. *Am.J.Epidemiol.,* 149**,** 777-784.

SANGAMNATDEJ, S., PAESEN, G. C., SLOVAK, M. & NUTTALL, P. A. (2002) A high affinity serotonin- and histamine-binding lipocalin from tick saliva. *Insect Mol Biol,* 11**,** 79-86.

SANTONI, V., MOLLOY, M. & RABILLOUD, T. (2000) Membrane proteins and proteomics: un amour impossible? *Electrophoresis,* 21**,** 1054-70.

SCHOELER, G. B., MANWEILER, S. A. & WIKEL, S. K. (1999) Ixodes scapularis: effects of repeated infestations with pathogen-free nymphs on macrophage and T lymphocyte cytokine responses of BALB/c and C3H/HeN mice. *Exp Parasitol,* 92**,** 239-48.

SEIXAS, A., LEAL, A. T., NASCIMENTO-SILVA, M. C., MASUDA, A., TERMIGNONI, C. & DA SILVA VAZ, I., JR. (2008) Vaccine potential of a tick vitellin-degrading enzyme (VTDCE). *Vet Immunol Immunopathol*.

SENNELS, L., SALEK, M., LOMAS, L., BOSCHETTI, E., RIGHETTI, P. G. & RAPPSILBER, J. (2007) Proteomic analysis of human blood serum using peptide library beads. *J Proteome Res,* 6**,** 4055-62.

SHARP, P. J., SMITH, D. R., BACH, W., WAGLAND, B. M. & COBON, G. S. (1991) Purified glutathione S-transferases from parasites as candidate protective antigens. *Int J Parasitol,* 21**,** 839-46.

SHAW, M. K., TILNEY, L. G. & MCKEEVER, D. J. (1993) Tick salivary gland extract and interleukin-2 stimulation enhance susceptibility of lymphocytes to infection by Theileria parva sporozoites. *Infect Immun,* 61**,** 1486-95.

SHIMOKAWA, H., FLAVAHAN, N. A., LORENZ, R. R. & VANHOUTTE, P. M. (1988) Prostacyclin releases endothelium-derived relaxing factor and potentiates its action in coronary arteries of the pig. *Br J Pharmacol,* 95**,** 1197-203.

SIMO, C., BACHI, A., CATTANEO, A., GUERRIER, L., FORTIS, F., BOSCHETTI, E., PODTELEJNIKOV, A. & RIGHETTI, P. G. (2008) Performance of combinatorial peptide libraries in capturing the low-abundance proteome of red blood cells. 1. Behavior of mono- to hexapeptides. *Anal Chem,* 80**,** 3547-56.

SKALLOVA, A., IEZZI, G., AMPENBERGER, F., KOPF, M. & KOPECKY, J. (2008) Tick saliva inhibits dendritic cell migration, maturation, and function while promoting development of Th2 responses. *J Immunol,* 180**,** 6186-92.

SOJKA, D., FRANTA, Z., HORN, M., HAJDUSEK, O., CAFFREY, C. R., MARES, M. & KOPACEK, P. (2008) Profiling of proteolytic enzymes in the gut of the tick Ixodes ricinus reveals an evolutionarily conserved network of aspartic and cysteine peptidases. *Parasit Vectors,* 1**,** 7.

SORGINE, M. H., LOGULLO, C., ZINGALI, R. B., PAIVA-SILVA, G. O., JULIANO, L. & OLIVEIRA, P. L. (2000) A heme-binding aspartic proteinase from the eggs of the hard tick Boophilus microplus. *J Biol Chem,* 275**,** 28659-65.

STARK, K. R. & JAMES, A. A. (1996) Anticoagulants in Vector Arthropods. *Parasitology Today,* 12**,** 430-437.

STARK, K. R. & JAMES, A. A. (1998) Isolation and characterization of the gene encoding a novel factor Xa-directed anticoagulant from the yellow fever mosquito, Aedes aegypti. *J Biol Chem,* 273**,** 20802-9.

STONE, B. F., BINNINGTON, K. C., GAUCI, M. & AYLWARD, J. H. (1989) Tick/host interactions for Ixodes holocyclus: role, effects, biosynthesis and nature of its toxic and allergenic oral secretions. *Experimental & Applied Acarology,* 7**,** 59-69.

STONE, B. F., BINNINGTON, K. C., SAUER, J. R. & HAIR, J. A. (1986) The paralyzing toxin and other immunogens of the tick Ixodes holocyclus and the role of the salivary gland in their biosynthesis. *Morphology,Physiology and Behavioural Biology of Ticks.* Chichester, Ellis Harwood.

SUGINO, M., IMAMURA, S., MULENGA, A., NAKAJIMA, M., TSUDA, A., OHASHI, K. & ONUMA, M. (2003) A serine proteinase inhibitor (serpin) from ixodid tick Haemaphysalis longicornis; cloning and preliminary assessment of its suitability as a candidate for a tick vaccine. *Vaccine,* 21**,** 2844-51.

SUN, W., GAO, S., WANG, L., CHEN, Y., WU, S., WANG, X., ZHENG, D. & GAO, Y. (2006) Microwave-assisted protein preparation and enzymatic digestion in proteomics. *Mol Cell Proteomics,* 5**,** 769-76.

SWANEY, D. L., MCALISTER, G. C., WIRTALA, M., SCHWARTZ, J. C., SYKA, J. E. & COON, J. J. (2007) Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors. *Anal Chem,* 79**,** 477-85.

SYKA, J. E. P., COON, J. J., SCHROEDER, M. J., SHABANOWITZ, J. & HUNT, D. F. (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *PNAS,* 101**,** 9528-9533.

TELLAM, R. L., KEMP, D., RIDING, G., BRISCOE, S., SMITH, D., SHARP, P., IRVING, D. & WILLADSEN, P. (2002) Reduced oviposition of Boophilus microplus feeding on sheep vaccinated with vitellin. *Vet Parasitol,* 103**,** 141-56.

THOMPSON, D. M., KHALIL, S. M., JEFFERS, L. A., SONENSHINE, D. E., MITCHELL, R. D., OSGOOD, C. J. & MICHAEL ROE, R. (2007) Sequence and the developmental and tissue-specific regulation of the first complete vitellogenin messenger RNA from ticks responsible for heme sequestration. *Insect Biochem Mol Biol,* 37**,** 363-74.

THULASIRAMAN, V., LIN, S., GHEORGHIU, L., LATHROP, J., LOMAS, L., HAMMOND, D. & BOSCHETTI, E. (2005) Reduction of the concentration difference of proteins in biological liquids using a library of combinatorial ligands. *Electrophoresis,* 26**,** 3561-71.

THURN, M. J. (1994) Tick toxinology: Isolation and characterisation of the toxin from the Australian Paralysis Tick, *Ixodes holocyclus*. Sydney, University of Technology, Sydney.

TRIMNELL, A. R., DAVIES, G. M., LISSINA, O., HAILS, R. S. & NUTTALL, P. A. (2005) A cross-reactive tick cement antigen is a candidate broad-spectrum tick vaccine. *Vaccine,* 23**,** 4329-41.

TRIMNELL, A. R., HAILS, R. S. & NUTTALL, P. A. (2002) Dual action ectoparasite vaccine targeting 'exposed' and 'concealed' antigens. *Vaccine,* 20**,** 3560-8.

TSUDA, A., MULENGA, A., SUGIMOTO, C., NAKAJIMA, M., OHASHI, K. & ONUMA, M. (2001) cDNA cloning, characterization and vaccine effect analysis of Haemaphysalis longicornis tick saliva proteins. *Vaccine,* 19**,** 4287-4296.

TUCHOLSKA, M., BOWDEN, P., JACKS, K., ZHU, P., FURESZ, S., DUMBROVSKY, M. & MARSHALL, J. (2009) Human Serum Proteins Fractionated by Preparative Partition Chromatography Prior to LC-ESI-MS/MS. *J Proteome Res,* 8**,** 1143-55.

TURAPOV, O. A., MUKAMOLOVA, G. V., BOTTRILL, A. R. & PANGBURN, M. K. (2008) Digestion of Native Proteins for Proteomics Using a Thermocycler. *Anal Chem*.

UNTALAN, P. M., GUERRERO, F. D., HAINES, L. R. & PEARSON, T. W. (2005) Proteome analysis of abundantly expressed proteins from unfed larvae of the cattle tick, Boophilus microplus. *Insect Biochem Mol Biol,* 35**,** 141-51.

VALENZUELA, J. G., FRANCISCHETTI, I. M., PHAM, V. M., GARFIELD, M. K., MATHER, T. N. & RIBEIRO, J. M. (2002) Exploring the sialome of the tick Ixodes scapularis. *J Exp Biol,* 205**,** 2843-64.

VAN DE LOCHT, A., STUBBS, M. T., BODE, W., FRIEDRICH, T., BOLLSCHWEILER, C., HOFFKEN, W. & HUBER, R. (1996) The ornithodorin-thrombin crystal structure, a key to the TAP enigma? *EMBO J,* 15**,** 6011-7.

VAUGHAN, J. A. & AZAD, A. F. (1988) Passage of host immunoglobulin G from blood meal into hemolymph of selected mosquito species (Diptera: Culicidae). *J Med Entomol,* 25**,** 472-4.

VENNESTROM, J. & JENSEN, P. M. (2007) Ixodes ricinus: the potential of two-dimensional gel electrophoresis as a tool for studying host-vector-pathogen interactions. *Exp Parasitol,* 115**,** 53-8.

WANG, H. & NUTTALL, P. A. (1994) Excretion of host immunoglobulin in tick saliva and detection of IgG-binding proteins in tick haemolymph and salivary glands. *Parasitology,* 109 (Pt 4)**,** 525-30.

WANG, H. & NUTTALL, P. A. (1995a) Immunoglobulin-G binding proteins in the ixodid ticks, Rhipicephalus appendiculatus, Amblyomma variegatum and Ixodes hexagonus. *Parasitology,* 111 (Pt 2)**,** 161-5.

WANG, H. & NUTTALL, P. A. (1995b) Immunoglobulin G binding proteins in male Rhipicephalus appendiculatus ticks. *Parasite Immunol,* 17**,** 517-24.

WANG, H. & NUTTALL, P. A. (1999) Immunoglobulin-binding proteins in ticks: new target for vaccine development against a blood-feeding parasite. *Cell Mol Life Sci,* 56**,** 286-95.

WANG, N., XIE, C., YOUNG, J. B. & LI, L. (2009) Off-line two-dimensional liquid chromatography with maximized sample loading to reversed-phase liquid chromatography-electrospray ionization tandem mass spectrometry for shotgun proteome analysis. *Anal Chem,* 81**,** 1049-60.

WANG, X., COONS, L. B., TAYLOR, D. B., STEVENS, S. E., JR. & GARTNER, T. K. (1996) Variabilin, a novel RGD-containing antagonist of glycoprotein IIb-IIIa and platelet

aggregation inhibitor from the hard tick Dermacentor variabilis. *J Biol Chem,* 271**,** 17785-90.

WARIDEL, P., FRANK, A., THOMAS, H., SURENDRANATH, V., SUNYAEV, S., PEVZNER, P. & SHEVCHENKO, A. (2007) Sequence similarity-driven proteomics in organisms with unknown genomes by LC-MS/MS and automated de novo sequencing. *Proteomics,* 7**,** 2318-29.

WASMUTH, J. D. & BLAXTER, M. L. (2004) prot4EST: translating expressed sequence tags from neglected genomes. *BMC Bioinformatics,* 5**,** 187.

WATTENBERG, A., ORGAN, A. J., SCHNEIDER, K., TYLDESLEY, R., BORDOLI, R. & BATEMAN, R. H. (2002) Sequence dependent fragmentation of peptides generated by MALDI quadrupole time-of-flight (MALDI Q-TOF) mass spectrometry and its implications for protein identification. *J Am Soc Mass Spectrom,* 13**,** 772-83.

WAXMAN, L. & CONNOLLY, T. M. (1993) Isolation of an inhibitor selective for collagen-stimulated platelet aggregation from the soft tick Ornithodoros moubata. *J Biol Chem,* 268**,** 5445-9.

WAXMAN, L., SMITH, D. E., ARCURI, K. E. & VLASUK, G. P. (1990) Tick anticoagulant peptide (TAP) is a novel inhibitor of blood coagulation factor Xa. *Science,* 248**,** 593-6.

WEI, C. L., KAINUMA, M. & HERSHEY, J. W. (1995) Characterization of yeast translation initiation factor 1A and cloning of its essential gene. *J Biol Chem,* 270**,** 22788-94.

WEI, S. H., CLARK, A. G. & SYVANEN, M. (2001) Identification and cloning of a key insecticide-metabolizing glutathione S-transferase (MdGST-6A) from a hyper insecticide-resistant strain of the housefly Musca domestica. *Insect Biochem Mol Biol,* 31**,** 1145-53.

WEIGELT, H., ADDICKS, K., HAUCK, G. & LUBBERS, D. W. (1979) Vitalmicroscopic studies in regard to the role of intraendothelian reactive structures in the inflammatory process. *Bibl Anat***,** 11-20.

WEISS, B. L. & KAUFMAN, W. R. (2004) Two feeding-induced proteins from the male gonad trigger engorgement of the female tick Amblyomma hebraeum. *Proc Natl Acad Sci U S A,* 101**,** 5874-9.

WIELSCH, N., THOMAS, H., SURENDRANATH, V., WARIDEL, P., FRANK, A., PEVZNER, P. & SHEVCHENKO, A. (2006) Rapid validation of protein identifications with the borderline statistical confidence via de novo sequencing and MS BLAST searches. *J Proteome Res,* 5**,** 2448-56.

WIKEL, S. K. (1996) Host immunity to ticks. *Annu.Rev.Entomol.,* 41**,** 1-22.

WIKEL, S. K. (1999) Tick modulation of host immunity: an important factor in pathogen transmission. *Int.J.Parasitol.,* 29**,** 851-859.

WIKEL, S. K., RAMACHANDRA, R. N., BERGMAN, D. K., BURKOT, T. R. & PIESMAN, J. (1997) Infestation with pathogen-free nymphs of the tick Ixodes scapularis induces host resistance to transmission of Borrelia burgdorferi by ticks. *Infect.Immun.,* 65**,** 335-338.

WILKINSON, B. & GILBERT, H. F. (2004) Protein disulfide isomerase. *Biochim Biophys Acta,* 1699**,** 35-44.

WILLIAMS, T. J. (1978) The role of prostaglandins in inflammation. *Ann R Coll Surg Engl,* 60**,** 198-201.

WILM, M., SHEVCHENKO, A., HOUTHAEVE, T., BREIT, S., SCHWEIGERER, L., FOTSIS, T. & MANN, M. (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature,* 379**,** 466-9.

YAN, Y., LIU, S., SONG, G., XU, Y. & DISSOUS, C. (2005) Characterization of a novel vaccine candidate and serine proteinase inhibitor from Schistosoma japonicum (Sj serpin). *Vet Parasitol,* 131**,** 53-60.

YANG, W., GOBERT, G. N. & MCMANUS, D. P. (1997) Oral vaccination of mice with recombinant Schistosoma japonicum proteins induces specific anti-parasite antibodies and damage to adult worms after a challenge infection. *Int J Parasitol,* 27**,** 843-53.

YERGEY, A. L., COORSSEN, J. R., BACKLUND, P. S., JR., BLANK, P. S., HUMPHREY, G. A., ZIMMERBERG, J., CAMPBELL, J. M. & VESTAL, M. L. (2002) De novo sequencing of peptides using MALDI/TOF-TOF. *J Am Soc Mass Spectrom,* 13**,** 784-91.

YOU, M., XUAN, X., TSUJI, N., KAMIO, T., IGARASHI, I., NAGASAWA, H., MIKAMI, T. & FUJISAKI, K. (2001) Molecular characterization of a troponin I-like protein from the hard tick Haemaphysalis longicornis. *Insect Biochem Mol Biol,* 32**,** 67-73.

ZEIDNER, N., MBOW, M. L., DOLAN, M., MASSUNG, R., BACA, E. & PIESMAN, J. (1997) Effects of Ixodes scapularis and Borrelia burgdorferi on modulation of the host immune response: induction of a TH2 cytokine response in Lyme disease-susceptible (C3H/HeJ) mice but not in disease-resistant (BALB/c) mice. *Infect Immun,* 65**,** 3100-6.

ZHAO, Z. & RUSSELL, P. J. (2003) Trypsin activity assay in substrate-specific one- and two-dimensional gels: a powerful method to separate and characterize novel proteases in active form in biological samples. *Electrophoresis,* 24**,** 3284-8.

ZIESKE, L. R. (2006) A perspective on the use of iTRAQ reagent technology for protein complex and profiling studies. *J Exp Bot,* 57**,** 1501-8.

ZUBAREV, R. A., ZUBAREV, A. R. & SAVITSKI, M. M. (2008) Electron Capture/Transfer versus Collisionally Activated/Induced Dissociations: Solo or Duet? *J Am Soc Mass Spectrom,* 19**,** 753-61.