

Kaavya Krishna Kumar,<sup>‡</sup>  
David A. Jacques,<sup>‡</sup>§ J. Mitchell  
Guss<sup>a</sup> and David A. Gell<sup>b\*</sup>

<sup>a</sup>School of Molecular Bioscience, University of  
Sydney, Sydney, NSW 2006, Australia, and

<sup>b</sup>Menzies Research Institute, University of  
Tasmania, Hobart, TAS 7000, Australia

‡ Current address: Walter and Eliza Hall  
Institute of Medical Research, Australia.

§ Current address: MRC Laboratory of  
Molecular Biology, Francis Crick Avenue,  
Cambridge Biomedical Campus, Cambridge  
CB2 0QH, England.

Correspondence e-mail: david.gell@utas.edu.au

Received 19 April 2014

Accepted 26 May 2014

**PDB reference:** human  $\alpha$ -haemoglobin  
complexed with the first NEAT domain of IsdH  
from *S. aureus*, 3s48

## The structure of $\alpha$ -haemoglobin in complex with a haemoglobin-binding domain from *Staphylococcus aureus* reveals the elusive $\alpha$ -haemoglobin dimerization interface

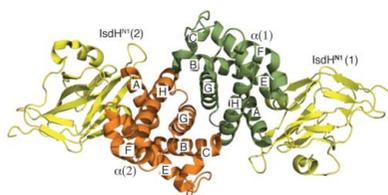
Adult haemoglobin (Hb) is made up of two  $\alpha$  and two  $\beta$  subunits. Mutations that reduce expression of the  $\alpha$ - or  $\beta$ -globin genes lead to the conditions  $\alpha$ - or  $\beta$ -thalassaemia, respectively. Whilst both conditions are characterized by anaemia of variable severity, other details of their pathophysiology are different, in part owing to the greater stability of the  $\beta$  chains that is conferred through  $\beta$  self-association. In contrast,  $\alpha$  subunits interact weakly, and in the absence of stabilizing quaternary interactions the  $\alpha$  chain ( $\alpha$ ) is prone to haem loss and denaturation. The molecular contacts that confer weak self-association of  $\alpha$  have not been determined previously. Here, the first structure of an  $\alpha_2$  homodimer is reported in complex with one domain of the Hb receptor from *Staphylococcus aureus*. The  $\alpha_2$  dimer interface has a highly unusual, approximately linear, arrangement of four His side chains within hydrogen-bonding distance of each other. Some interactions present in the  $\alpha_1\beta_1$  dimer interface of native Hb are preserved in the  $\alpha_2$  dimer. However, a marked asymmetry is observed in the  $\alpha_2$  interface, suggesting that steric factors limit the number of stabilizing interactions that can form simultaneously across the interface.

### 1. Introduction

The major form of adult haemoglobin (HbA) is a tetramer of two  $\alpha$  Hb ( $\alpha$ ) and two  $\beta$  Hb ( $\beta$ ) subunits (Perutz *et al.*, 1960). Each subunit interacts with its non-identical partner through two different interfaces. The  $\alpha_1\beta_1$  (and  $\alpha_2\beta_2$ ) interface is extremely stable ( $K_a = \sim 5 \times 10^{11} M^{-1}$ ; Valdes & Ackers, 1977b) and remains essentially unchanged by oxygen binding. The  $\alpha_1\beta_2$  (and  $\alpha_2\beta_1$ ) interface undergoes a reorganization coupled to oxygen binding involving a  $\sim 15^\circ$  rotation of the two  $\alpha\beta$  dimers.

Mutations that reduce expression of the  $\alpha$ - or  $\beta$ -globin genes lead to the condition  $\alpha$ - or  $\beta$ -thalassaemia, respectively (Nathan & Gunn, 1966). Insufficiency of  $\alpha$  or  $\beta$  production leads to anaemia that can range from mild to fatal (Clegg & Weatherall, 1976). In addition to the insufficient production of normal Hb, thalassaemia mutations lead to a relative excess of the partner Hb chain, which has cytotoxic effects. Free Hb chains are unstable compared with native Hb owing to the absence of native quaternary interactions. Quaternary interactions reduce chain unfolding and precipitation and reduce haem loss (Hargrove *et al.*, 1997), which is associated with the production of harmful reactive species catalyzed by haem iron (Rifkind *et al.*, 2004; Rachmilewitz & Schrier, 2001). Notably, free globin chains undergo self-association (Bucci *et al.*, 1965) and differences in the self-association properties of  $\alpha$  or  $\beta$  contribute different stabilities and different pathologies in  $\alpha$ - or  $\beta$ -thalassaemia.

In  $\alpha$ -thalassaemia, excess  $\beta$  self-associates into stable  $\beta_4$  tetramers (HbH). The dimer–tetramer association is more rapid than the monomer–dimer interaction (Philo *et al.*, 1988), giving the appearance of a monomer–tetramer reaction ( $K_a = \sim 4 \times 10^{16} M^{-3}$ ; Valdes & Ackers, 1977a, 1978).  $\beta_4$  tetramers exhibit noncooperative and high-affinity O<sub>2</sub> binding and therefore do not effectively release O<sub>2</sub> to the tissues. Formation of  $\beta_4$  achieves a 20-fold stabilization against haem



**Table 1**

Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

X-ray source	Australian Synchrotron MX2
Wavelength (Å)	0.9537
Temperature (°C)	−173
Detector	ADSC Quantum 315r CCD
Total rotation range (°)	180
Space group	$I4_1$
Unit-cell parameters (Å)	$a = b = 115.4, c = 142.4$
Resolution range (Å)	50.00–3.05 (3.13–3.05)
No. of observations	40834
No. of unique reflections	17737
Completeness (%)	99.5 (99.4)
Multiplicity	2.3 (2.2)
$\langle I/\sigma(I) \rangle$	12.9 (2.1)
Wilson $B$ value (Å <sup>2</sup> )	119.3
$R_{\text{merge}}^{\dagger}$	0.265 (0.635)
Reflections in working set‡	16833 (1244)
Reflections in test set	904 (61)
Contents of the asymmetric unit	$2\alpha + 2\text{IsdH}^{\text{N1}}$
No. of atoms	
Total (non-H)	4486
Protein	4398
Water§	2
Others (haem)	86
$R_{\text{cryst}}$	0.242 (0.358)
$R_{\text{free}}$	0.268 (0.369)
R.m.s.d., bond lengths (Å)	0.004
R.m.s.d., bond angles (°)	0.639
$\langle B \rangle$ (Å <sup>2</sup> )	
IsdH <sup>N1</sup> (2), chain $A$	82
IsdH <sup>N1</sup> (1), chain $B$	249
Chain $C$	114
Chain $D$	93
Cruickshank's DPI¶ (Å)	0.4
PDB code	3s48

<sup>†</sup>  $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ . <sup>‡</sup> Friedel pairs were kept separate for scaling, but were merged by *REFMAC5* during structure refinement. <sup>§</sup> A single water molecule was placed in the distal haem pocket of each globin chain, coordinating the haem  $\text{Fe}^{\text{III}}$ . <sup>¶</sup> Diffraction precision indicator as output from *REFMAC5*.

loss compared with  $\beta$  monomers, and consequently HbA and  $\beta_4$  lose haem from the  $\beta$  chains at comparable rates (Hargrove *et al.*, 1996, 1997). Nevertheless,  $\beta_4$  protein slowly forms intracellular precipitates, causing oxidative damage to the membrane cytoskeleton of circulating red blood cells, which manifests as haemolytic anaemia (Advani *et al.*, 1992). In foetal development the foetal  $\beta$  chain,  $\gamma$ , similarly forms  $\gamma_4$  tetramers (Hb Bart's; Kidd *et al.*, 2001).

Compared with  $\beta$ ,  $\alpha$  undergoes comparatively weak self-association to form dimers ( $K_a = \sim 5\text{--}9 \times 10^3 \text{ M}^{-1}$ ; Valdes & Ackers, 1977a; Windsor *et al.*, 1992), coupled with a  $\sim 25$ -fold increase in haem loss compared with HbA or  $\beta_4$  (Hargrove *et al.*, 1997). As a result,  $\alpha$  precipitation and oxidative damage occur early in erythroid cell development (Polliack *et al.*, 1974). Damage to, and apoptosis of, erythroid precursors leads to compensatory proliferation of erythroid progenitors in the bone marrow, liver and spleen, with consequent disruption to the structure and function of these organs (a condition known as ineffective erythropoiesis; Rachmilewitz & Schrier, 2001).

In addition to its role in disease, globin chain self-association has significance for the Hb assembly pathway. In normal erythropoiesis,  $\alpha$  is produced in a slight excess over  $\beta$  (Vasseur *et al.*, 2011) and in this way  $\alpha$  can act as a chaperone for  $\beta$  by outcompeting the formation of nonfunctional  $\beta_4$  (Miele *et al.*, 2001). In turn,  $\alpha$  is stabilized by its own chaperone,  $\alpha$  Hb stabilizing protein (AHSP; Kihm *et al.*, 2002; Gell *et al.*, 2002). A moderately strong interaction between AHSP and oxygenated  $\alpha$  ( $K_a = 5 \times 10^7 \text{ M}^{-1}$ ) out-competes weak  $\alpha_2$  self-association and prevents  $\alpha$  precipitation, while at the same time permitting rapid kinetics of  $\alpha$  release to  $\beta$  (Mollan *et al.*, 2012).

The structures of  $\beta_4$  and  $\gamma_4$  have been determined (Borgstahl *et al.*, 1994a,b; Kidd *et al.*, 2001), allowing Kidd and coworkers to identify features at the N- and C-termini of  $\beta$  and  $\gamma$  that help to explain why these chains form tetramers but  $\alpha$  does not. Here, we present the first crystal structure containing an  $\alpha_2$  dimer and identify features at the  $\alpha_2$  dimer interface that contribute to weak dimerization. The crystals contain  $\alpha$  bound to the first near-iron transporter (NEAT) domain of iron-regulated surface determinant H (IsdH) from *Staphylococcus aureus* (Pilpa *et al.*, 2006), and were obtained during investigation of Hb capture by this bacterium.

## 2. Materials and methods

### 2.1. Protein production

Hb was purified from blood as reported previously (Gell *et al.*, 2002). The splitting of carbonmonoxy-liganded Hb into  $\alpha$  and  $\beta$  chains was carried out using the well established *p*-hydroxymercuribenzoate (PMB) method (Bucci & Frontice, 1965). The resulting  $\alpha^{\text{PMB}}$  and  $\beta^{\text{PMB}}$  chains were separated over DEAE Sepharose resin (GE Healthcare). PMB was removed from  $\alpha$  by overnight incubation at 4°C with dithiothreitol followed by purification on an SP Sepharose column and storage at  $-80^\circ\text{C}$ . For protein crystallography, purified carbonmonoxy  $\alpha$  was converted to the oxygenated form by passing a pure stream of oxygen over a protein solution held on ice and illuminated with a focused beam from a 50 W halogen lamp. The oxygenated  $\alpha$  was converted to ferric (met)  $\alpha$  by addition of excess potassium ferricyanide in 20 mM sodium phosphate pH 7.0. The reaction was monitored to completion by UV-visible spectroscopy and  $\alpha$  was isolated using a Sephadex G-25 column. The concentration of  $\alpha$  was estimated from the concentration of the associated haem group, measured at 390 nm, from unfolded globin samples in 6 M guanidine hydrochloride (extinction coefficient  $37\,800 \text{ M}^{-1} \text{ cm}^{-1}$ ).

IsdH<sup>N1</sup> (IsdH residues 86–229) from *S. aureus* strain TCH1516 was cloned into pET-15b (Novagen) for expression with an N-terminal hexa-His tag. IsdH<sup>N1</sup> was expressed and purified as described previously (Pilpa *et al.*, 2006) to yield a final product with the additional N-terminal sequence MGSSHHHHHSSGLVPRGSHM.

### 2.2. Crystallization

Ferric  $\alpha$  at  $5 \text{ mg ml}^{-1}$  (20 mM sodium phosphate pH 7.0) was mixed with one molar equivalent of IsdH<sup>N1</sup> prepared in the same buffer and crystallization screening was performed by the hanging-drop vapour-diffusion method in 96-well plates using a Mosquito nanolitre liquid-handling robot (TTP LabTech). Protein complex (400 nl) was mixed with commercially available crystallization screens (The JCSG+, PACT and Classics Suites, Qiagen) in a 1:1 ratio and was incubated at 25°C. Crystals of ferric  $\alpha$ -IsdH<sup>N1</sup> appeared within a week in 0.2 M sodium sulfate, 0.1 M bis-tris propane, 20% PEG 3350 pH 6.5. On further optimization, single crystals appeared in 0.2 M sodium sulfate, 0.1 M bis-tris propane, 16% PEG 3350 pH 6.5. A single crystal was transferred into a cryoprotectant solution consisting of 30%(v/v) glycerol in the buffer from the crystallization condition and flash-cooled in a nitrogen stream ( $-173^\circ\text{C}$ ).

### 2.3. Data collection and processing

Data were recorded (over a  $\varphi$  range of  $180^\circ$ ) to a resolution of 3.05 Å using an ADSC Quantum 315r detector on the MX2 beamline of the Australian Synchrotron at a wavelength of 0.95370 Å using the *Blu-Ice* control system (McPhillips *et al.*, 2002). X-ray diffraction data

were indexed and scaled using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

### 2.4. Structure solution and refinement

Molecular replacement (MR) was performed using the IsdH<sup>N1</sup> and  $\alpha$  subunit (minus the haem group) of AHSP- $\alpha$ -IsdH<sup>N1</sup> (PDB entry 3ovu, D. A. Jacques, K. Krishna Kumar, T. T. Caradoc-Davies, D. B. Langley, J. P. Mackay, J. M. Guss & D. A. Gell, unpublished work) as

search models. A unique solution was found by *Phaser* (McCoy, 2007) in the tetragonal space group *I*<sub>4</sub> with two molecules of IsdH<sup>N1</sup> and two molecules of  $\alpha$  in the crystallographic asymmetric unit. The structure was refined using *REFMAC5* (Murshudov *et al.*, 2011), with manual map inspection and model building being performed in *Coot* (Emsley *et al.*, 2010). The quality of the model was regularly checked for steric clashes, incorrect stereochemistry and rotamer outliers using *MolProbity* (Chen *et al.*, 2010). All structural figures were produced using *PyMOL* (Schrödinger, <http://www.pymol.org>). Refined atomic coordinates and experimental structure factors have been deposited in the Protein Data Bank (PDB entry 3s48). Data-collection and refinement statistics are given in Table 1.

### 2.5. Naming conventions

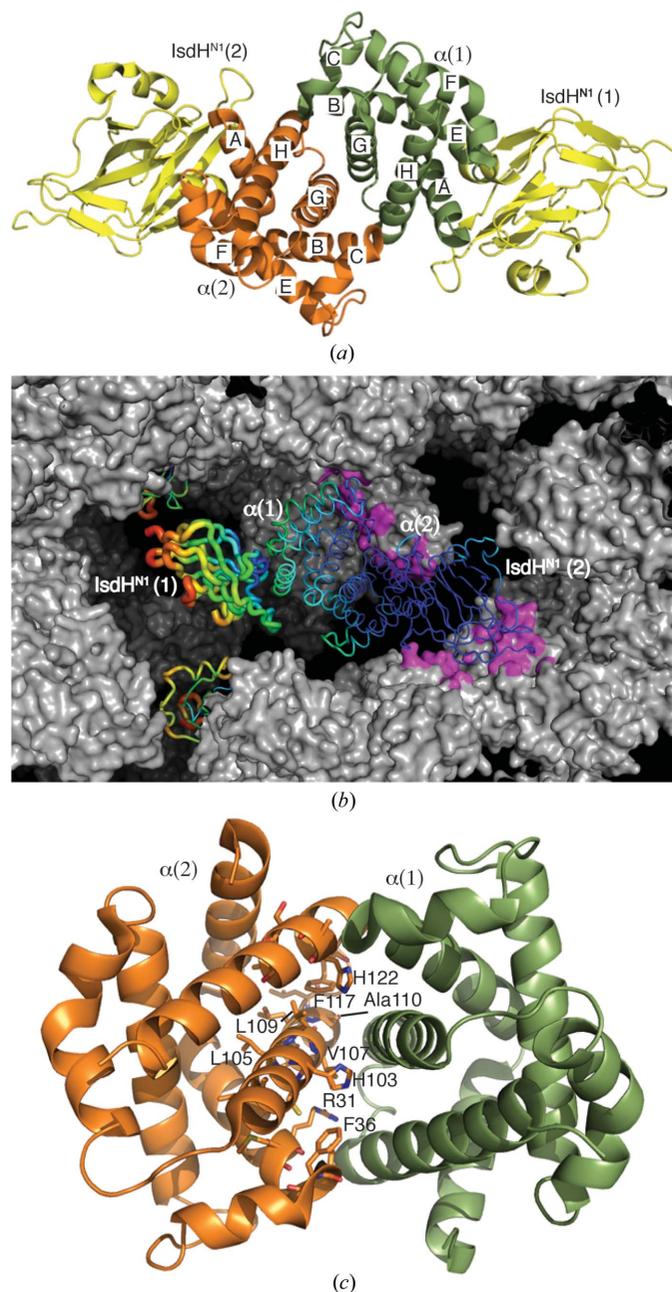
In addition to the usual convention of numbering amino-acid residues sequentially from the N-terminus of a peptide, it is conventional in the globin field to identify the amino-acid position relative to the structure of sperm whale myoglobin. For example, F8 indicates a residue that is functionally equivalent to the eighth residue in helix F of sperm whale myoglobin (the haem-ligating proximal His). Within the HbA tetramer, interfaces between the four subunits are designated using upper-case numerals according to convention in the field, *e.g.* the  $\alpha$ 1 subunit contributes to the  $\alpha$ 1 $\beta$ 1,  $\alpha$ 1 $\beta$ 2 and  $\alpha$ 1 $\alpha$ 2 subunit interfaces. The oligomeric state of globin assemblies is designated using subscript numerals by convention, *e.g.*  $\alpha$ <sub>2</sub> $\beta$ <sub>2</sub> refers to a complex of two  $\alpha$  and two  $\beta$  subunits (that is, the HbA tetramer).

## 3. Results and discussion

### 3.1. Crystal structure of the $\alpha$ -IsdH<sup>N1</sup> complex

The structure of  $\alpha$ -IsdH<sup>N1</sup> was determined to a resolution of 3.05 Å and refined to an *R*<sub>cryst</sub> and *R*<sub>free</sub> of 0.242 and 0.268, respectively. The crystallographic asymmetric unit contains an  $\alpha$ <sub>2</sub> dimer, with each  $\alpha$  monomer bound to one molecule of IsdH<sup>N1</sup> (Fig. 1*a*). One IsdH<sup>N1</sup> domain [IsdH<sup>N1</sup>(1)] does not make any intermolecular crystal contacts (Fig. 1*b*). Disorder in this region is reflected in substantially higher *B* factors (Fig. 1*b* and Table 1, chain *B*) and the relatively high *R*<sub>merge</sub> value.

The  $\alpha$ -IsdH<sup>N1</sup> interface is not significantly different from that observed in a previous structure in which IsdH<sup>N1</sup> is bound to an  $\alpha\beta$  dimer (PDB entry 3szk; r.m.s.d. of 0.2 Å for 141 C $\alpha$  atoms; Krishna Kumar *et al.*, 2011) and is not discussed further. The structure of the  $\alpha$ <sub>2</sub> dimer has not previously been observed and reveals that the  $\alpha$ <sub>2</sub> dimer interface is equivalent to the  $\alpha$ 1 $\beta$ 1 dimer interface in HbA. Residues at, or close to, the  $\alpha$ <sub>2</sub> dimer interface have previously been identified by NMR chemical shift perturbation (Fig. 1*c*; Dickson *et al.*, 2013), confirming that the crystal structure accurately represents the  $\alpha$ <sub>2</sub> interaction in solution. In a recent analysis of 113 crystal structures of protein dimers from the PDB (Chen *et al.*, 2013), the weakest interaction (between a bromodomain and an acetylated peptide from histone H3) had a *K*<sub>a</sub> of  $\sim 1 \times 10^3 M^{-1}$  (VanDemark *et al.*, 2007). A number of electron-transport complexes have successfully been crystallized; these typically display *K*<sub>a</sub> in the range  $10^3$ – $10^6 M^{-1}$  and are notable as physiologically important ‘weak’ interactions. The *K*<sub>a</sub> of the  $\alpha$ <sub>2</sub> dimer interface is estimated at  $5$ – $9 \times 10^3 M^{-1}$  (Valdes & Ackers, 1977*a*; Windsor *et al.*, 1992), placing this at the weak-interaction end of the spectrum of crystallized protein complexes.



**Figure 1**  
The structure of the  $\alpha$ -IsdH<sup>N1</sup> complex. (*a*) Two  $\alpha$ -IsdH<sup>N1</sup> dimers are present in the crystallographic asymmetric unit. (*b*) The crystal packing is shown. One complete ( $\alpha$ -IsdH<sup>N1</sup>)<sub>2</sub> complex in the centre of the figure is represented as a tube (wider tubes and hot colours indicate higher C $\alpha$ -atom *B* factors). Crystal contacts are shown as magenta patches on the grey surface of symmetry-related molecules. IsdH<sup>N1</sup>(1) projects into a cavity in the crystal and has substantially higher *B* factors. Its nearest neighbours are other symmetry-related IsdH<sup>N1</sup>(1) subunits (also shown in tube representation). (*c*) Residues close to the  $\alpha$  dimerization interface identified by NMR (Dickson *et al.*, 2013) are shown in stick representation on the  $\alpha$ (2) chain.

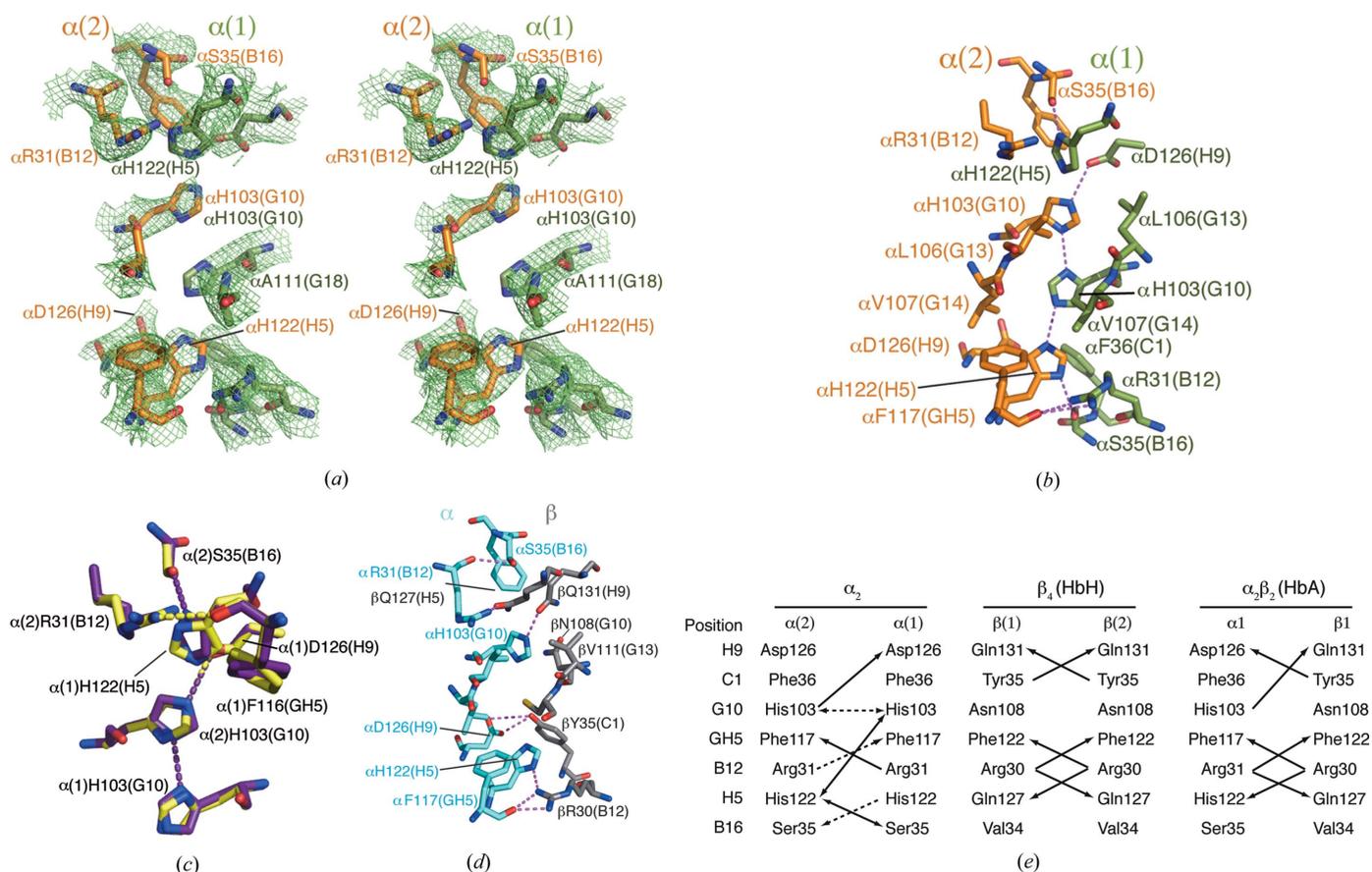
**Table 2**

*MolProbity* (Chen *et al.*, 2010) evaluation of 3s48 and 3s48 re-refined by *PDB\_REDO* (Joosten *et al.*, 2012).

	3s48	3s48 re-refined by <i>PDB_REDO</i>
$R/R_{\text{free}}$	0.2420/0.2680	0.2218/0.2452
Poor rotamers (goal <1%)	1 (0.21%)	47 (9.67%)
Ramachandran outliers (goal <0.05%)	1 (0.18%)	4 (0.72%)
Ramachandran favoured (goal >98%)	521 (94.04%)	513 (92.60%)
$C^\beta$ deviations > 0.25 Å (goal 0)	0 (0.00%)	6 (1.11%)
Bad backbone bonds (goal 0%)	0/4611 (0.00%)	2/4650 (0.04%)
Bad backbone angles (goal <0.1%)	0/6301 (0.00%)	6/6351 (0.09%)

### 3.2. Buried His residues at the $\alpha_2$ interface

A highly unusual arrangement of four buried His side chains, His103(G10) and His122(H5) from each chain, interdigitate across the  $\alpha_2$  dimer interface (Fig. 2*a*). Optimized hydrogen-bonding patterns were investigated using *MolProbity* (Chen *et al.*, 2010), *PDB2PQR* (Dolinsky *et al.*, 2004) and *WHAT IF* (Vriend, 1990). These analyses show hydrogen-bond interactions between  $N^{\delta 1}$  of  $\alpha(2)$ His103 and  $N^{\delta 1}$  of  $\alpha(1)$ His103 and between  $N^{\epsilon 2}$  of  $\alpha(1)$ His103 and  $N^{\epsilon 2}$  of  $\alpha(2)$ His122, forming a novel ‘histidine-zipper’ arrangement (Fig. 2*b*).


**Figure 2**

The  $\alpha_2$  dimer interface. (a) The OMIT map electron density of the  $\alpha_2$  dimer interface contoured at  $1\sigma$  (mesh), showing the four buried His side chains. The OMIT map was calculated using the *SFCHECK* program from *CCP4* (Winn *et al.*, 2011). (b) Hydrogen-bonding interactions at the  $\alpha_2$  dimer interface (PDB entry 3s48) (violet) and the structure re-refined by *PDB\_REDO* (yellow). (c) Comparison of interfacial residues and hydrogen-bonding interactions in PDB entry 3s48 (violet) and the structure re-refined by *PDB\_REDO* (yellow). (d) Hydrogen-bonding interactions at the  $\alpha_1\beta_1$  interface of HbA (PDB entry 2dn3; Park *et al.*, 2006). (e) A comparison of the hydrogen-bonding interactions at the  $\alpha_1\beta_1$ -equivalent interfaces of  $\alpha_2$ ,  $\beta_4$  and  $\alpha_2\beta_2$ . Double-headed arrows indicate hydrogen bonds where the donor/acceptor atoms could potentially be reversed, for example owing to protonation of  $\alpha(2)$ His103 under the influence of  $\alpha(1)$ Asp126. Dashed lines indicate hydrogen bonds that appear in 3s48 or the *PDB\_REDO* re-refinement of 3s48 but not both.

hydrogen-bonding distance of the  $\alpha$  carbonyl of  $\alpha(1)$ Phe117, and the imidazole groups of  $\alpha(1)$ His122,  $\alpha(2)$ His103 and  $\alpha(1)$ His103 were rotated 20–50°, which resulted in the loss of a hydrogen bond between  $\alpha(1)$ His103 and  $\alpha(2)$ His103 that is present in the 3s48 structure (Fig. 2c). Thus, experimental uncertainty in the atomic coordinates is compatible with several hydrogen-bonding patterns, one of these being an extended network involving  $\alpha(2)$ His103,  $\alpha(1)$ His103 and  $\alpha(2)$ His122 (Fig. 2b). Overall, the *PDB\_REDO* re-refinement did not improve the model and therefore the results were not deposited.

A pair of buried His residues, hydrogen bonded through their N<sup>δ1</sup> atoms, is found in the bacterial ammonium transporter (AmtB; Javelle *et al.*, 2006; Liao *et al.*, 2013; Zheng *et al.*, 2004), but we are not aware of structures that have three consecutive hydrogen-bonded His residues, as may occur in the  $\alpha_2$  dimer. However, it has previously been speculated that a hydrogen-bonded His-zipper motif occurs at the dimerization face of the bacterial protein RopE (repetitive organellar protein E) from *Plasmodium chabaudi* (Werner *et al.*, 1998).

### 3.3. Asymmetry of the $\alpha_2$ dimer

In native Hb, the largest contribution to subunit interactions, based on structure and thermodynamic measurements, comes from hydrophobic contacts, with a smaller number of electrostatic interactions (Mrabet *et al.*, 1986; Perutz *et al.*, 1968; Valdes & Ackers, 1977a). The total surface area buried on formation of the  $\alpha_2$  dimer interface is 1473 Å<sup>2</sup>, which is a reduction of 18% compared with the  $\alpha_1\beta_1$  interface in Hb. The reduction in nonpolar buried surface is only slightly greater at 24% (Table 3). Systematic analyses show a relationship between buried surface area and affinity, but also that dissociation constants can vary over four orders of magnitude for the same buried surface area (Chen *et al.*, 2013). Remarkably, the stability of the  $\alpha_1\beta_1$  and  $\alpha_2$  dimers differs by almost eight orders of magnitude indicating that, in this case, buried surface area is an extremely poor indicator of affinity. Analysis of interface packing using the program *SC* from the *CCP4* package (Winn *et al.*, 2011) indicated lower shape complementarity for the  $\alpha_2$  dimerization interface; however, this may reflect the low resolution of data in addition to intrinsic properties of the interface (Lawrence & Colman, 1993).

The arrangement of the two  $\alpha$  chains deviates notably from the  $C_2$  point-group symmetry that is expected for a homodimeric complex: the atoms in one subunit deviate on average by ~2 Å from the position expected in a truly symmetrical dimer. By comparison the  $\alpha_1\beta_1$ -equivalent dimers in  $\beta_4$  and  $\gamma_4$  are highly symmetric. At the native  $\alpha_1\beta_1$  interface, the G helices of the two chains are in close proximity across the dimer axis, raising the possibility that steric clashes could potentially occur in this region in a homodimeric complex. There is evidence of this in the  $\beta_4$  structures, where the side chains of  $\beta$ Cys112(G14) adopt two different rotamers to avoid clashing (Borgstahl *et al.*, 1994a,b). In  $\alpha$ , residues at positions G10, G13 and G14 are larger side chains (His, Leu and Val) than their counterparts in  $\beta$  (Asn, Val and Cys), potentially introducing steric clashes that prevent symmetrical packing of the  $\alpha$  monomers. It has been shown that substitution of  $\alpha$ His103(G10) with the more bulky Tyr destabilizes the native dimer interface (Hoyer *et al.*, 2002), indicating that the side chain at G10 does not reorient to relieve the steric clash. Overall, symmetric homodimers have been postulated to be more stable (Blundell & Srinivasan, 1996).

A comparison of hydrogen-bonding interactions at the  $\alpha_1\beta_1$ -equivalent interfaces of  $\alpha_2$ ,  $\beta_4$  and HbA highlights the asymmetric

**Table 3**

Comparison of the physicochemical properties of  $\alpha_1\beta_1$ -equivalent interfaces in  $\alpha$ , HbA,  $\beta_4$  and  $\gamma_4$ .

	IsdH <sup>N1-<math>\alpha</math></sup> ( $\alpha_2$ interface)	HbA	$\beta_4$	$\gamma_4$
Buried area upon complex formation <sup>†</sup> (Å <sup>2</sup> )	1473	1802	1698	1627
Polar buried area upon complex formation <sup>†</sup> (Å <sup>2</sup> )	995	1172	1182	1109
Nonpolar buried area upon complex formation <sup>†</sup> (Å <sup>2</sup> )	478	630	516	518
No. of residues at the interface <sup>†</sup>	38	49	44	45
Shape complementarity by <i>SC</i> <sup>‡</sup>	0.80	0.65	0.74	0.73

<sup>†</sup> Calculated using *COCOMAPS* (Vangone *et al.*, 2011). <sup>‡</sup> Calculated using *SC* from *CCP4* (Lawrence & Colman, 1993).

nature of the  $\alpha_2$  contacts (Fig. 2e). The asymmetric interactions of His side chains and unpaired negative charge on  $\alpha(2)$ Asp126(H9) have already been discussed. In addition, bilateral interactions of Arg(B12) with the backbone carbonyl of Phe(GH5) occur in  $\beta_4$  and HbA, but only one of these interactions is present in 3s48. Disruption of this interaction in Hb Prato, in which  $\alpha$ Arg31(B12) is replaced by Ser, leads to an unstable Hb (Marinucci *et al.*, 1979). The *PDB\_REDO* re-refinement of 3s48 permits bilateral interactions of  $\alpha$ Arg31(B12), but disrupts symmetrical interactions of  $\alpha$ Ser35(B12).

## 4. Conclusion

The structure reveals the  $\alpha_2$  dimer interface, through which  $\alpha$  self-associates weakly in solution. The  $\alpha_2$  dimer interface reveals a novel hydrogen-bonding network involving up to three of the four buried His side chains. To our knowledge, such a ‘histidine zipper’ has not been described previously. Marked asymmetry in the interface contacts suggests that steric factors reduce the number of stabilizing contacts that can be formed by  $\alpha$  homodimers compared with  $\beta$  homodimers or  $\alpha\beta$  heterodimers, with the consequence of reduced  $\alpha$  stability in  $\beta$ -thalassaemia.

We thank Dr Joel Mackay (the University of Sydney) for supporting KKK to carry out aspects of this work in his laboratory and Dr Tom Caradoc-Davies at the MX beamlines at the Australian Synchrotron for help with data collection. This work was supported in part by the University of Tasmania (to DAG).

## References

- Advani, R., Sorenson, S., Shinar, E., Lande, W., Rachmilewitz, E. & Schrier, S. L. (1992). *Blood*, **79**, 1058–1063.
- Bas, D. C., Rogers, D. M. & Jensen, J. H. (2008). *Proteins*, **73**, 765–783.
- Blundell, T. L. & Srinivasan, N. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 14243–14248.
- Borgstahl, G. E., Rogers, P. H. & Arnone, A. (1994a). *J. Mol. Biol.* **236**, 817–830.
- Borgstahl, G. E., Rogers, P. H. & Arnone, A. (1994b). *J. Mol. Biol.* **236**, 831–843.
- Bucci, E. & Frontice, C. (1965). *J. Biol. Chem.* **240**, PC551–PC552.
- Bucci, E., Fronticelli, C., Chiancone, E., Wyman, J., Antonini, E. & Rossi-Fanelli, A. (1965). *J. Mol. Biol.* **12**, 183–192.
- Chen, J., Sawyer, N. & Regan, L. (2013). *Protein Sci.* **22**, 510–515.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Clegg, J. B. & Weatherall, D. J. (1976). *Br. Med. Bull.* **32**, 262–269.
- Dickson, C. F., Rich, A. M., D’Avigdor, W. M., Collins, D. A., Lowry, J. A., Mollan, T. L., Khandros, E., Olson, J. S., Weiss, M. J., Mackay, J. P., Lay, P. A. & Gell, D. A. (2013). *J. Biol. Chem.* **288**, 19986–20001.
- Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. (2004). *Nucleic Acids Res.* **32**, W665–W667.

- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Gell, D., Kong, Y., Eaton, S. A., Weiss, M. J. & Mackay, J. P. (2002). *J. Biol. Chem.* **277**, 40602–40609.
- Hargrove, M. S., Barrick, D. & Olson, J. S. (1996). *Biochemistry*, **35**, 11293–11299.
- Hargrove, M. S., Whitaker, T., Olson, J. S., Vali, R. J. & Mathews, A. J. (1997). *J. Biol. Chem.* **272**, 17385–17389.
- Hoyer, J. D., McCormick, D. J., Snow, K., Kwon, J. H., Booth, D., Duarte, M., Grayson, G., Kubik, K. S., Holmes, M. W. & Fairbanks, V. F. (2002). *Hemoglobin*, **26**, 175–179.
- Javelle, A., Lupo, D., Zheng, L., Li, X.-D., Winkler, F. K. & Merrick, M. (2006). *J. Biol. Chem.* **281**, 39492–39498.
- Joosten, R. P., Joosten, K., Murshudov, G. N. & Perrakis, A. (2012). *Acta Cryst.* **D68**, 484–496.
- Kidd, R. D., Baker, H. M., Mathews, A. J., Brittain, T. & Baker, E. N. (2001). *Protein Sci.* **10**, 1739–1749.
- Kihm, A. J., Kong, Y., Hong, W., Russell, J. E., Rouda, S., Adachi, K., Simon, M. C., Blobel, G. A. & Weiss, M. J. (2002). *Nature (London)*, **417**, 758–763.
- Krishna Kumar, K., Jacques, D. A., Pishchany, G., Caradoc-Davies, T., Spirig, T., Malmirchegini, G. R., Langley, D. B., Dickson, C. F., Mackay, J. P., Clubb, R. T., Skaar, E. P., Guss, J. M. & Gell, D. A. (2011). *J. Biol. Chem.* **286**, 38439–38447.
- Lawrence, M. C. & Colman, P. M. (1993). *J. Mol. Biol.* **234**, 946–950.
- Liao, S.-M., Du, Q.-S., Meng, J.-Z., Pang, Z.-W. & Huang, R.-B. (2013). *Chem. Cent. J.* **7**, 44.
- Marinucci, M., Mavilio, F., Massa, A., Gabbianelli, M., Fontanarosa, P. P., Camagna, A., Ignesti, C. & Tentori, L. (1979). *Biochim. Biophys. Acta*, **578**, 534–540.
- McCoy, A. J. (2007). *Acta Cryst.* **D63**, 32–41.
- McPhillips, T. M., McPhillips, S. E., Chiu, H.-J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M. & Kuhn, P. (2002). *J. Synchrotron Rad.* **9**, 401–406.
- Miele, G., Manson, J. & Clinton, M. (2001). *Nature Med.* **7**, 361–364.
- Mollan, T. L., Khandros, E., Weiss, M. J. & Olson, J. S. (2012). *J. Biol. Chem.* **287**, 11338–11350.
- Mrabet, N. T., McDonald, M. J., Turci, S., Sarkar, R., Szabo, A. & Bunn, H. F. (1986). *J. Biol. Chem.* **261**, 5222–5228.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Nathan, D. G. & Gunn, R. B. (1966). *Am. J. Med.* **41**, 815–830.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Park, S.-Y., Yokoyama, T., Shibayama, N., Shiro, Y. & Tame, J. R. H. (2006). *J. Mol. Biol.* **360**, 690–701.
- Perutz, M. F., Muirhead, H., Cox, J. M. & Goaman, L. C. G. (1968). *Nature (London)*, **219**, 131–139.
- Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G. & North, A. C. (1960). *Nature (London)*, **185**, 416–422.
- Philo, J. S., Lary, J. W. & Schuster, T. M. (1988). *J. Biol. Chem.* **263**, 682–689.
- Pilpa, R. M., Fadeev, E. A., Villareal, V. A., Wong, M. L., Phillips, M. & Clubb, R. T. (2006). *J. Mol. Biol.* **360**, 435–447.
- Polliack, A., Yataganas, X. & Rachmilewitz, E. A. (1974). *Ann. N. Y. Acad. Sci.* **232**, 261–282.
- Rachmilewitz, E. A. & Schrier, S. (2001). *Pathophysiology of  $\beta$ -Thalassemia*. Cambridge University Press.
- Rieder, R. F., Oski, F. A. & Clegg, J. B. (1969). *J. Clin. Invest.* **48**, 1627–1642.
- Rifkind, J. M., Ramasamy, S., Manoharan, P. T., Nagababu, E. & Mohanty, J. G. (2004). *Antioxid. Redox Signal.* **6**, 657–666.
- Valdes, R. Jr & Ackers, G. K. (1977a). *J. Biol. Chem.* **252**, 74–81.
- Valdes, R. Jr & Ackers, G. K. (1977b). *J. Biol. Chem.* **252**, 88–91.
- Valdes, R. Jr & Ackers, G. K. (1978). *Proc. Natl Acad. Sci. USA*, **75**, 311–314.
- VanDemark, A. P., Kasten, M. M., Ferris, E., Heroux, A., Hill, C. P. & Cairns, B. R. (2007). *Mol. Cell*, **27**, 817–828.
- Vangone, A., Spinelli, R., Scarano, V., Cavallo, L. & Oliva, R. (2011). *Bioinformatics*, **27**, 2915–2916.
- Vasseur, C., Pissard, S., Domingues-Hamdi, E., Marden, M. C., Galactéros, F. & Baudin-Creuzat, V. (2011). *Am. J. Hematol.* **86**, 199–202.
- Vriend, G. (1990). *Protein Eng.* **4**, 221–223.
- Werner, E. B., Taylor, W. R. & Holder, A. A. (1998). *Mol. Biochem. Parasitol.* **94**, 185–196.
- Windsor, W. T., Philo, J. S., Potschka, M. & Schuster, T. M. (1992). *Biophys. Chem.* **43**, 61–71.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Zheng, L., Kostrewa, D., Bernèche, S., Winkler, F. K. & Li, X.-D. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 17090–17095.