Structures of haemoglobin from woolly mammoth in liganded and unliganded states

Hiroki Noguchi,[a] Kevin L. Campbell,[b] Chien Ho,[c] Satoru Unzai,[a] Sam-Yong Park[a]* and Jeremy R. H. Tame[a]*

[a]Protein Design Laboratory, Yokohama City University, Suehiro 1-7-29, Tsurumi-ku, Yokohama 230-0045, Japan, [b]Department of Biological Sciences, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada, and [c]Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Correspondence e-mail: park@tsurumi.yokohama-cu.ac.jp, jtame@tsurumi.yokohama-cu.ac.jp

The haemoglobin (Hb) of the extinct woolly mammoth has been recreated using recombinant genes expressed in *Escherichia coli*. The globin gene sequences were previously determined using DNA recovered from frozen cadavers. Although highly similar to the Hb of existing elephants, the woolly mammoth protein shows rather different responses to chloride ions and temperature. In particular, the heat of oxygenation is found to be much lower in mammoth Hb, which appears to be an adaptation to the harsh high-latitude climates of the Pleistocene Ice Ages and has been linked to heightened sensitivity of the mammoth protein to protons, chloride ions and organic phosphates relative to that of Asian elephants. To elucidate the structural basis for the altered homotropic and heterotropic effects, the crystal structures of mammoth Hb have been determined in the deoxy, carbonmonoxy and aquo-met forms. These models, which are the first structures of Hb from an extinct species, show many features reminiscent of human Hb, but underline how the delicate control of oxygen affinity relies on much more than simple overall quaternary-structure changes.

1. Introduction

The great majority of mammals produce a single major haemoglobin (Hb) component which represents 90% or more of the Hb in the red cells. This protein must therefore adapt to the habitat of the animal in order to both extract sufficient oxygen from the air and deliver it to the tissues efficiently. For instance, vertebrates that live at high altitudes, including high-flying birds, have Hbs with relatively high oxygen affinity (Hiebl *et al.*, 1989; Jessen *et al.*, 1991; Weber, 2007), and evolution has found several elegant solutions to the problem of adjusting this property of the protein (Perutz, 1983). Most vertebrate Hbs are heterotetramers composed of two α-type and two β-type subunits, each of which carries a single haem group for the reversible binding of oxygen molecules. Oxygen transport is effected by the Hb molecule binding oxygen cooperatively, so that it saturates in the lungs and releases oxygen readily within the body. Purified human Hb is found to have a very high oxygen affinity, but this is reduced inside red cells by a roughly equimolar concentration of 2,3-diphospho-glycerate (DPG, also known as BPG). This polyanion binds to the protein near the N-terminus of the β subunits and lowers oxygen affinity. Near actively respiring tissues, oxygen release is also assisted by the presence of carbon dioxide (Bohr *et al.*, 1904), which can bind directly to Hb and, *via* its conversion to carbonic acid, also lowers the pH of the blood, leading to the well known Bohr effect. Finally, Hb affinity can be further

modulated by chloride ions, although the precise mechanisms are under debate and remain to be elucidated. The molecular structure of Hb from humans and other animals is known through the pioneering work of Perutz, whose group discovered that the protein adopts two basic quaternary conformations (Perutz, 1970, 1972). These were identified with the T and R states of the allosteric model proposed by Monod, Wyman and Changeux (Monod *et al.*, 1965), so that cooperatively was suggested to arise as the low-affinity T state switched to the high-affinity R state with increasing oxygenation. While this essential picture has been thoroughly tested, it clearly neglects other very important effects. The oxygen affinity of the haems is not determined solely by the quaternary state of the protein, but, as noted above, by heterotropic ligands which bind at various positions on the protein surface (Imai, 1982). It has been demonstrated, for example, that Hb in the R state may have an affinity just as low as the T-state protein: the artificial effector L35 binds to both the T state and R state and reduces the oxygen affinity of both to the same level (Shibayama *et al.*, 2002; Yokoyama *et al.*, 2006). The textbook description of Hb lays strong emphasis on the allosteric switch of the protein, but it is now clear that dynamic effects are also at work, control-ling the oxygen affinity of the protein within each quaternary state (Schay *et al.*, 2006; Ho *et al.*, 2011; Yonetani *et al.*, 2002).

The regulation of oxygen affinity by heterotropic effector molecules can also vary among mammalian Hbs. Hb from ungulates such as cows, for example, is found to respond weakly to organophosphates such as DPG; these proteins have an intrinsically low oxygen affinity which makes such effectors unnecessary (Perutz & Imai, 1980). The low intrinsic oxygen affinity of bovine Hb is believed to arise largely from changes in the β-type subunit of the protein, although it has been demonstrated that mutating human HbA to give it a similar N-terminus to the β subunit does not transfer this effect (Fronticelli *et al.*, 1995). Although the pH dependence (Bohr effect) of bovine Hb is similar to that of human Hb, the enthalpy of oxygenation (ΔH) is only about one-third as much at pH 9, where enthalpy effects of protonation can be ignored (roughly -29 kJ mol^{-1} instead of -84 kJ mol^{-1}; Razynska *et al.*, 1990). A reduced enthalpy is highly desirable for animals living in cold environments (Giardina *et al.*, 1990; Weber & Campbell, 2011) since enthalpy is also a measure of the change of free energy of a reaction with temperature, a relationship quantified by the van't Hoff equation. The Hbs of several Arctic species, notably reindeer, have been studied in some detail (Giardina *et al.*, 1990). Unlike human HbA, which has a similar ΔH in the T and R states, for reindeer Hb the enthalpy of oxygenation becomes much smaller with each oxygen ligand. Thus, for human Hb the oxygenation curve is shifted with temperature but maintains the same shape, whereas for reindeer Hb it changes shape with temperature. It is clear that there are numerous possible routes to reducing the ΔH of oxygenation while leaving the Gibbs free energy (ΔG) unchanged, exploiting the well known phenomenon of entropy–enthalpy compensation. The observed heat change on binding of any protein to a ligand in fact reflects a number of different factors including solvent rearrangement (Chervenak & Toone, 1995) and homotropic effects not associated with effector binding.

The ancestors of woolly mammoths (*Mammuthus primigenius*) evolved in subtropical Africa and were phylogenetically closer to modern Asian (also called Indian) elephants than African elephants are. Woolly mammoths only colonized high-latitude environments in the early Pleistocene Ice Ages some 1.2–2.0 million years ago. Presumably under strong selective pressures associated with this new environment, woolly mammoth Hb acquired three
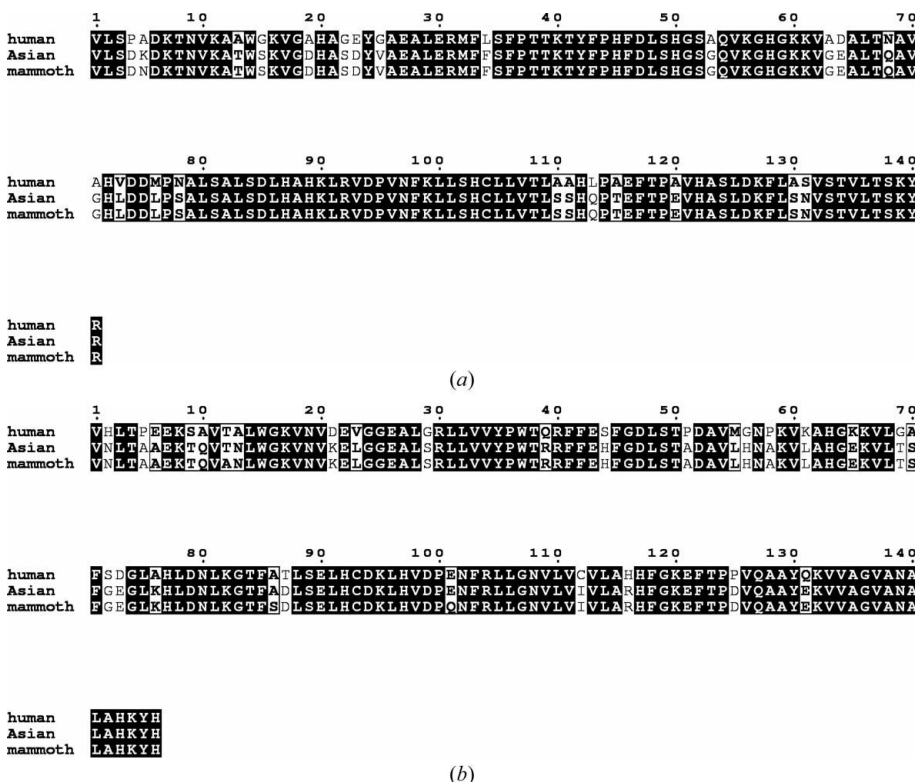
Figure 1
(*a*) An alignment of the α-globin sequences of human HbA, Asian (Indian) elephant (*Elephas maximus*) and woolly mammoth (*Mammothus primigenius*). The sequences from *E. maximus* and mammoth differ at one residue (Asn5→Lys), and similarly the modern African elephant (*Loxodonta africana*) shows one amino-acid difference, Ser49→Gly. Residues that are identical in all three sequences are shown in white on black. The figure was produced with *ESPript* (Gouet *et al.*, 1999). (*b*) An equivalent alignment of the β or β/δ subunits. The mammoth β/δ subunit has acquired three mutations relative to the Asian (Indian) elephant, Thr12→Ala, Ala86→Ser and Glu101→Gln. The single point mutation between modern African and Asian elephants in this subunit, Asp52→Glu, corroborates the evidence from the α subunits that the mammoth and the Asian elephant descended from a common ancestor after the African elephant had diverged.

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the outer shell.

	Met	CO	Deoxy
Data collection			
Resolution range (Å)	50.0–1.50 (1.53–1.50)	50.0–1.55 (1.58–1.55)	46.51–2.20 (2.32–2.20)
Space group	$C2$	$C2$	$P1$
Unit-cell parameters			
a (Å)	109.4	109.5	53.6
b (Å)	61.5	61.8	59.1
c (Å)	53.5	53.6	62.1
α (°)			96.7
β (°)	110.1	110.4	114.4
γ (°)			113.9
No. of measured reflections	224380	237928	75756
No. of unique reflections	53178	47308	28619
Completeness (%)	96.7 (96.2)	97.4 (83.2)	94.4 (93.9)
Mean $I/\sigma(I)$	50.09 (14.14)	51.4 (7.99)	6.6 (2.7)
Multiplicity	4.3	2.6	2.6
R_{merge}† (%)	4.7 (12.6)	5.0 (18.7)	9.8 (32.7)
Refinement statistics			
Resolution range (Å)	20.0–1.50	20.0–1.55	20.0–2.20
R factor‡ (%)	18.8	17.2	20.6
R_{free}‡ (%)	22.0	19.9	26.9
Solvent molecules	242	244	209
R.m.s. deviations from ideal			
Bond lengths (Å)	0.023	0.024	0.012
Bond angles (°)	2.059	2.104	1.359
Chiral volumes (Å³)	0.129	0.154	0.093
Ramachandran plot, residues in (%)			
Most favourable region	98.2	98.2	97.5
Additional allowed region	1.8	1.8	2.5

† $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation, $\langle I(hkl)\rangle$ is the mean value for that reflection and the summations are over all reflections. ‡ R factor $= \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. The free R factor was calculated with 5% of the data excluded from the refinement.

amino-acid substitutions, all of which are located on the chimeric β/δ globin chain. Notable among these mutations is the Gluβ/δ101(G3)→Gln replacement situated within the central cavity of the Hb tetramer near subunit interfaces of the protein that markedly alter its physiochemical properties. In the absence of chloride or the polyanionic effector DPG, mammoth Hb has substantially higher oxygen affinity than elephant Hb, but a very similar heat of oxygenation (Campbell *et al.*, 2010; Yuan *et al.*, 2011). Chloride ions and DPG exert a stronger effect on it, however, so that with these effectors present the oxygen affinity of mammoth Hb is close to that of elephant Hb under the same conditions at 310 K but with a significantly smaller heat release on oxygen binding. A sequence alignment of human and modern Asian elephant globins is shown in Fig. 1.

Five human variant Hbs are known in which the highly conserved β101 position is altered, and all of them show strong increases in intrinsic oxygen affinity. Hb Rush, which carries the same Gluβ101→Gln mutation as mammoth Hb, is slightly unstable and leads to mild haemolytic anaemia (Adams *et al.*, 1974). Alone among the five variants, Hb Rush also possesses a proton-linked chloride-binding site (Shih *et al.*, 1985). Despite considerable study, the stereochemical basis underlying this difference remains unclear. Here, we have attempted to match the physiochemical properties of woolly mammoth

Hb to the structures of the protein in the T and R states. We find no evidence of surface chloride-binding sites and that the unique Gluβ101(G3)→Gln mutation imposes no significant structural change on the protein.

2. Materials and methods

2.1. Protein crystallization

Mammoth Hb was produced as described by Campbell *et al.* (2010) and stored in the carbonmonoxy form as pellets frozen in liquid nitrogen. To prepare crystals, the protein was dialyzed into 10 mM Tris–HCl pH 7.5 and concentrated to 60 mg ml^{-1}. CO was removed as necessary using a table lamp to illuminate a sample rotated on an ice-water bath under a stream of air. Deoxy and CO crystals were grown in rubber-stoppered tubes holding 50 µl samples as described previously (Park *et al.*, 2006). Deoxy crystals were grown under nitrogen using 2.8 M ammonium sulfate pH 6.5 as a precipitant. Crystals of the aquo-met (oxidized) form of the protein were grown while exposed to air using 2.4 M sodium/potassium phosphate pH 6.7 and the CO form was subsequently also crystallized under the same conditions using 0.1 M Tris–HCl pH 6.5, 1.9 M ammonium sulfate, and the CO form was subsequently crystallized using 2.2 M sodium/potassium phosphate pH 6.7. Crystals were harvested using standard nylon cryoloops (Hampton Research) in mother liquor containing 20% glycerol (25% in the case of the met form) before cooling to 100 K.

2.2. Data collection and structure refinement

Data were collected on beamline BL17A of the Photon Factory, Tsukuba, Japan using an ADSC CCD detector and were processed with *HKL*-2000 (Otwinowski & Minor, 1997). Structure determination was carried out using the *CCP*4 suite of programs (Winn *et al.*, 2011). *MOLREP* (Vagin & Teplyakov, 2010) was used to build initial models by molecular replacement using PDB entries 2dn1 and 2dn2 (Park *et al.*, 2006) as models of human oxy and deoxy Hb. The models were manipulated with *Coot* (Emsley *et al.*, 2010), which was also used for model validation. Refinement was carried out with *REFMAC* (Murshudov *et al.*, 2011). Default restraints (geometric and thermal parameters) were used. In all cases 5% of reflections were used to calculate a free R factor. No σ-factor cutoff or resolution cutoff was applied at any stage. Solvent water molecules were modelled into the map where geometrically reasonable with at least 1σ $2mF_o - DF_c$ electron density to support their inclusion. Structures and structure factors have been deposited in the PDB with codes 3vre (deoxy form), 3vrf (carbonmonoxy form) and 3vrg (met form). Data-collection and refinement statistics are given in Table 1.

3. Results

3.1. Overall structure

The DNA sequences of component globin chains from woolly mammoth have been determined previously (Campbell

Table 2
Indicator hydrogen-bond lengths (Å) in different liganded states of mammoth Hb.

Values in parentheses indicate symmetry-equivalent interactions. A cutoff distance of 3.5 Å was used. Distances are rounded to one decimal place.

	CO	Met	Deoxy
Lys$\alpha_1$40$\cdots$His $\beta/\delta_2$146	—	—	2.6 (2.8)
Tyr$\alpha_1$42$\cdots$Asp$\beta/\delta_2$99	—	—	2.6 (2.6)
Asp$\alpha_1$94$\cdots$Trp$\beta/\delta_2$37	3.7	3.7	2.9 (3.2)
Asp$\alpha_1$94$\cdots$Asn$\beta/\delta_2$102	2.7	2.8	—
Arg$\alpha_1$141$\cdots$Asp$\beta/\delta_2$126	—	—	2.7 (2.4)
Asp$\beta/\delta_1$94$\cdots$His$\beta/\delta_1$146	—	—	2.9 (2.7)
Trp$\beta/\delta_1$37$\cdots$Asn$\beta/\delta_1$102	3.0	2.9	—

et al., 2010). The sequences are entries D3U1H8 and D3U1H9 in the UniProt database. Rather than a typical β subunit, the β-type chain of mammoth Hb is the product of a chimeric fusion gene (*HBB/D*) that arose *via* an unequal crossover event between the parental *HBB* (β) and *HBD* (δ) loci that predates the diversification of paenungulate (elephants, sea cows and hyraxes) mammals (Opazo *et al.*, 2009). As in living elephants, mammoth Hb contains 141 residues in the α subunit and 146 residues in the β/δ subunit. Notably, the mammoth β/δ subunit shows three amino-acid changes compared with the Asian elephant protein, Thr12(A9) to Ala, Ala86(F2) to Ser and Glu101(G3) to Gln, while the α subunit of the Asian species evolved a single replacement, Asn5(A3) to Lys, following its divergence from the mammoth lineage (Campbell *et al.*, 2010).

Despite over 60 million years of independent evolution, the mammoth Hb protein is 81% identical to human HbA, and the crystallographic models are consequently strikingly similar to the human protein in overall structure. One of the most notable features of sequence comparison between the human and mammoth proteins is that HbA has 14 proline residues per $\alpha\beta$ dimer, whereas the mammoth protein (and modern elephant Hb) has only nine, implying a slightly more flexible structure. In the deoxy model, the largest deviation between the C$^\alpha$ traces of mammoth and human Hb is found in the D helix of one β/δ subunit, where Proβ51(D2) is replaced by alanine in mammoth Hb. Additionally, the mammoth protein has Aspβ/δ52(D3) close to Glu$\alpha_2$120(H3), although the aspartate makes a hydrogen bond to Hisβ/δ56(DE2). This Asp–His salt bridge is maintained in the R state and thus is unlikely to contribute to the Bohr effect, which is markedly reduced in Asian elephants relative to HbA (Yuan *et al.*, 2011). Organic anions such as inositol hexaphosphate bind to a single site near the N-termini of the β subunits of T-state Hb (Arnone, 1972). It has previously been suggested that changes at the N-termini of the β chains are responsible for the functional differences between human and bovine Hb, but there are no obvious structural differences between the human and mammoth proteins here. The mutation of Proβ5(A2) in HbA to alanine does not shift the C$^\alpha$ trace of the protein, but Hisβ2(NA2) is replaced by Asnβ/δ2, which presumably contributes to the weaker binding of organic phosphates to elephantid Hbs (Yuan *et al.*, 2011). A further example of the

Table 3
Comparison of mammoth Hb models with human and bovine Hb.

Root-mean-square deviations (Å) of the tetramer using the 'Tame frame' (148 C$^\alpha$ atoms from residues α23–48, α57–63, α101–111, α118–125, β51–57, β110–116 and β119–132). The PDB models used were 2dn2 (human deoxy; Park *et al.*, 2006), 2dn1 (human oxy; Park *et al.*, 2006), 1hda (bovine deoxy; Perutz *et al.*, 1993), 1fsx (bovine carbonmonoxy; Safo & Abraham, 2001) and 1bbb (human R2 state; Silva *et al.*, 1992). Figures are quoted to two decimal places. Values in parentheses show the overlaps between alternative $\alpha\beta$ dimers where these are found in the asymmetric unit.

	Tetramer	Dimer
Deoxy mammoth–deoxy human	0.58	0.55 (0.46)
Deoxy mammoth–deoxy bovine	0.58	0.55 (0.41)
Deoxy human–deoxy bovine	0.30	0.22 (0.25)
CO mammoth–CO human	0.95	0.61
CO mammoth–CO bovine	0.88	0.58
CO human–CO bovine	0.82	0.33
CO mammoth–human R2	1.26	0.56
Oxy human–human R2	1.14	0.29
CO bovine–human R2	0.75	0.26

replacement of a proline in HbA by alanine is found at β58(E2), but this also has minimal effect on the protein structure. Compensating mutations are also found, for example the closely apposed Aspβ21(B3) and Lysβ65(E9) of HbA are replaced by Lys and Glu, respectively, in mammoth Hb. The residues forming the haem pockets of both subunit types are conserved between mammoth and human Hb, and the structures of these regions also overlap extremely well.

The allosteric mechanism of HbA involves a concerted shift in the contacts between one $\alpha\beta$ dimer and the other (Perutz *et al.*, 1998). Interactions between the α C helix and the β FG corner change substantially, while those at the β C helix and α FG corner show much more modest changes. These two regions are therefore known as the switch and flexible slide regions, respectively (Baldwin & Chothia, 1979), and interactions at these sites are largely preserved between fish and mammalian Hbs (Tame *et al.*, 1996; Yokoyama *et al.*, 2004). Since fish and mammalian lineages diverged hundreds of millions of years ago, it is unsurprising to find that mammoth Hb and HbA share the same key residues forming hydrogen bonds characteristic of the T and R states (Park *et al.*, 2006). These are listed in Table 2. Tyr$\alpha_1$42(C7) forms a hydrogen bond to Asp$\beta/\delta_2$99(G1) in the T state which is used as a marker in NMR studies to determine the quaternary state (Fung & Ho, 1975). The region around this residue overlaps extremely well with human HbA. Asp$\alpha_1$94(G1) is another highly conserved residue which hydrogen bonds to Trp$\beta_2$37(C3) in the T state, as illustrated in the electron-density map around these residues in deoxy and carbonmonoxy mammoth Hb (Fig. 2). The preservation of these core residues and their interactions shows that the allosteric mechanism is essentially the same as that of human Hb. Using the 'Tame frame' of 148 C$^\alpha$ atoms (Park *et al.*, 2006), which is similar to the BGH frame (Baldwin & Chothia, 1979), to compare different Hb molecules, we find r.m.s.d.s of 0.58 Å between mammoth and human Hb or mammoth and bovine Hb and of 0.30 Å between human and bovine Hb (Table 3), differences that are close to the expected experimental error.
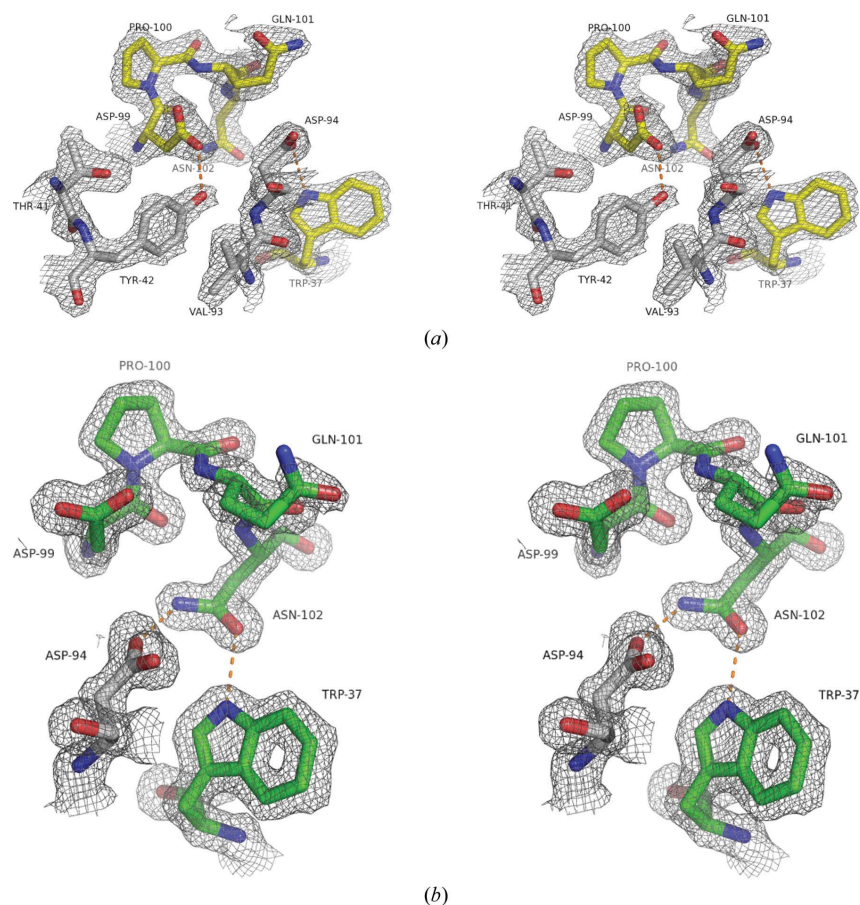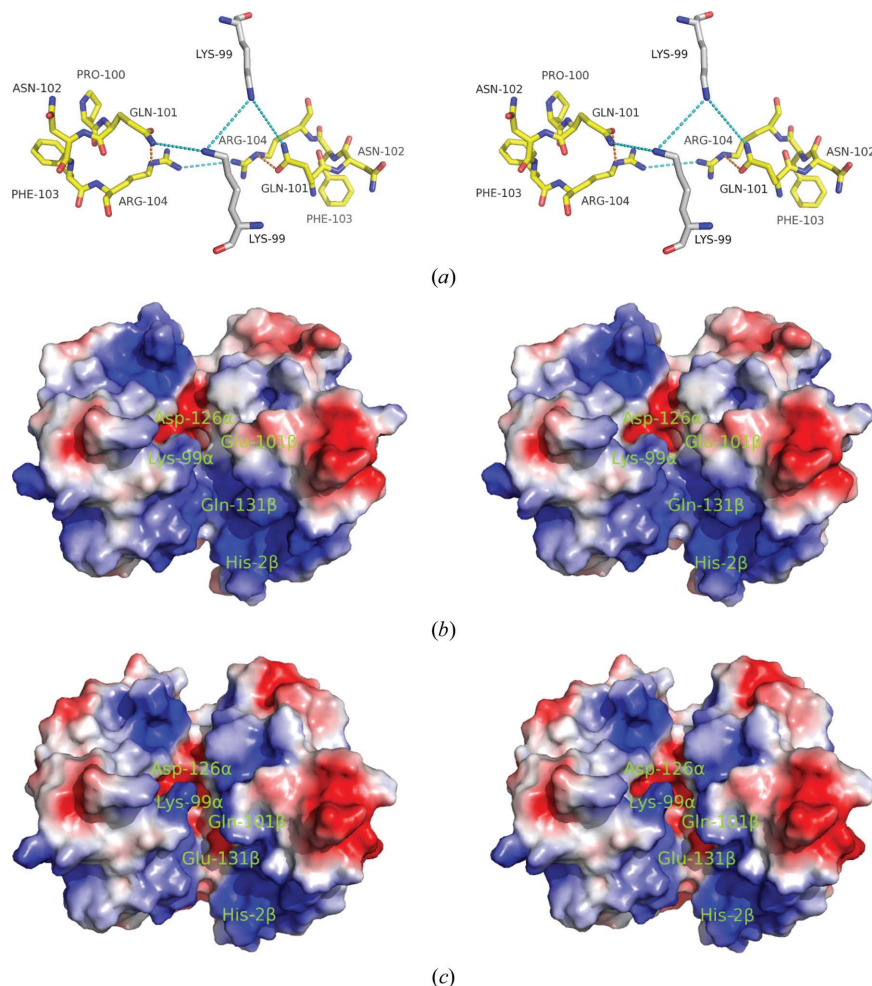
Figure 2

(a) A stereoview of the $\alpha_1\beta/\delta_2$ interface of deoxy woolly mammoth Hb showing the $2mF_o - DF_c$ electron-density map at a contour level of 1σ. The structure is shown as a stick model coloured by atom type: oxygen, red; nitrogen, blue. C atoms of the α subunit are coloured white and C atoms of the β/δ subunit are coloured yellow. Hydrogen bonds used as markers of the T state are shown as orange dotted lines. Glnβ/δ101, the principal mutation in mammoth Hb, can be seen at the top of the figure. Figures were produced using *PyMOL* (DeLano, 2002). (b) A stereoview of the $\alpha_1\beta/\delta_2$ interface of carbonmonoxy woolly mammoth Hb showing the $2mF_o - DF_c$ electron-density map at a contour level of 1σ. C atoms of the α subunit are coloured white and C atoms of the β/δ subunit are coloured green. Hydrogen bonds used as markers of the R state are shown as orange dotted lines.
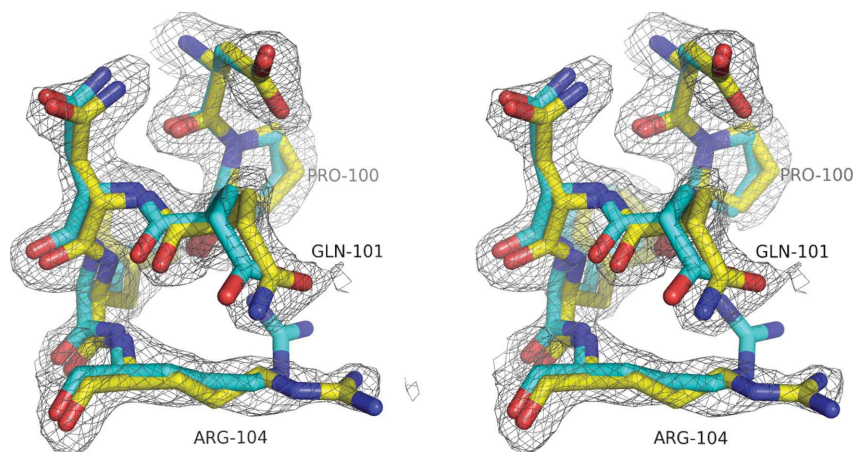
Figure 3

A stereoview showing the overlap of deoxy woolly mammoth Hb and deoxy HbA around Glnβ/δ101. C atoms of the human protein are coloured cyan and the $2mF_o - DF_c$ electron-density map is shown at a contour level of 1σ. Argβ/δ104 adopts a more extended conformation in mammoth Hb, pointing towards its symmetry mate on the opposite side of the central cavity.

3.2. Quaternary structure

There has been considerable debate concerning the nature of the liganded state of Hb (Tame, 1999). Early data from the Perutz group suggested that haem ligands or oxidation to the aquo-met (Fe^{3+}) state drive the protein into the R conformation. Arnone and coworkers later found that liganded HbA can crystallize in a different form called 'R2' (Silva *et al.*, 1992). It was subsequently suggested that the R structure is in fact an artifact of the crystallization conditions and that the R2 structure more accurately reflects the solution structure of the molecule (Srinivasan & Rose, 1994). Unfortunately, minor errors such as a peptide flip and incorrect side-chain rotamers in the highly cited 1983 crystal structure of oxy-HbA (Shaanan, 1983) have been taken as hallmarks of the R state by some authors (Srinivasan & Rose, 1994; Safo & Abraham, 2005), and the lack of a consistent definition of the quaternary conformations has greatly confused the literature. We use the pattern of hydrogen bonding to class a model as T or R state, rather than rotation angles or side-chain rotamers; the residues involved in forming these interactions are all well conserved over millions of years of evolution and are readily modelled in the electron-density maps. Such a definition allows significant conformational variation within both the T and the R states. A number of slightly different crystal structures of liganded Hb have been published, suggesting that the protein samples a relatively large conformational space. It remains debatable whether some of these forms can usefully be called novel quaternary states, since they are little different from the well known R structure. However, NMR analysis of human oxy-Hb in solution shows that it samples a considerable conformational space and is not well represented by either the R or the R2 form alone (Gong *et al.*, 2006; Lukin *et al.*, 2003). More recently, molecular-dynamics simulations have shown that T-state HbA rapidly relaxes to the R state, but does not sample the R2 state on the timescale of the experiment (Hub *et al.*, 2010). The liganded mammoth Hb structures refined here are closer to the R conformation of HbA than the R2 conformation. Like other animal Hbs (Tame, 1999), mammoth Hb supports the idea that liganded Hb is conformationally flexible but generally maintains the hydrogen-bonding

(a)

(b)

(c)

Figure 4

(a) A stereoview of the central cavity of deoxy woolly mammoth Hb. C atoms of the α subunit are coloured white and C atoms of the β/δ subunit are coloured yellow. Orange dotted lines indicate distances of between 3 and 4 Å in length and light blue dotted lines indicate distances of between 4 and 7.2 Å in length. In human HbA, the charge on the pairs of lysine and arginine residues is partly countered by Gluβ/δ101. (b) A stereoview of the molecular surface of one $\alpha\beta$ dimer of deoxy human HbA coloured by electrostatic potential (red, negative; blue, positive). Colour saturation corresponds to an electron energy of $\pm 3kT$. The α chain is on the left; the central cavity is shown running vertically and the DPG binding site is at the bottom. Aspα126 is responsible for the negatively charged pocket between the two subunits and Lysα99 projects from the α subunit towards Gluβ101. The surface potential was calculated using *APBS* (Baker *et al.*, 2001). (c) A similar figure to (b) but showing deoxy woolly mammoth Hb. The electrostatic potential is altered from the normal pattern by two mutations: Gluβ/δ101→Gln and Glnβ/δ131→Glu. The potential is essentially unchanged around Asnα131 and Valβ/δ1, which are the sites of chloride binding in other animal Hbs. Lysα99 shows a higher potential in woolly mammoth Hb than in HbA.

pattern shown for the carbonmonoxy and aquo-met models (Table 2).

3.3. Differences between elephant and mammoth Hb

Although the T and R structures of mammoth Hb are highly similar to those of human Hb, several important distinctions are apparent. Histidine residues are important to Hb since they mediate much of the Bohr effect (Busch *et al.*, 1991; Fang *et al.*, 1999; Sun *et al.*, 1997). As in HbA, the C-terminal histidine of the β/δ subunit, Hisβ/δ146(HC3), makes a salt

bridge with Aspδ94(FG1) in the T state but not in the R state, so that oxygenation releases protons. Mammoth Hb has 19 histidine residues in the $\alpha\beta$ dimer, two of which, β/δ44(CD3) and β/δ56(DE2), are not present in HbA. Conversely, HbA has Hisβ2(NA2), which is Asn in the mammoth protein, and Hisβ116(G18), which is Arg in the mammoth protein. The histidine residues all have very similar interactions in common except for Hisα50(CE8). This imidazole side chain makes a stable hydrogen bond to Gluα30(B11) of the same subunit in both proteins, but the histidine rotamer is different in one α chain in the asymmetric unit of the deoxy form since Proβ/δ125(H3) in HbA is changed to Asp in mammoth Hb, pulling Hisα30(B11) towards it. However, the interaction between Asp$\beta/\delta_1$125 and His$\alpha_1$50 is unlikely to contribute to the Bohr effect since the histidine maintains a hydrogen bond to Gluα30 in both the T and R states and the second pK_a of the imidazole well above the physiological pH.

The principal feature of interest is how with very modest changes to the protein surface the thermodynamics of oxygen binding may be adapted to an Arctic environment. In the case of mammoth Hb, the enthalpy change of oxygen binding seems to be partly reduced compared with ancestral Hbs through changes in the chloride effect (Campbell *et al.*, 2010). Many vertebrate Hbs, including human Hb, show reduced oxygen affinity in the presence of chloride ions, although high concentrations of chloride, typically 100 mM, are required to observe this effect fully. Since chloride binding is very weak, minor changes to the protein may influence it appreciably. The mutation in mammoth Hb at the highly conserved residue Gluβ/δ101(G3) is therefore highly surprising. In HbA, this glutamate residue lies close to Argβ104(G6) (Fig. 3), towards which it is pushed by the perfectly conserved Asp$\alpha_2$94(G1) and Asp$\beta_1$99(G1). The arginine side chain shows different conformations in different T-state structures so that it may form, for example, a hydrogen bond to the carboxyl O atom of Proβ100(G2), but it generally prefers to bind to the side chain of Asnβ139(H17) (Park *et al.*, 2006; Tame & Vallone, 2000). In the deoxy mammoth Hb, Glnβ/δ101(G3) exerts no electrostatic pull on the arginine, allowing it to stretch across the central cavity of the tetramer towards its symmetry mate. The two copies of Argβ/δ104(G6) approach within 5 Å of each other in the T state, whereas in the R-state structure the closest approach of these symmetry-

related arginine residues is almost 6.5 Å. This mutation is therefore expected to destabilize the T state relative to the R state, which is consistent with the elevated intrinsic oxygen affinity of the protein.

The Gluβ101 residue of human Hb adopts a very similar position to that of Glnβ/δ101 in mammoth Hb, so the effects of this mutation appear to arise from electrostatic changes and their influence on nearby residues. It can be seen from Fig. 4 that Glnβ/δ101 lies between Lysα99(G6) and Argβ/δ104(G6), both of which are highly conserved among vertebrates. The mammoth Gluβ/δ101→Gln mutation therefore leads to an increased and concentrated positive charge within the central cavity, which may well draw chloride ions into the protein and increase its sensitivity to this allosteric effector. Mammoth Hb has a significantly higher oxygen affinity than elephant Hb in the absence of heterotropic effectors, but a comparable affinity in the presence of chloride and DPG (Yuan *et al.*, 2011; Campbell *et al.*, 2010), so it appears that repulsion between Lysα99(G6) and Argβ/δ104(G6) (and their symmetry mates) is alleviated by these counterions. Exactly the same behaviour has been reported for Hb Rush by Shih *et al.* (1985), who noted that under stripped conditions the affinity for the first oxygen ligand K_1 is increased but K_4 is unchanged; however, the increased positive charge within the central cavity increases proton-linked chloride binding and restores oxygen affinity to normal in the presence of 100 mM chloride. This extra chloride binding reduces the heat of oxygenation exactly as found for mammoth Hb. No site for the extra proton binding has been identified previously, but the structure of mammoth Hb shows that Lysα99(G6) adopts the same position as in human HbA and this residue appears to be the most plausible candidate.

The residue Alaβ/δ12(A9) lies on the surface of mammoth Hb within helix A, close to two lysine residues: Lysβ/δ8(A5) and Lysβ/δ76(E20) (Fig. 5a). The C$^\alpha$ trace of mammoth Hb is very close to that of human HbA around the N-terminus of the non-α subunit, but is slightly different from that of bovine Hb. Human β globin (like elephant β/δ globin) has a threonine at position 12, but an alanine at position 76. The lysine residues are shared by bovine Hb and have been strongly implicated in chloride ion binding to this protein (Fronticelli, 1990). In bovine Hb, however, main-chain shifts around the N-terminus push Lysβ8(A5) towards Thrβ12(A9) so that the lysine carbonyl O atom makes hydrogen bonds to both the threonine side chain and N atom. Other Hbs only have the main-chain interaction within the A helix. The loss of the threonine at position β/δ12 may have the effect of removing this chloride ion-binding site, although such binding may in any case require the small shift of the A helix found in bovine Hb. We have previously suggested (Campbell *et al.*, 2010) that the mammoth Thrβ/δ12→Ala mutation may allow the Lysβ/δ8 side chain to interact more strongly with the carboxyl group of Aspβ/δ79 on the same subunit and weaken the repulsion of anions around the DPG binding site. Some evidence for such an effect can be seen in one $\alpha\beta$ dimer in the deoxy mammoth Hb model (Fig. 5b) but not the other. Even where Lysβ/δ8 does approach Aspβ/δ79(EF3), the hydrogen bond is almost 3.5 Å in length. Although the crystal structures do not suggest any gain or loss of function arising from the Thrβ/δ12→Ala mutation, some effect may occur in the presence of DPG. Finally, the replacement of Alaβ/δ86(F2) by Ser in mammoth Hb is unusual but this surface residue apparently forms no important interactions and the structure gives no suggestion that the functional properties of the protein would be altered in any way. To our knowledge, this region of the protein has not been implicated in chloride or DPG binding in any vertebrate Hb.
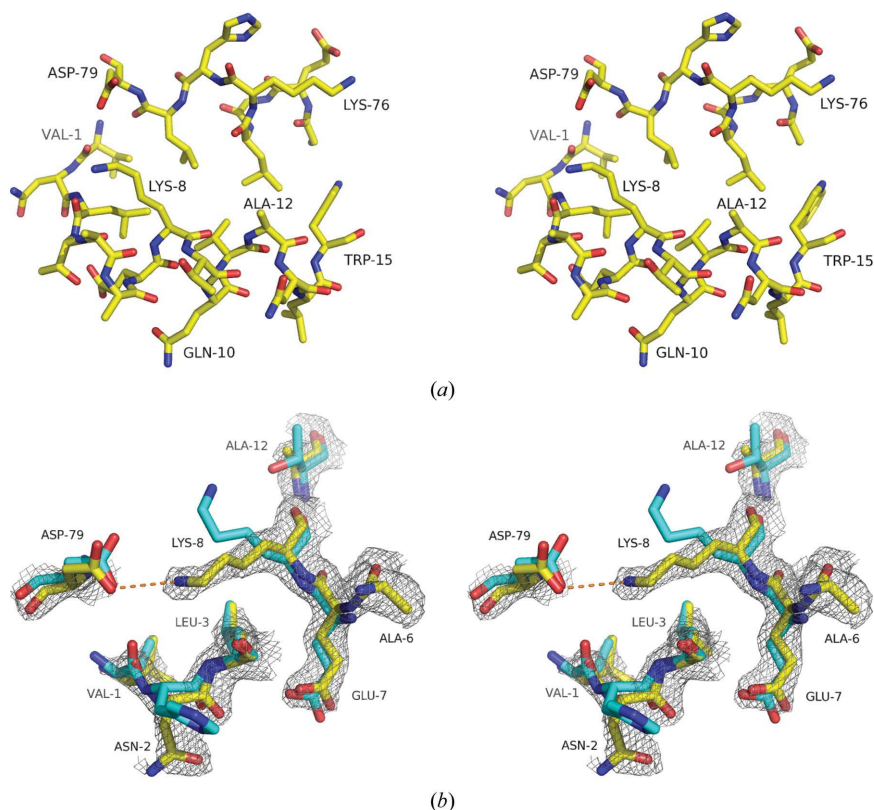
Figure 5
(*a*) The surface of the β/δ subunit of deoxy woolly mammoth Hb, showing the proposed chloride ion-binding site of bovine Hb (Fronticelli *et al.*, 1995). In the subunit shown, Lys8 makes no interaction with Asp79 of the same subunit. (*b*) Stereoview of the deoxy woolly mammoth Hb β/δ subunit not shown in (*a*), with the $2mF_o - DF_c$ map superposed at a level of 1σ. In this case Lysβ/δ8 forms a salt bridge with Aspβ/δ79. Superimposed on this model is deoxy HbA (C atoms coloured cyan). In HbA Lysβ8 can hydrogen bond to Thrβ12.

4. Discussion

The chloride effect of HbA was originally explained, like the DPG effect, by preferential binding to the low-affinity T state changing the allosteric equilibrium of the protein. It was realised early on that the

oxygen affinity of R-state Hb is dependent on the solution conditions, but this effect was largely neglected until the global allostery model was put forward by Yonetani and coworkers (Tsuneshige *et al.*, 2002; Yonetani & Tsuneshige, 2003). Unlike DPG, which can be visualized in crystallographic electron-density maps (Arnone, 1972), chloride proved more difficult to detect in the case of bovine Hb. This led Perutz and coworkers to propose a 'new' type of allosteric effect in which chloride ions simply acted to weaken the repulsion of positively charged groups within the central cavity of the protein (Perutz *et al.*, 1993); since these charges move closer together in the T state than the R state, chloride ions were proposed to stabilize the T state. It was found that even bromide ions could not be detected in the crystal structure of bovine Hb, which was interpreted as demonstrating that there are no fixed binding sites for chloride or bromide within the T-state structure of bovine or human Hb. This view has remained controversial, not least because of the negative nature of the evidence. Considerable experimental work has been carried out on the chloride effect of bovine Hb (Fronticelli, 1990). Like human Hb, bovine Hb loses about 2.2 chloride ions on oxygenation near neutral pH in the presence of 100 mM chloride (Perutz *et al.*, 1994; Fronticelli *et al.*, 1995). Mutagenic experiments have implicated the lysine residues at positions $\beta 8$ and $\beta 76$ on the outer surface of bovine Hb in chloride binding (Fronticelli *et al.*, 1995), implying a very different mechanism to that of Perutz. As mentioned above, HbA has alanine at position $\beta 76$, suggesting that the chloride effect may occur through quite different residues in different Hbs. Some data indicate that chloride ions can interact with other Hbs at specific sites. For example, the frog *Telmatobius* lives high in the Andes around Lake Titicaca and its unusually high oxygen-affinity Hb has been found to be unresponsive to chloride (Weber *et al.*, 2002). Sequence comparisons with normal frog Hbs strongly suggest that chloride binding is lost owing to a single mutation in the α chain, Ser$\alpha 131$(H14)$\rightarrow$Ala, within the central cavity. In HbA, this serine lies close to the N-terminus of the β subunit of the opposite $\alpha\beta$ dimer, a site implicated in chloride binding to the T state (O'Donnell *et al.*, 1979). This replacement of one uncharged residue by another smaller one can hardly block chloride binding through steric or electrostatic effects. Ser$\alpha 131$(H14) is not preserved in bovine or Asian elephant Hb, being replaced in both cases by asparagine. The chloride effect of elephant Hb is notably low (Yuan *et al.*, 2011), yet the positively charged residues within the central cavity are preserved between bovine and elephant Hbs. This result may possibly be reconciled with the Perutz model by the extra negative charge in mammoth and elephant Hb at residue $\beta/\delta 139$ (see below). Comparing elephant and mammoth Hb, it can be seen that the mammoth protein is more sensitive to chloride, while retaining Asn$\alpha 131$(H14). This residue lies over 20 Å from Glu$\beta/\delta_1 101$ and Glu$\beta/\delta_2 101$, so chloride binding at the site lost from *Telmatobius* Hb can be ruled out. Furthermore, from the structures of mammoth Hb presented here, it is highly unlikely that the protein surface interacts with chloride ions in the manner suggested for bovine Hb. Additional evidence against these surface sites has recently

emerged from a comparative study of mole Hbs (Signore *et al.*, 2012).

We are therefore led to conclude that the functional differences between mammoth Hb and elephant Hb arise primarily from the Gln$\beta/\delta 101$ mutation, with Thr$\beta/\delta 12\rightarrow$Ala possibly contributing a smaller effect. The strong sequence conservation around the site of the mutation suggests that minor changes such as the additional hydrogen bond between His$\alpha_1 50$ and Asp$\beta/\delta_1 125$ (noted above) may play a role. The replacement of His$\beta/\delta 2$ by Asn will compensate to some extent for the additional positive charge within the central cavity, but this residue lies at the mouth of the cavity over 20 Å from Gln$\beta/\delta 101$. Within 12 Å of this residue, HbA has a glutamine at position $\beta 131$(H9) which is replaced by glutamate in mammoth and elephant. This has the effect of restoring charge balance within the tetramer and probably stabilizes the protein to the Glu$\beta/\delta 101\rightarrow$Gln mutation. [Human Hb carrying the Gln$\beta 101$ mutation (Hb Rush) is slightly unstable.] At the same time, this extra negative charge within Asian elephant Hb may explain its weak response to chloride ions. Ancestral (and modern) elephant Hb therefore appears to carry a $\beta/\delta 131$ mutation predisposing it to develop a large chloride effect through a single-residue change at $\beta/\delta 101$, while having almost no chloride effect of its own. The structures of woolly mammoth Hb highlight how poorly we understand the chloride binding of different Hbs and suggest that different animal Hbs have evolved quite different interactions with this fundamental heterotropic effector.

References

Adams, J. G., Winter, W. P., Tausk, K. & Heller, P. (1974). *Blood*, **43**, 261–269.

Arnone, A. (1972). *Nature (London)*, **237**, 146–149.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.

Baldwin, J. & Chothia, C. (1979). *J. Mol. Biol.* **129**, 175–220.

Bohr, C., Hasselbalch, K. & Krogh, A. (1904). *Skand. Arch. Physiol.* **16**, 401–412.

Busch, M. R., Mace, J. E., Ho, N. T. & Ho, C. (1991). *Biochemistry*, **30**, 1865–1877.

Campbell, K. L., Roberts, J. E., Watson, L. N., Stetefeld, J., Sloan, A. M., Signore, A. V., Howatt, J. W., Tame, J. R. H., Rohland, N., Shen, T. J., Austin, J. J., Hofreiter, M., Ho, C., Weber, R. E. & Cooper, A. (2010). *Nature Genet.* **42**, 536–540.

Chervenak, M. C. & Toone, E. J. (1995). *Biochemistry*, **34**, 5685–5695.

DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Fang, T.-Y., Zou, M., Simplaceanu, V., Ho, N. T. & Ho, C. (1999). *Biochemistry*, **38**, 13423–13432.

Fronticelli, C. (1990). *Biophys. Chem.* **37**, 141–146.

Fronticelli, C., Sanna, M. T., Perez-Alvarado, G. C., Karavitis, M., Lu, A. L. & Brinigar, W. S. (1995). *J. Biol. Chem.* **270**, 30588–30592.

Fung, L. W. & Ho, C. (1975). *Biochemistry*, **14**, 2526–2535.

Giardina, B., Condò, S. G., Petruzzelli, R., Bardgard, A. & Brix, O. (1990). *Biophys. Chem.* **37**, 281–286.

Gong, Q., Simplaceanu, V., Lukin, J. A., Giovannelli, J. L., Ho, N. T. & Ho, C. (2006). *Biochemistry*, **45**, 5140–5148.

Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.

Hiebl, I., Weber, R. E., Schneeganss, D. & Braunitzer, G. (1989). *Biol. Chem. Hoppe Seyler*, **370**, 699–706.

Ho, C., Yuan, Y. & Simplaceanu, V. (2011). *Hemoglobin: Recent Developments and Topics*, edited by M. Nagai, pp. 17–36. Kerala: Research Signpost.

Hub, J. S., Kubitzki, M. B. & de Groot, B. L. (2010). *PLoS Comput. Biol.* **6**, e1000774.

Imai, K. (1982). *Allosteric Effects in Haemoglobin*. Cambridge University Press.

Jessen, T. H., Weber, R. E., Fermi, G., Tame, J. & Braunitzer, G. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 6519–6522.

Lukin, J. A., Kontaxis, G., Simplaceanu, V., Yuan, Y., Bax, A. & Ho, C. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 517–520.

Monod, J., Wyman, J. & Changeux, J. P. (1965). *J. Mol. Biol.* **12**, 88–118.

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.* D**67**, 355–367.

O'Donnell, S., Mandaro, R., Schuster, T. M. & Arnone, A. (1979). *J. Biol. Chem.* **254**, 12204–12208.

Opazo, J. C., Sloan, A. M., Campbell, K. L. & Storz, J. F. (2009). *Mol. Biol. Evol.* **26**, 1469–1478.

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. (1970). *Nature (London)*, **228**, 726–739.

Perutz, M. F. (1972). *Nature (London)*, **237**, 495–499.

Perutz, M. F. (1983). *Mol. Biol. Evol.* **1**, 1–28.

Perutz, M. F., Fermi, G., Poyart, C., Pagnier, J. & Kister, J. (1993). *J. Mol. Biol.* **233**, 536–545.

Perutz, M. F. & Imai, K. (1980). *J. Mol. Biol.* **136**, 183–191.

Perutz, M. F., Shih, D. T. & Williamson, D. (1994). *J. Mol. Biol.* **239**, 555–560.

Perutz, M. F., Wilkinson, A. J., Paoli, M. & Dodson, G. G. (1998). *Annu. Rev. Biophys. Biomol. Struct.* **27**, 1–34.

Razynska, A., Fronticelli, C., Di Cera, E., Gryczynski, Z. & Bucci, E. (1990). *Biophys. Chem.* **38**, 111–115.

Safo, M. K. & Abraham, D. J. (2001). *Protein Sci.* **10**, 1091–1099.

Safo, M. K. & Abraham, D. J. (2005). *Biochemistry*, **44**, 8347–8359.

Schay, G., Smeller, L., Tsuneshige, A., Yonetani, T. & Fidy, J. (2006). *J. Biol. Chem.* **281**, 25972–25983.

Shaanan, B. (1983). *J. Mol. Biol.* **171**, 31–59.

Shibayama, N., Miura, S., Tame, J. R. H., Yonetani, T. & Park, S.-Y. (2002). *J. Biol. Chem.* **277**, 38791–38796.

Shih, D. T.-B., Jones, R. T., Imai, K. & Tyuma, I. (1985). *J. Biol. Chem.* **260**, 5919–5924.

Signore, A. V., Stetefeld, J., Weber, R. E. & Campbell, K. L. (2012). *J. Exp. Biol.* **215**, 518–525.

Silva, M. M., Rogers, P. H. & Arnone, A. (1992). *J. Biol. Chem.* **267**, 17248–17256.

Srinivasan, R. & Rose, G. D. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 11113–11117.

Sun, D. P., Zou, M., Ho, N. T. & Ho, C. (1997). *Biochemistry*, **36**, 6663–6673.

Tame, J. R. H. (1999). *Trends Biochem. Sci.* **24**, 372–377.

Tame, J. R. H. & Vallone, B. (2000). *Acta Cryst.* D**56**, 805–811.

Tame, J. R. H., Wilson, J. C. & Weber, R. E. (1996). *J. Mol. Biol.* **259**, 749–760.

Tsuneshige, A., Park, S. & Yonetani, T. (2002). *Biophys. Chem.* **98**, 49–63.

Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.

Weber, R. E. (2007). *Respir. Physiol. Neurobiol.* **158**, 132–142.

Weber, R. E. & Campbell, K. L. (2011). *Acta Physiol. (Oxf.)*, **202**, 549–562.

Weber, R. E., Ostojic, H., Fago, A., Dewilde, S., Van Hauwaert, M. L., Moens, L. & Monge, C. (2002). *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **283**, R1052–R1060.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Yokoyama, T., Chong, K. T., Miyazaki, G., Morimoto, H., Shih, D. T.-B., Unzai, S., Tame, J. R. H. & Park, S.-Y. (2004). *J. Biol. Chem.* **279**, 28632–28640.

Yokoyama, T., Neya, S., Tsuneshige, A., Yonetani, T., Park, S.-Y. & Tame, J. R. H. (2006). *J. Mol. Biol.* **356**, 790–801.

Yonetani, T., Park, S. I., Tsuneshige, A., Imai, K. & Kanaori, K. (2002). *J. Biol. Chem.* **277**, 34508–34520.

Yonetani, T. & Tsuneshige, A. (2003). *C. R. Biol.* **326**, 523–532.

Yuan, Y., Shen, T.-J., Gupta, P., Ho, N. T., Simplaceanu, V., Tam, T. C. S., Hofreiter, M., Cooper, A., Campbell, K. L. & Ho, C. (2011). *Biochemistry*, **50**, 7350–7360.