Chen, J., Morral, N. & Engel, D. A. (2007). Transcription releases protein VII from adenovirus chromatin. *Virology* **369**, 411–422.

Daniell, E., Groff, D. E. & Fedor, M. J. (1981). Adenovirus chromatin structure at different stages of infection. *Mol Cell Biol* **1**, 1094–1105.

Déry, C. V., Toth, M., Brown, M., Horvath, J., Allaire, S. & Weber, J. M. (1985). The structure of adenovirus chromatin in infected cells. *J Gen Virol* **66**, 2671–2684.

Fedor, M. J. & Daniell, E. (1983). Ionic and nonionic interactions in adenoviral nucleoprotein complexes. *J Virol* **47**, 370–375.

Greber, U. F., Webster, P., Weber, J. & Helenius, A. (1996). The role of the adenovirus protease on virus entry into cells. *EMBO J* **15**, 1766–1777.

Gyurcsik, B., Haruki, H., Takahashi, T., Mihara, H. & Nagata, K. (2006). Binding modes of the precursor of adenovirus major core protein VII to DNA and template activating factor I: implication for the mechanism of remodeling of the adenovirus chromatin. *Biochemistry* **45**, 303–313.

Haruki, H., Gyurcsik, B., Okuwaki, M. & Nagata, K. (2003). Ternary complex formation between DNA–adenovirus core protein VII and TAF-Iβ/SET, an acidic molecular chaperone. *FEBS Lett* **555**, 521–527.

Haruki, H., Okuwaki, M., Miyagishi, M., Taira, K. & Nagata, K. (2006). Involvement of template-activating factor I/SET in transcription of adenovirus early genes as a positive-acting factor. *J Virol* **80**, 794–801.

Hindley, C. E., Davidson, A. D. & Matthews, D. A. (2007). Relationship between adenovirus DNA replication proteins and nucleolar proteins B23.1 and B23.2. *J Gen Virol* **88**, 3244–3248.

Hingorani, K., Szebeni, A. & Olson, M. O. (2000). Mapping the functional domains of nucleolar protein B23. *J Biol Chem* **275**, 24451–24457.

Hisaoka, M., Ueshima, S., Murano, K., Nagata, K. & Okuwaki, M. (2010). Regulation of nucleolar chromatin by B23/nucleophosmin jointly depends upon its RNA binding activity and transcription factor UBF. *Mol Cell Biol* **30**, 4952–4964.

Johnson, J. S., Osheim, Y. N., Xue, Y., Emanuel, M. R., Lewis, P. W., Bankovich, A., Beyer, A. L. & Engel, D. A. (2004). Adenovirus protein VII condenses DNA, represses transcription, and associates with transcriptional activator E1A. *J Virol* **78**, 6459–6468.

Kawase, H., Okuwaki, M., Miyaji, M., Ohba, R., Handa, H., Ishimi, Y., Fujii-Nakata, T., Kikuchi, A. & Nagata, K. (1996). NAP-I is a functional homologue of TAF-I that is required for replication and transcription of the adenovirus genome in a chromatin-like structure. *Genes Cells* **1**, 1045–1056.

Komatsu, T., Haruki, H. & Nagata, K. (2011). Cellular and viral chromatin proteins are positive factors in the regulation of adenovirus gene expression. *Nucleic Acids Res* **39**, 889–901.

Levy, A. & Noll, M. (1981). Chromatin fine structure of active and repressed genes. *Nature* **289**, 198–203.

Martin-Fernandez, M., Longshaw, S. V., Kirby, I., Santis, G., Tobin, M. J., Clarke, D. T. & Jones, G. R. (2004). Adenovirus type-5 entry and disassembly followed in living cells by FRET, fluorescence anisotropy, and FLIM. *Biophys J* **87**, 1316–1327.

Matsumoto, K., Nagata, K., Ui, M. & Hanaoka, F. (1993). Template activating factor I, a novel host factor required to stimulate the adenovirus core DNA replication. *J Biol Chem* **268**, 10582–10587.

Matsumoto, K., Okuwaki, M., Kawase, H., Handa, H., Hanaoka, F. & Nagata, K. (1995). Stimulation of DNA transcription by the replication factor from the adenovirus genome in a chromatin-like structure. *J Biol Chem* **270**, 9645–9650.

Matthews, D. A. (2001). Adenovirus protein V induces redistribution of nucleolin and B23 from nucleolus to cytoplasm. *J Virol* **75**, 1031–1038.

Murano, K., Okuwaki, M., Hisaoka, M. & Nagata, K. (2008). Transcription regulation of the rRNA gene by a multifunctional nucleolar protein, B23/nucleophosmin, through its histone chaperone activity. *Mol Cell Biol* **28**, 3114–3126.

Nagata, K., Kawase, H., Handa, H., Yano, K., Yamasaki, M., Ishimi, Y., Okuda, A., Kikuchi, A. & Matsumoto, K. (1995). Replication factor encoded by a putative oncogene, *set*, associated with myeloid leukemogenesis. *Proc Natl Acad Sci U S A* **92**, 4279–4283.

Nakanishi, Y., Maeda, K., Ohtsuki, M., Hosokawa, K. & Natori, S. (1986). *In vitro* transcription of a chromatin-like complex of major core protein VII and DNA of adenovirus serotype 2. *Biochem Biophys Res Commun* **136**, 86–93.

Nakano, M. Y., Boucke, K., Suomalainen, M., Stidwill, R. P. & Greber, U. F. (2000). The first step of adenovirus type 2 disassembly occurs at the cell surface, independently of endocytosis and escape to the cytosol. *J Virol* **74**, 7085–7095.

Okuda, M., Horn, H. F., Tarapore, P., Tokuyama, Y., Smulian, A. G., Chan, P.-K., Knudsen, E. S., Hofmann, I. A., Snyder, J. D. & other authors (2000). Nucleophosmin/B23 is a target of CDK2/cyclin E in centrosome duplication. *Cell* **103**, 127–140.

Okuwaki, M., Iwamatsu, A., Tsujimoto, M. & Nagata, K. (2001a). Identification of nucleophosmin/B23, an acidic nucleolar pro tein, as a stimulatory factor for *in vitro* replication of adenovirus DNA complexed with viral basic core proteins. *J Mol Biol* **311**, 41–55.

Okuwaki, M., Matsumoto, K., Tsujimoto, M. & Nagata, K. (2001b). Function of nucleophosmin/B23, a nucleolar acidic protein, as a histone chaperone. *FEBS Lett* **506**, 272–276.

Okuwaki, M., Tsujimoto, M. & Nagata, K. (2002). The RNA binding activity of a ribosome biogenesis factor, nucleophosmin/B23, is modulated by phosphorylation with a cell cycle-dependent kinase and by association with its subtype. *Mol Biol Cell* **13**, 2016–2030.

Okuwaki, M., Kato, K., Shimahara, H., Tate, S. & Nagata, K. (2005). Assembly and disassembly of nucleosome core particles containing histone variants by human nucleosome assembly protein I. *Mol Cell Biol* **25**, 10639–10651.

Samad, M. A., Okuwaki, M., Haruki, H. & Nagata, K. (2007). Physical and functional interaction between a nucleolar protein nucleophosmin/B23 and adenovirus basic core proteins. *FEBS Lett* **581**, 3283–3288.

Savkur, R. S. & Olson, M. O. (1998). Preferential cleavage in pre-ribosomal RNA by protein B23 endoribonuclease. *Nucleic Acids Res* **26**, 4508–4515.

Sergeant, A., Tigges, M. A. & Raskas, H. J. (1979). Nucleosome-like structural subunits of intranuclear parental adenovirus type 2 DNA. *J Virol* **29**, 888–898.

Spector, D. J. (2007). Default assembly of early adenovirus chromatin. *Virology* **359**, 116–125.

Sung, M. T., Lischwe, M. A., Richards, J. C. & Hosokawa, K. (1977). Adenovirus chromatin I. Isolation and characterization of the major core protein VII and precursor Pro-VII. *J Biol Chem* **252**, 4981–4987.

Sung, M. T., Cao, T. M., Coleman, R. T. & Budelier, K. A. (1983). Gene and protein sequences of adenovirus protein VII, a hybrid basic chromosomal protein. *Proc Natl Acad Sci U S A* **80**, 2902–2906.

Tate, V. E. & Philipson, L. (1979). Parental adenovirus DNA accumulates in nucleosome-like structures in infected cells. *Nucleic Acids Res* **6**, 2769–2785.

Trotman, L. C., Mosberger, N., Fornerod, M., Stidwill, R. P. & Greber, U. F. (2001). Import of adenovirus DNA involves the nuclear pore complex receptor CAN/Nup214 and histone H1. *Nat Cell Biol* **3**, 1092–1100.

Xue, Y., Johnson, J. S., Ornelles, D. A., Lieberman, J. & Engel, D. A. (2005). Adenovirus protein VII functions throughout early phase and interacts with cellular proteins SET and pp32. *J Virol* **79**, 2474–2483.

Yu, Y., Maggi, L. B., Jr, Brady, S. N., Apicelli, A. J., Dai, M. S., Lu, H. & Weber, J. D. (2006). Nucleophosmin is essential for ribosomal protein L5 nuclear export. *Mol Cell Biol* **26**, 3798–3809.

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

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

Received 19 April 2012
Accepted 28 June 2012

PDB References: woolly mammoth haemoglobin, deoxy form, 3vre; carbonmonoxy form, 3vrf; met form, 3vrg

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 aquomet 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-diphosphoglycerate (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, controlling 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

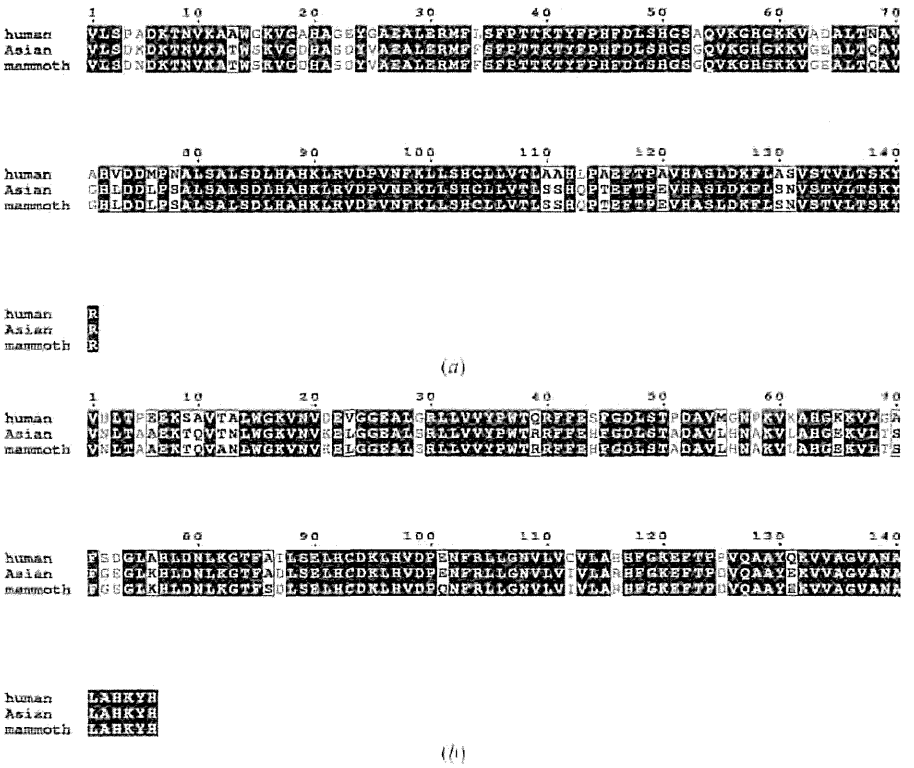

Figure 1
(*a*) An alignment of the α-globin sequences of human HbA, Asian (Indian) elephant (*Elephas maximus*) and woolly mammoth (*Mammothus primigenius*). The sequences from *E. maximus* and mammoth differ at one residue (Asn5→Lys), and similarly the modern African elephant (*Loxodonta africana*) 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	*C*2	*C*2	*P*1
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 (Å^3)	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 m*M* 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···His $\beta/\delta_2$146	—	—	2.6 (2.8)
Tyr$\alpha_1$42···Asp$\beta/\delta_2$99	—	—	2.6 (2.6)
Asp$\alpha_1$94···Trp$\beta/\delta_2$37	3.7	3.7	2.9 (3.2)
Asp$\alpha_1$94···Asn$\beta/\delta_2$102	2.7	2.8	—
Arg$\alpha_1$141···Asp$\beta/\delta_2$126	—	—	2.7 (2.4)
Asp$\beta/\delta_1$94···His$\beta/\delta_1$146	—	—	2.9 (2.7)
Trp$\beta/\delta_1$37···Asn$\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^α 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^α 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 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.

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^α 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

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^α 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.

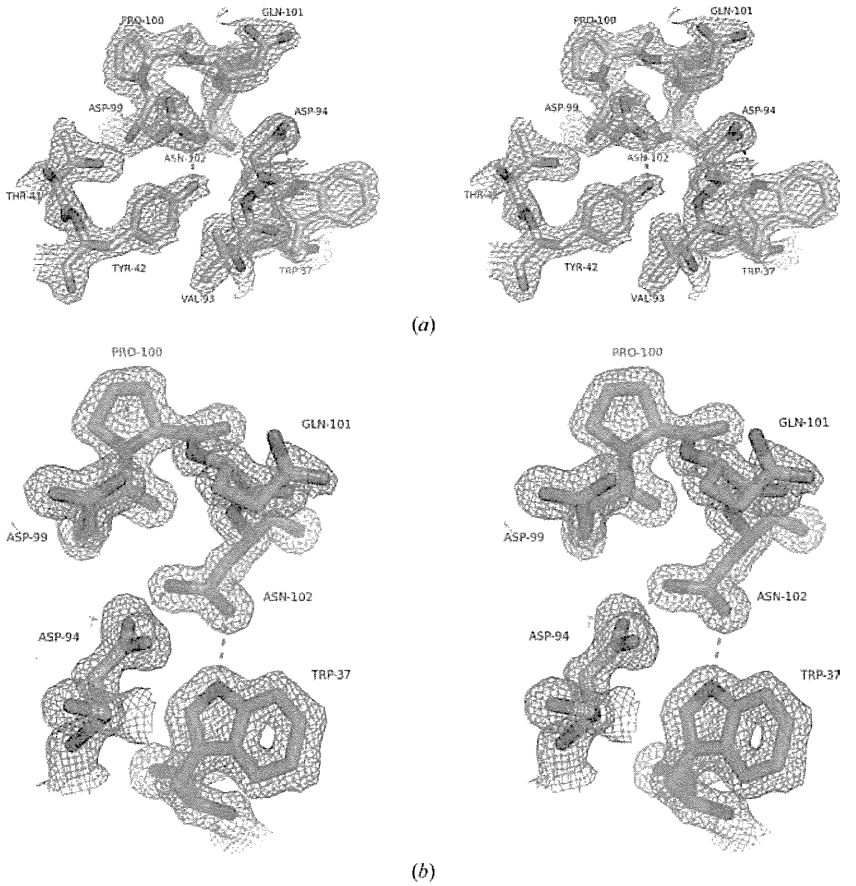

Figure 2
(*a*) A stereoview of the $\alpha_1\beta/\delta_2$ interface of deoxy woolly mammoth Hb showing the $2mF_o - DF_c$ electron-density map at a contour level of 1σ. The structure is shown as a stick model coloured by atom type: oxygen, red; nitrogen, blue. C atoms of the α subunit are coloured 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.

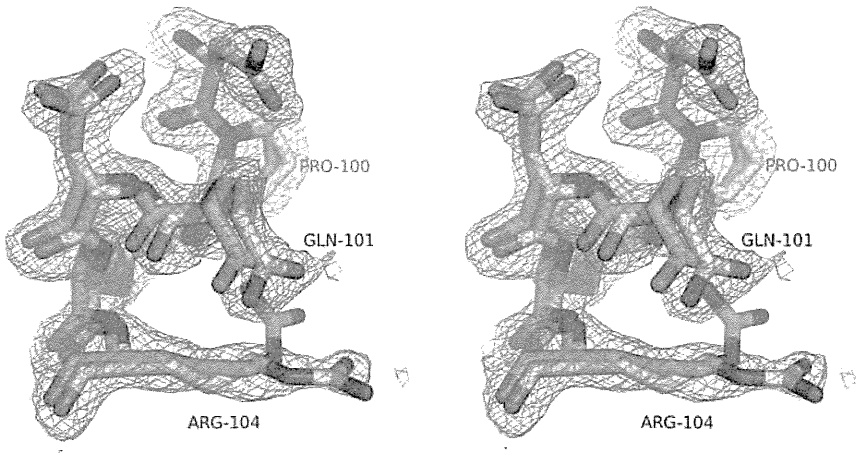

Figure 3
A stereoview showing the overlap of deoxy woolly mammoth Hb and deoxy HbA around Glnβ/δ101. C atoms of the human protein are coloured cyan and the $2mF_o - DF_c$ electron-density map is shown at a contour level of 1σ. Argβ/δ104 adopts a more extended conformation in mammoth Hb, pointing towards 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

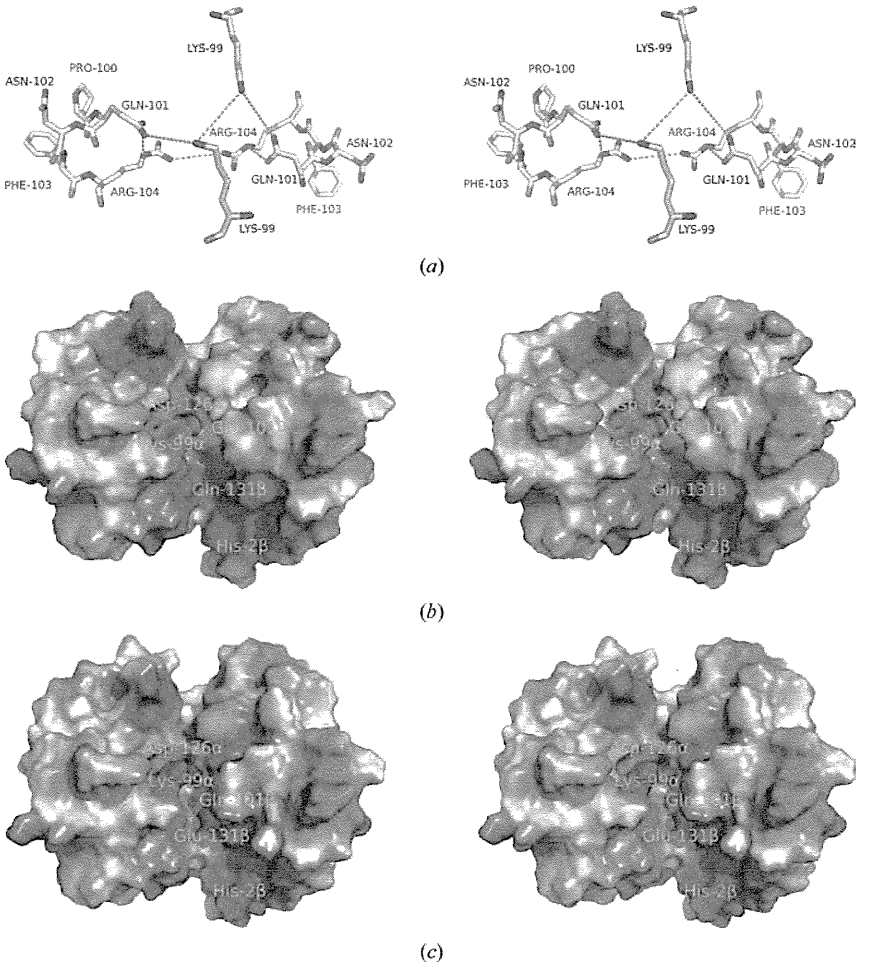

Figure 4
(*a*) A stereoview of the central cavity of deoxy woolly mammoth Hb. C atoms of the α subunit are coloured white and C atoms of the β/δ subunit are coloured yellow. Orange dotted lines indicate distances of between 3 and 4 Å in length and light blue dotted lines indicate distances of between 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 m*M* 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. 5*a*). The C^{α} 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. 5*b*) 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.

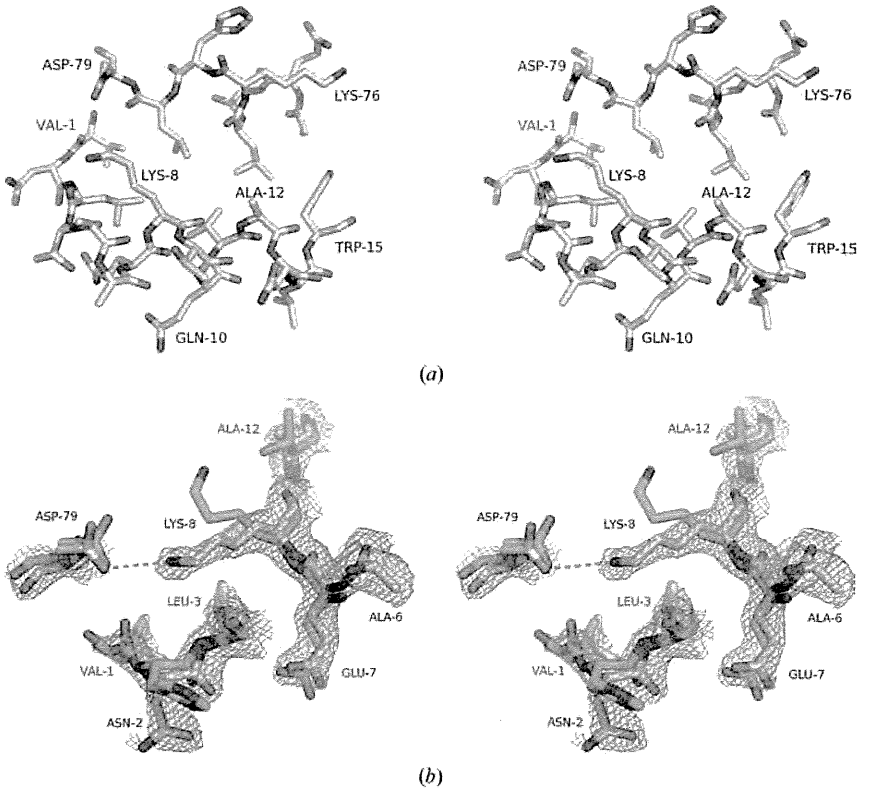

Figure 5
(*a*) The surface of the β/δ subunit of deoxy woolly mammoth Hb, showing the proposed chloride ion-binding site of bovine Hb (Fronticelli *et al.*, 1995). In the subunit shown, Lys8 makes no interaction with Asp79 of the same subunit. (*b*) Stereoview of the deoxy woolly mammoth Hb β/δ subunit 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 m*M* chloride (Perutz *et al.*, 1994; Fronticelli *et al.*, 1995). Mutagenic experiments have implicated the lysine residues at positions β8 and β76 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 β76, 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, Serα131(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. Serα131(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 β/δ139 (see below). Comparing elephant and mammoth Hb, it can be seen that the mammoth protein is more sensitive to chloride, while retaining Asnα131(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 Glnβ/δ101 mutation, with Thrβ/δ12$\rightarrow$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$125 (noted above) may play a role. The replacement of Hisβ/δ2 by Asn will compensate to some extent for the additional positive charge within the central cavity, but this residue lies at the mouth of the cavity over 20 Å from Glnβ/δ101. Within 12 Å of this residue, HbA has a glutamine at position β131(H9) which is replaced by glutamate in mammoth and elephant. This has the effect of restoring charge balance within the tetramer and probably stabilizes the protein to the Gluβ/δ101$\rightarrow$Gln mutation. [Human Hb carrying the Glnβ101 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 β/δ131 mutation predisposing it to develop a large chloride effect through a single-residue change at β/δ101, 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.

http://dx.doi.org/10.1016/j.jmb.2012.07.012 J. Mol. Biol. (2012) 423, 351–364

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

Crystal Structures of Penicillin-Binding Protein 3 (PBP3) from Methicillin-Resistant *Staphylococcus aureus* in the Apo and Cefotaxime-Bound Forms

Hisashi Yoshida[1], Fumihiro Kawai[1], Eiji Obayashi[1], Satoko Akashi[2], David I. Roper[3], Jeremy R. H. Tame[1]* and Sam-Yong Park[1]*

[1]*Protein Design Laboratory, Yokohama City University, Suehiro 1-7-29, Tsurumi-ku, Yokohama 230–0045, Japan*
[2]*Structural Biology Laboratory, Yokohama City University, Suehiro 1-7-29, Tsurumi-ku, Yokohama 230–0045, Japan*
[3]*School of Life Sciences, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*

Received 30 April 2012;
received in revised form
4 July 2012;
accepted 16 July 2012
Available online
27 July 2012

Edited by R. Huber

Keywords:
antibiotic;
MRSA;
protein complex;
analytical ultracentrifugation;
mass spectrometry

Staphylococcus aureus is a widespread Gram-positive opportunistic pathogen, and a methicillin-resistant form (MRSA) is particularly difficult to treat clinically. We have solved two crystal structures of penicillin-binding protein (PBP) 3 (PBP3) from MRSA, the apo form and a complex with the β-lactam antibiotic cefotaxime, and used electrospray mass spectrometry to measure its sensitivity to a variety of penicillin derivatives. PBP3 is a class B PBP, possessing an N-terminal non-penicillin-binding domain, sometimes called a dimerization domain, and a C-terminal transpeptidase domain. The model shows a different orientation of its two domains compared to earlier models of other class B PBPs and a novel, larger N-domain. Consistent with the nomenclature of "dimerization domain", the N-terminal region forms an apparently tight interaction with a neighboring molecule related by a 2-fold symmetry axis in the crystal structure. This dimer form is predicted to be highly stable in solution by the PISA server, but mass spectrometry and analytical ultracentrifugation provide unequivocal evidence that the protein is a monomer in solution.

Introduction

Peptidoglycan is the principal component of the bacterial cell wall synthesized by penicillin-binding proteins (PBPs).[1,2] In the case of the Gram-positive *Staphylococcus aureus*, it forms a layer approximately 30 nm thick around the cell. *S. aureus* commonly colonizes human skin and upper respiratory tract, especially the nasal passages, but may cause serious infections under certain circumstances, such as injury or surgery. A wide variety of clinical conditions are due to *S. aureus* infection, which may also produce toxic shock syndrome. Strains resistant to benzylpenicillin arose in the 1950s but were successfully treated with methicillin, a simple derivative of benzylpenicillin carrying two extra methoxy groups that rendered the drug less readily degradable by β-lactamases. MRSA (*m*ethicillin-*r*esistant *S. aureus*) initially arose due to the acquisition of the *mecA* gene that encodes PBP2a,[3] and now multi-resistant MRSA presents a very significant clinical challenge.[4] Methicillin resistance may alternatively be due to the secretion of β-lactamases.[5] Since peptidoglycan is

*Corresponding authors. E-mail addresses:
jtame@tsurumi.yokohama-cu.ac.jp;
park@tsurumi.yokohama-cu.ac.jp.

Present addresses: F. Kawai, Department of Biology and Biochemistry, University of Houston, Houston, TX 77204, USA; E. Obayashi, Department of Biochemistry, Shimane University School of Medicine, Izumo 693–8501, Japan.

Abbreviations used: PBP, penicillin-binding protein; Se-Met, selenomethionine; TEV, tobacco etch virus; nanoESI, nanoelectrospray ionization; PDB, Protein Data Bank.

0022-2836/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.

almost invariably required for bacterial viability and absent from eukaryotes, PBPs remain attractive drug targets. In both Gram-positive and Gram-negative bacteria, the stem peptide of the peptidoglycan precursor ends with two D-alanine residues, the last of which is released in the transpeptidation reaction. β-Lactam antibiotics are suicide substrates that mimic this dipeptide,[6] forming a covalent adduct with PBPs that blocks further transpeptidase or carboxypeptidase activity.

The transpeptidation active sites of all PBPs studied to date have several strictly conserved motifs. The active-site serine residue is found at the start of helix α2 (by the standard naming convention), three residues before a conserved lysine, to make an SxxK motif. A second motif, SxN, is found on a loop between helices α4 and α5, and a KTG motif is located very close to the active site on strand β3. Resistance to β-lactams may arise through a variety of mechanisms, such as the expression of hydrolases that rapidly destroy the drug. Intrinsic resistance may also develop within the PBPs themselves, lowering their affinity for substrates sufficiently to function in the presence of antibiotics. These mutations may operate through rather indirect mechanisms, such as altered protein flexibility, which are not immediately apparent from simple visual inspection of protein models.[7]

PBPs are classed in two major groups by size.[8] Low-molecular-weight PBPs are single-domain proteins that principally function as DD-carboxypeptidases to restrict the level of cross-linking; they are often nonessential.[9] HMW PBPs (*h*igh-*m*olecular-*w*eight *PBPs*) carry several domains, only one of which acts on peptidoglycan peptides. Class A HMW PBPs also possess the essential transglycosylation activity, whereas the extra domain of class B HMW PBPs has unknown function. *S. aureus* has four PBPs (PBP1–PBP4), of which only PBP4 is a low-molecular-weight PBP. It has been implicated in the drug resistance of community-acquired MRSA,[10] and its crystal structure has been determined.[11] PBP3 (encoded by the *pbpC* gene) has a C-terminal penicillin-binding domain, and its N-terminal domain is well conserved when compared to other species, so that *S. aureus* PBP3 is found almost unaltered in several species of *Bacillus* (such as *Bacillus cereus* and *Bacillus anthracis*) and has 44% identity with PBP2a of *Bacillus subtilis*.[12] Note that the numbering of PBPs arose historically from migration by SDS-PAGE and, therefore, may not reflect sequence similarity across species—PBP3 of *S. aureus* is similar to PBP2 of *Escherichia coli*, and PBP3 of *E. coli* is similar to PBP1 of *S. aureus*. PBP3 of *E. coli* is part of the divisome and required for construction of the septum prior to cell division; it is associated with the flippase FtsW[13] and other enzymes with roles in cell wall synthesis.[14,15] PBP3 of *S. aureus* was reported to be involved in cell division and essential,[16] but later work with a deletion mutant showed that it is not required for cell survival.[12] The deletion mutant showed normal growth with no antibiotics present but severe shape abnormalities and septation defects under sublethal concentrations of methicillin, suggesting that PBP1 and PBP2 may functionally replace PBP3.[12] PBP3 is more sensitive to methicillin than either PBP1 or PBP2, but loss of PBP3 appears to increase methicillin resistance to some degree.

Crystal structures are known for class A and class B PBPs including *E. coli* (PBP1b[17]), *Neisseria gonorrhoeae* (NG-PBP2[18]), *Pseudomonas aeruginosa* (PA-PBP3[19,20]), *Streptococcus pneumoniae* (SP-PBP1a,[21] SP-PBP2x,[22] and SP-PBP2b[7]), and *S. aureus* (SA-PBP2[23] and SA-PBP2a[24]). In this paper, we report the crystal structure of PBP3 from *S. aureus*, determined using heavy atoms and anomalous scattering, and compare it with previously refined HMW PBP models. We also report the reactivity of SA-PBP3 with a variety of β-lactams, measured using mass spectrometry, and try to relate these to the protein structure.

Results

A plasmid was created to express, in *E. coli*, the portion of the *S. aureus* PBP3 gene expected to encode the structured region of the protein, omitting the N-terminal signal peptide and adjacent hydrophobic residues believed to anchor the protein to the cell membrane. This construct permitted high-level cytoplasmic expression of soluble protein, including the sequence from residue 46 to the C-terminus (residue 691), with a cleavable N-terminal histidine tag. Crystals were grown of the apo form. The crystals diffracted to a maximum resolution of 2.3 Å and contain two copies of the protein in the asymmetric unit. Phasing was attempted by molecular replacement using the known model of PBP3 from *S. pneumoniae* (SP-PBP2b)[7] and *P. aeruginosa* (PA-PBP3),[19] but this failed to give convincing solutions using the entire structure as a search model. Solutions were obtained using the C-terminal β-lactam-binding domain alone, but not the N-domain, and molecular replacement ultimately failed to yield a satisfactory model. Therefore, crystals were prepared with selenomethionine (Se-Met) in place of methionine or with bound heavy atoms to allow experimental phase determination. The protein model was constructed *de novo* without reference to earlier models and was largely built with the automatic tools in the PHENIX suite.[25] Although much of the model is covered by electron density of high quality, several turn regions show notably poorly defined geometry and are apparently highly flexible. No residues are

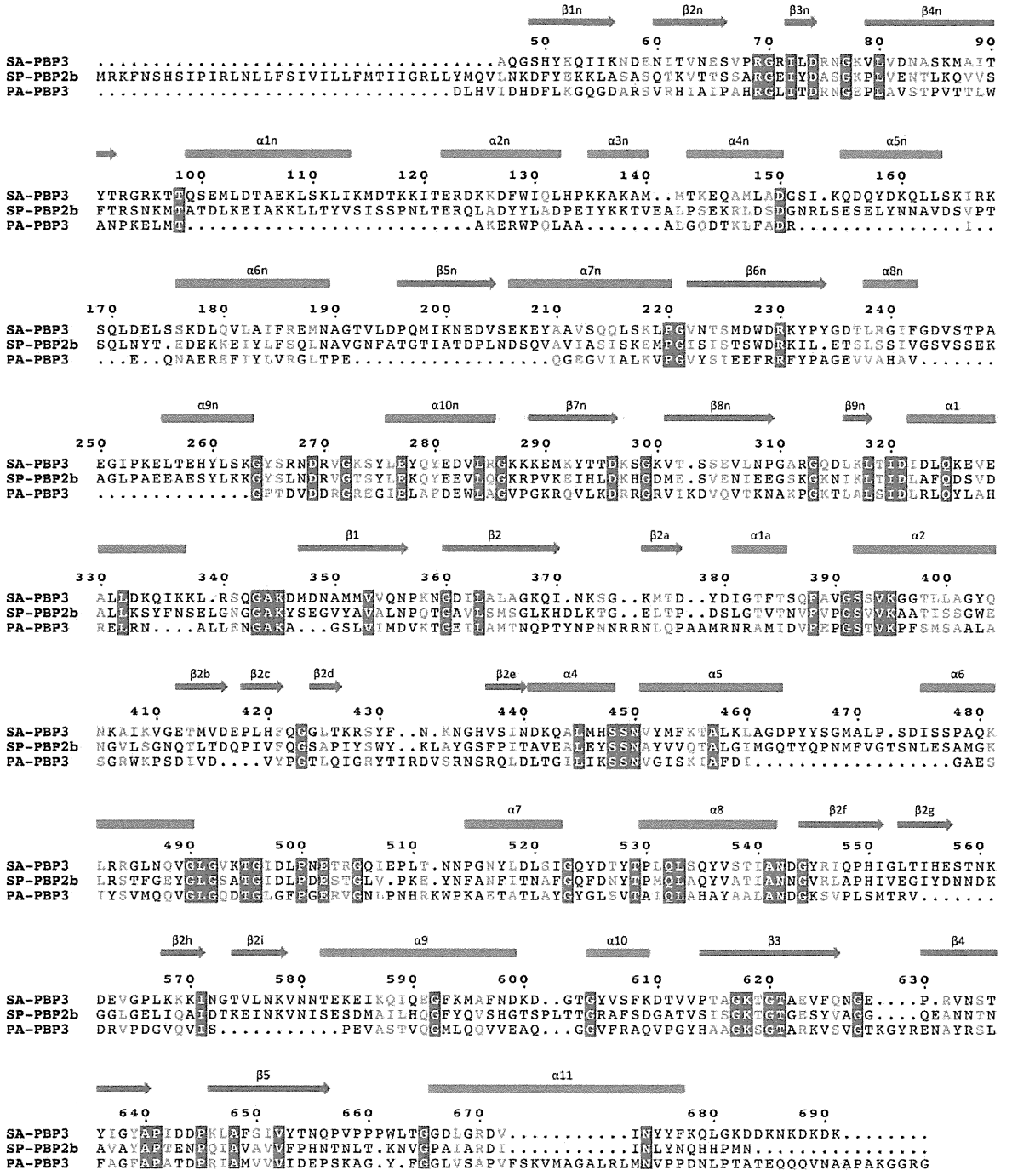

Fig. 1. A sequence and structural alignment of MRSA PBP3 with SP-PBP2b (*S. pneumoniae* R6) and PA-PBP3 (*P. aeruginosa*). α-Helices in the PBP3 model are indicated with red bars, and β-sheets are indicated with blue arrows. Residues common to all three sequences are shown in white on blue. The conserved SxxK, SxN, and KTG motifs at the active site are found at residues 392–395, 448–450, and 618–620, respectively. The active-site serine is Ser392. The figure was made with ESPript.[26]

missing from the model except residues 627–628 (chain A) and 602 (chain B) and at the C-terminus from residue 678 onwards. The secondary structure is shown in Fig. 1.

The crystal structure shows SA-PBP3 to resemble strongly other class B PBPs, having two domains, a C-terminal transpeptidase domain, and an N-terminal nonenzymatic domain. Overall, the two copies

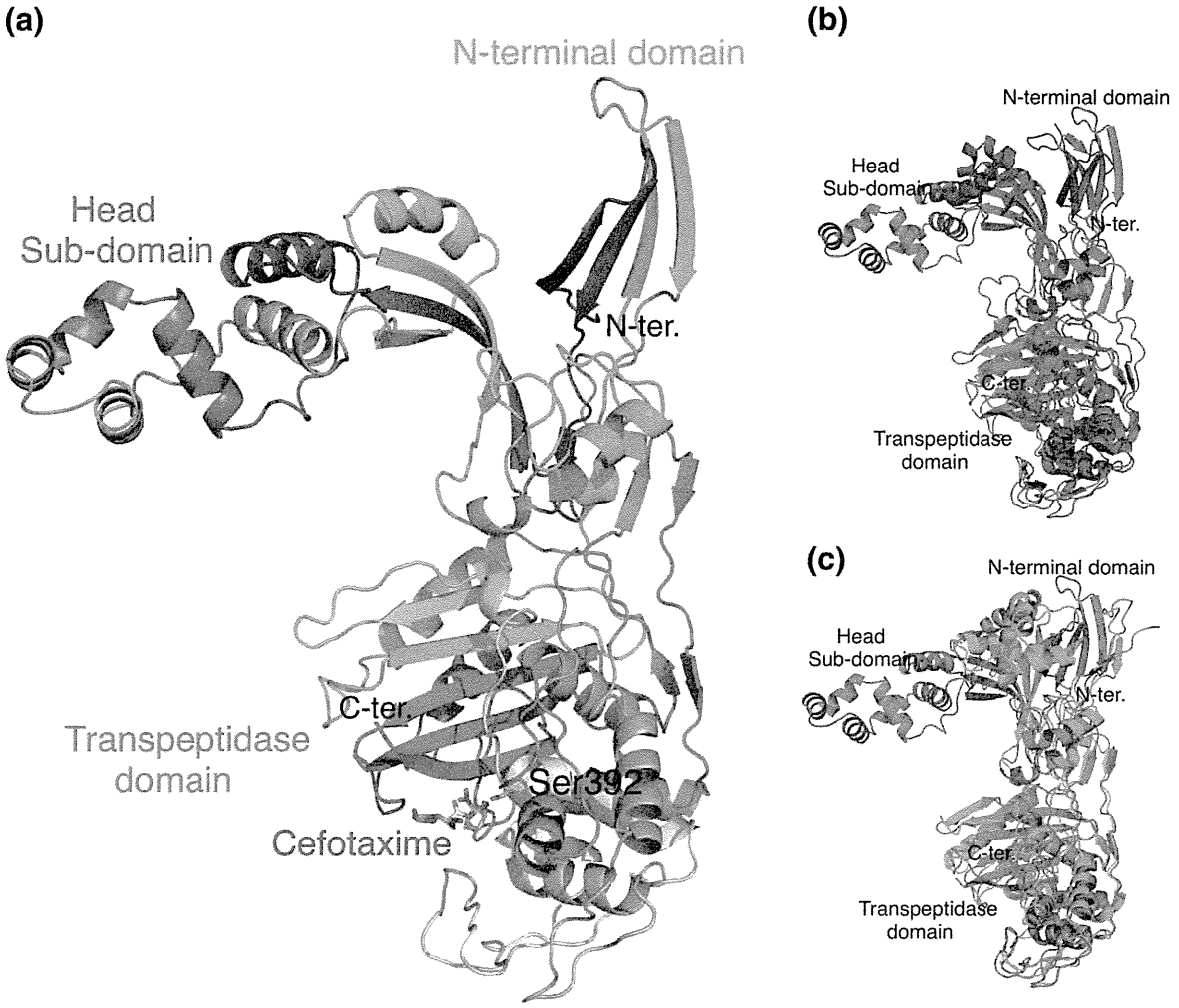

Fig. 2. The overall structure of PBP3. (a) The C^{α} trace of SA-PBP3, colored from blue (N-terminus) to red (C-terminus). The non-catalytic N-terminal domain includes a so-called head sub-domain (residues 77–237). The active site of the transpeptidation domain (Ser392) is shown with covalently bound cefotaxime. (b) Structural overlay of the overall structure of native PBP3 (red) with that of PBP3 of *P. aeruginosa* (blue; PDB ID 3OC2). Structural alignment was carried out automatically with SSM (secondary-structure *m*atching)[30] through Coot. We matched 379 residues of the penicillin-binding domain, with a sequence identity of 22.4% and a final rmsd of 2.3 Å. (c) Overlay of SA-PBP3 (red) with PBP2b from *S. pneumoniae* (green; PDB ID 2WAE). The two models overlap closely with SA-PBP3 over the penicillin-binding domain, but SA-PBP3 shows a considerably larger head domain, with a greater level of similarity to PA-PBP3 than SP-PBP2b. We overlaid 438 residues of the penicillin-binding domain with 38.4% sequence identity and 1.9 Å rmsd.

of the apoprotein in the asymmetric unit have a C^{α} rmsd of 2.6 Å. Non-crystallographic symmetry restraints were not used in the refinement. For the head domain (residues 46–76 and 238–319), the rmsd is 0.99 Å, and for the transpeptidase domain (residues 320–678), the rmsd is 1.3 Å. Comparing the A chains of the native and cefotaxime-bound models, the C^{α} rmsd is 0.39 Å, and for the B chains, the C^{α} rmsd is 0.58 Å, indicating that drug binding to the crystal causes a smaller conformational change than the inherent differences between the copies of the protein in the asymmetric unit. Structural overlays suggest that there is considerable flexibility about the inter-domain hinge of HMW PBPs,[19,27] but the non-penicillin-binding domain of *E. coli* PBP3 is required for correct folding of the protein.[28] In previously determined class B structures, the N-terminal domain includes a so-called "head domain" that is largely α-helical. In PA-PBP3, this contiguous and compact domain extends over residues 80–149,[19] includes four α-helices and three β-strands, and corresponds to the so-called PBP-dimer domain identified by Pfam.[29] It is slightly larger in PBP2b from *S. pneumoniae*[24] but substantially extended in the present SA-PBP3 structure by five extra α-helices extending from Glu122 to Ser175, apparently arising from a single insertion into an ancestral form. Although Pfam

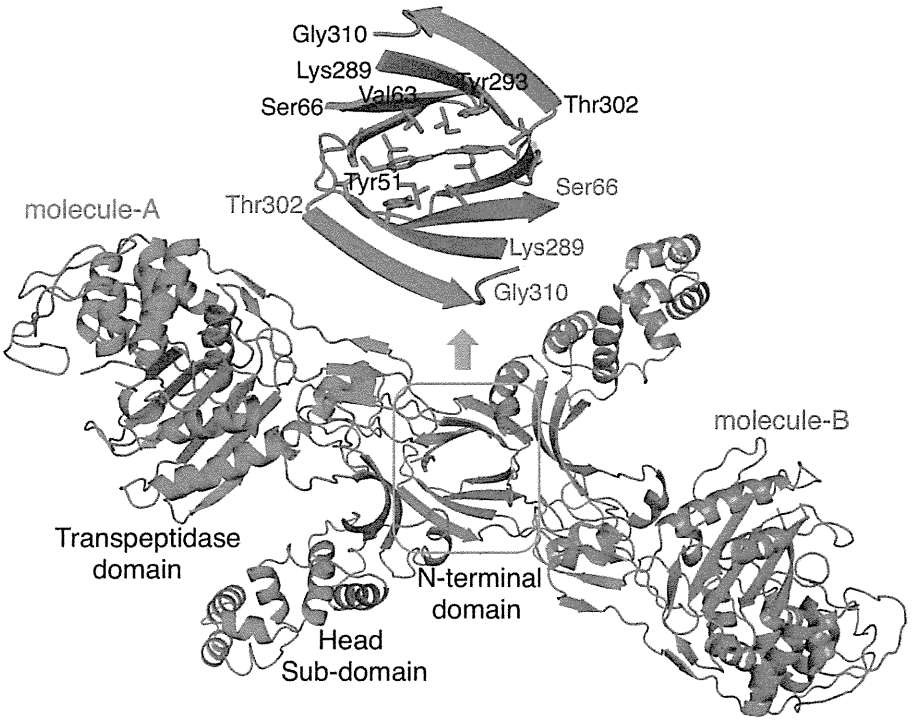

Fig. 3. The principal crystal contact between neighboring molecules. The C^{α} traces of the two molecules are colored red and blue. The self-complementary surface includes Tyr51, which hydrogen bonds to its symmetry equivalent, and also contacts Tyr293 of the opposite chain.

confidently identifies the PBP-dimer domain from the SA-PBP3 sequence, the match is only convincing over residues 179–247, which form widespread structural elements and not a compact domain. The extra C-terminal domain found in PBP2x[22] is not present in SA-PBP3 (Fig. 2).

An interesting feature of the SA-PBP3 structure is the contact formed between the two molecules in the asymmetric unit. Two β-hairpins on each molecule (residues 49–66 and 289–310) associate through non-crystallographic 2-fold symmetry to create a β-sandwich with a total buried surface area of 2730 Å^2 (Fig. 3). This interface is confidently predicted by PISA,[31] leading to stable dimer formation in solution, with a ΔG of dissociation of 6.9 kcal/mol. For comparison, PISA finds no likely oligomeric structure for PA-PBP3, which is believed to be a monomer from gel-filtration results.[19] However, analytical ultracentrifugation shows that the protein exists as a monomer in solution (Fig. 4). To confirm the solution state of the protein, we measured the mass by electrospray mass spectrometry, also indicating the absence of dimers (Fig. 4). Prediction of ligand binding affinity from buried surface areas, hydrogen bond counting, or similar analysis of static crystallographic models is popular but known to be highly unreliable.[32] SA-PBP3 is the first example, to our knowledge, of such a high predicted ΔG from the PISA server proving incorrect. The estimation of separate entropy and enthalpy changes on ligand binding from a static model of a complex, with no regard for experimental conditions, is an exercise in data fitting that is unjustified by thermodynamics.[33] While bioinformatics tools such as Pfam and PISA are without doubt highly useful, their results are not completely reliable and the "PBP-dimer" domain may, in fact, not always self-associate.

Active site

The active-site serine (Ser392) faces a groove running along strand β3 that is approximately 20 Å deep and 15 Å wide, allowing access for peptidoglycan and/or β-lactams (Fig. 5). β3 includes the KS/TG motif (residues 618–621). The SXN motif (residues 448–450) occurs between two helices and the catalytic serine. These three motifs interact through hydrogen bonds; in the apo form, for example, Lys618 hydrogen bonds to Ser448 side chain and the main-chain oxygen of Ser392. The side chain of Ser392 hydrogen bonds to Lys395 and faces away from the other motifs. The active site is open but is more occluded than in *E. coli* PBP4, which is readily inhibited by β-lactams.[34] In the case of SA-PBP2a, which is highly resistant to β-lactams, the active site undergoes significant conformational change on binding an inhibitor, the active-site serine moving up to 1.8 Å away from strand β3.[24] SA-PBP3

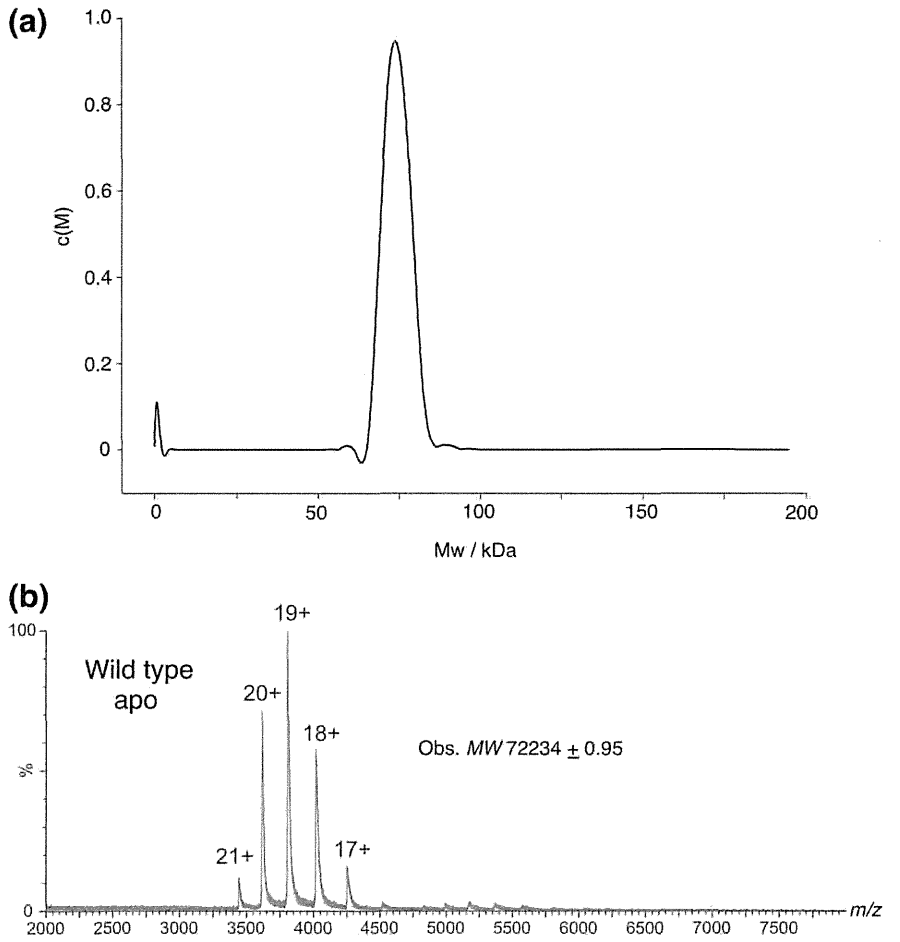

Fig. 4. The solution state of the protein. (a) The association state of the protein in solution was determined by analytical ultracentrifugation. One sharp peak was observed with a sedimentation coefficient of 4.04, consistent with a single species in solution with a molecular mass close to 75 kDa. The expected mass of the monomer is 72,234 Da. No evidence for higher-molecular-weight species was observed, indicating that the protein shows no self-association under the conditions used. (b) NanoESI mass spectrum of native PBP3. Molecular ions with overall charge of 17+ to 21+ were observed in the spectrum, and the molecular mass of PBP3 was calculated to be 72,234.34±0.95, suggesting that PBP3 exists only as a monomer under the experimental conditions in aqueous solution.

shows no such movements, consistent with its relative ease of inhibition,[35] and crystals of the apoenzyme could be soaked directly with substrate. (In the case of the class A enzyme *S. pneumoniae* PBP1b, it was necessary to use an Asn-to-Gly mutant to open up the active site and facilitate inhibitor soaking.[36]) Most of the residues in the immediate vicinity of any bound inhibitor are conserved between PBP2, SA-PBP2a, and SA-PBP3. However, even though Tyr446 of PBP2a interacts with the aromatic ring of bound methicillin,[24] the equivalent residue of PBP3, Tyr430, points away from the active site and sits between Phe341 and His447. It does not appear to play a role in enhancing substrate binding.

Another unique feature of the SA-PBP3 structure is the shortened α11 helix. In NG-PBP2 and SA-PBP2a, the equivalent helix lies near bound β-lactams, potentially sterically hindering access to the active site. Mutations in this region play an important role in antibiotic resistance in NG-PBP2, without causing any substantial change in local structure.[18] In SA-PBP3, this helix is shorter, and residues at its N-terminus form a loop (residues 655–664) that carries apolar residues including the sequence VPPPWL, which extends toward the active site and may influence the approach of hydrophobic substrates. This might suggest a preference of PBP3 for cephalexin, with a simple benzyl side group, to the more polar cefotaxime carrying a free amine on a thiazolidine ring. Cefotaxime, however, appears to be the more effective inhibitor from our results. This observation is not entirely surprising as it is found that disordered regions or effects far from the active site may have a considerable effect on the substrate-

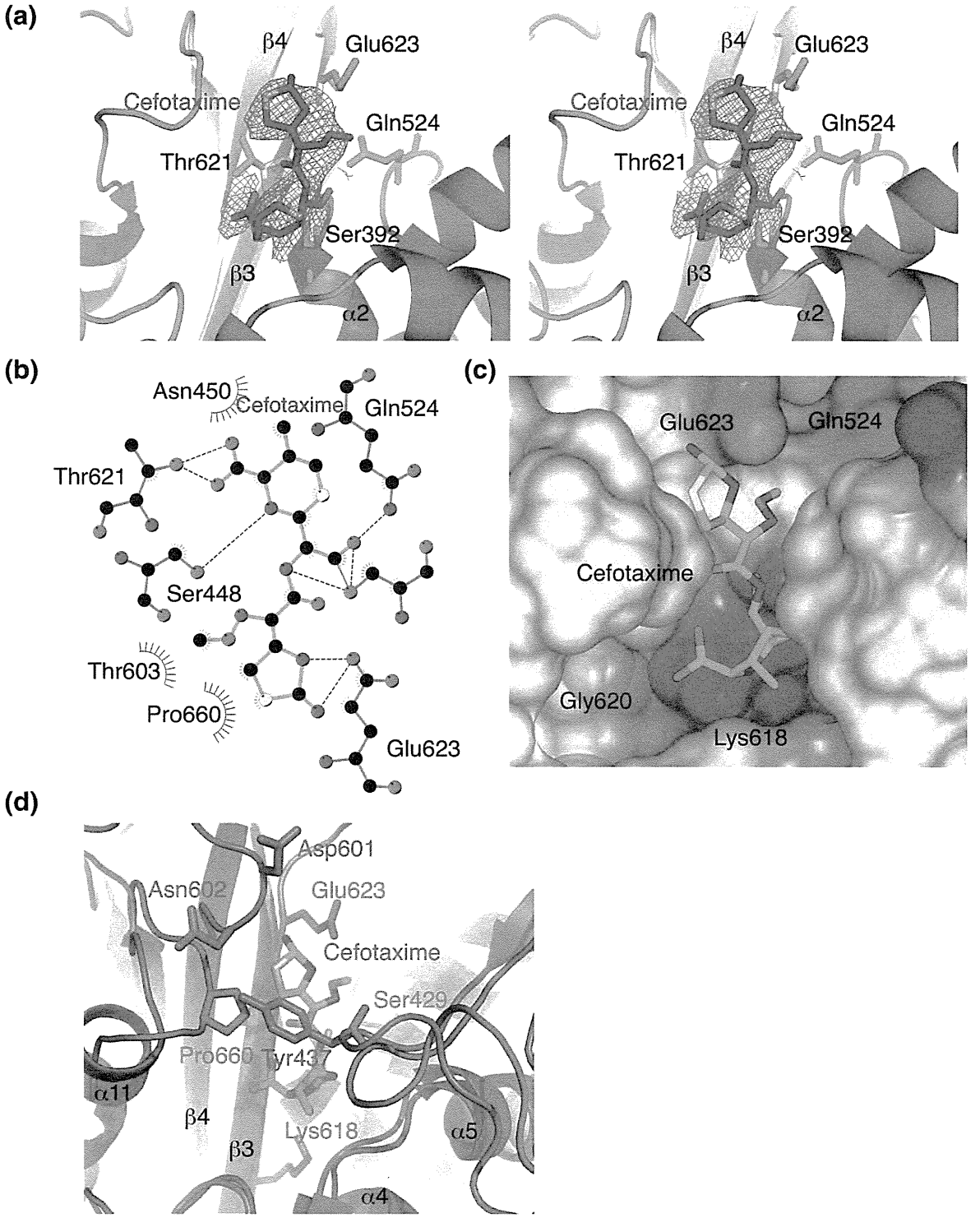

Fig. 5. The interaction of PBP3 with cefotaxime. (a) The 2mFo−DFc electron density omit map, shown in stereo and contoured at a level of 1.1 σ, covering the covalently bound residue of the drug in the active site. (b) Schematic diagram showing the protein–drug interactions. The hydrophobic interactions between cefotaxime and the protein (distance, 3.5–3.9 Å) are indicated with dotted arcs. Hydrogen bonds (2.5–3.4 Å) are shown as green dotted lines. (c) Stick representation of the drug over the molecular surface of the protein. The protein surface is colored by electrostatic potential (blue, positive), showing the charged patches near the active site. Carbon atoms of the drug are colored green; nitrogen, blue; oxygen, red; and sulfur, yellow. (d) Comparison of the active sites of the SA-PBP3 with cefotaxime complex (red) and the PBP2 of *S. aureus* (blue; PDB ID 2OLU). The models were overlapped by SSM. The penicillin-binding site is occluded by Tyr437 in the case of apo PBP2 but remains open in the apo form of SA-PBP3.

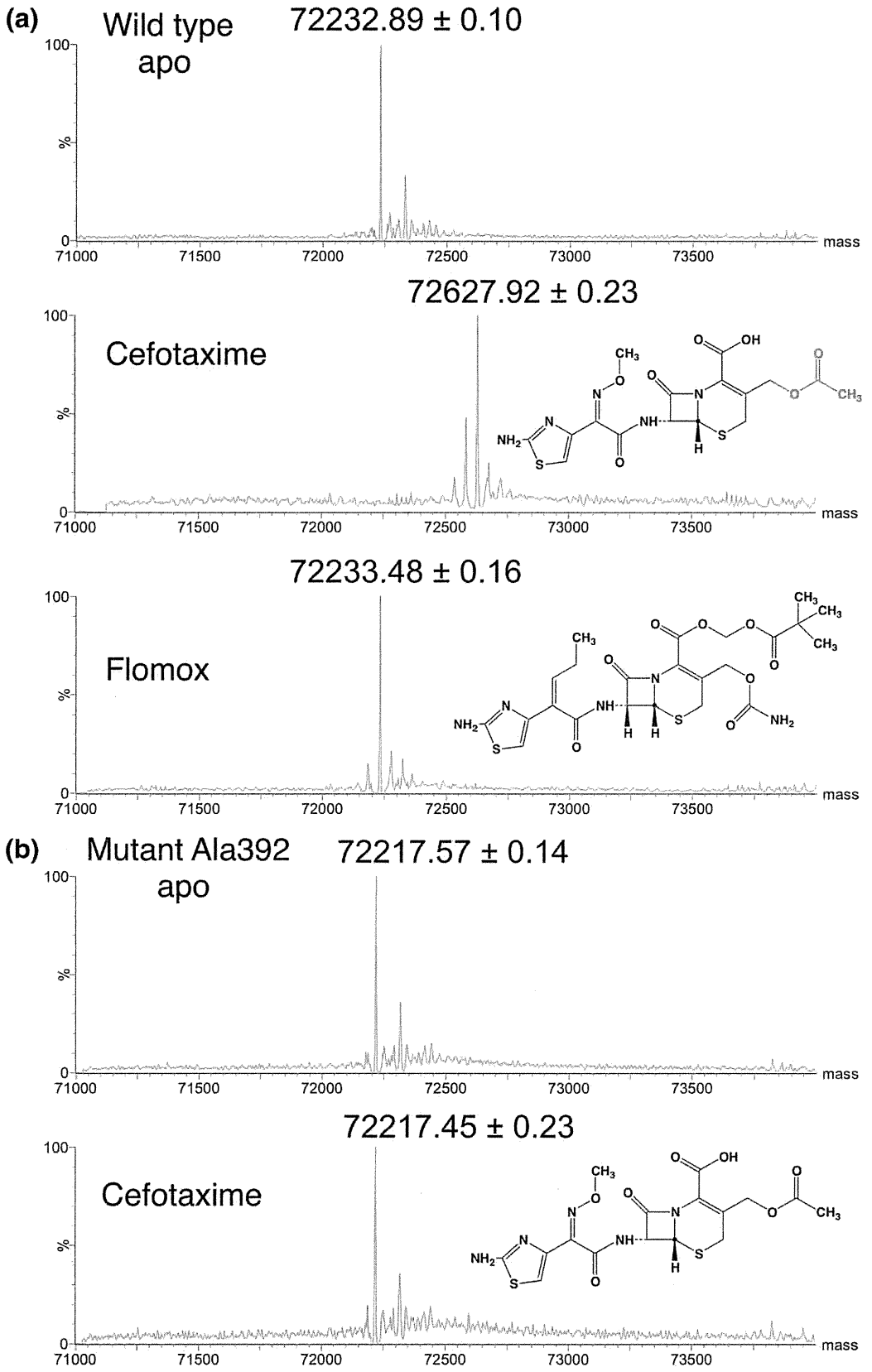

Fig. 6 (*legend on next page*)

binding properties of PBPs, making functional predictions from the structure unreliable.[37]

Cefotaxime-bound structure

Attempts were made to determine the structure of SA-PBP3 with cefotaxime, cephalexin, or ampicillin bound by soaking the apo crystals for 12 h in the presence of these compounds. Refinement proceeded readily from the apo model, but it became clear that the active site is empty in the case of crystals soaked in ampicillin or cephalexin, and therefore, these structures are no different from the apo model (data not shown). Cefotaxime remains covalently bound and clearly visible in the electron density map, although no density is observed for the acetyl group (Fig. 5). Overall, the changes to the structure on binding cefotaxime are very minor, including the conserved catalytic residues. Thr619 (of the KTG motif) and Thr621 rotate away from the drug. The C^{α} atom of Thr621 moves roughly 1 Å. Ser429 of the SXN loop turns away from the bound drug, and Glu623, shortly downstream from the KTG motif, turns slightly toward the solvent. Pro660 shifts about 1.3 Å, together with a slight twist of Trp662. None of these movements appears highly strained, and the active site of the apo form is primed for substrate binding.

Failure to observe antibiotics other than cefotaxime in the SA-PBP3 active site led us to examine the reactivity of the protein by mass spectrometry. The protein was exposed to a variety of compounds and then denatured with acetonitrile before mass measurement. A selection of the results is shown in Fig. 6. Every β-lactam tested except flomox (cefcapene pivoxil hydrochloride) showed covalent attachment to the protein. These compounds were, in ascending order of molecular mass, faropenem (285 Da), imipenem (299 Da), penicillin-G (334 Da), cephalexin (347 Da), ampicillin (349 Da), oxacillin (401 Da), cefoxitin (427 Da), and cefotaxime (455 Da). Flomox is substantially larger than the other drugs tested, with a molecular mass of 622 Da, and its size apparently prevents it from interacting productively with SA-PBP3. As a negative control, the S392A mutant was also tested, but with the active-site serine residue missing this protein showed no attachment of any β-lactam. The mass spectrometry results show that the absence of ampicillin or cephalexin from the active site of the co-crystal structure cannot be ascribed to a lack of reactivity with the protein. Instead, it appears that SA-PBP3 is able to hydrolyze away the drug residue over the 12-h period of soaking. Cefotaxime is more resilient to removal from the active site, presumably in part because of its larger size blocking the close approach of water molecules to the active-site serine residue, but the structures offer no convincing explanation for the longevity of the cefotaxime acyl complex. The mass spectrometry shows a mass increase of 395 Da after reaction with cefotaxime, rather than the expected 455 Da, corroborating the loss of an acetoxy group suggested by the electron density.

Discussion

In this paper, we have sought to understand the structure of PBP3 from *S. aureus* and the nature of its interactions with substrates. Despite decades of research, PBP substrate specificity remains a topic of debate due to the recalcitrance of these enzymes to catalyze *in vitro* the reactions that they readily perform *in vivo*.[38] Overall among PBPs, ease of inhibition shows some correlation with the openness of the active site, although it has been widely found that mutations throughout a PBP structure may influence its reactivity. Our crystal structure shows that the active site of SA-PBP3 is not sterically occluded in the apo form, and the active-site serine is in a suitable position for reaction with sufficiently small β-lactams, consistent with its relative sensitivity to such compounds.[35] Cefotaxime, developed in 1979, is a third-generation cephalosporin. Compounds in this group are generally known for their broad spectrum activity against Gram-negative pathogens rather than Gram-positive pathogens and for their improved resistance to β-lactamases than earlier generations of β-lactams due to their 7β-(aminothiazoyl)oxyiminoacetamido side chain.[39] Cephalosporins developed specifically to tackle MRSA have a long hydrophobic 3′ side chain with a basic group at the end; this fits the narrow and rather negatively charged cleft into which it must fit to bind SA-PBP2a.[24,39]

Some PBPs show subsites in the substrate pocket that are thought to be involved in substrate specificity via recognition of the stem peptide, for example, the

Fig. 6. NanoESI mass spectrum of PBP3 (wild type and mutant Ala392) in the denatured state. The molecular masses of the ligand-bound forms were obtained by deconvolution of the original nanoESI mass spectra. (a) Samples of wild-type SA-PBP3 were dialyzed with a variety of antibiotics before denaturation and analysis. The measured mass of the apoprotein showed excellent agreement with the expected value of 72,234 Da (upper panel). Increases in mass were seen with every β-lactam examined except flomox. Only the results with cefotaxime and flomox are shown, in the middle and lower panels, with the chemical structure of each compound. Cefotaxime has a molecular mass of 455 Da but gave an observed mass increase of 395 Da, suggesting loss of the acetoxy group highlighted in red. (b) Mutant Ala392. Samples of the mutant were treated identically with the wild-type protein, but no increase in mass was observed with any compound.