Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The structures of deoxy human haemoglobin and the mutant Hb Tyrα42His at 120 K

Jeremy R. H. Tame[a]* and Beatrice Vallone[b]

[a]Protonic Nanomachine Project, ERATO, 3-4 Hikaridai, Seika, Kyoto 619-0237, Japan, and [b]Dipartimento di Scienze Biochimiche, Universita di Roma 'La Sapienza', Piazzale A. Moro, 5 00185 Roma, Italy

Correspondence e-mail: jtame@npn.jst.go.jp

Received 14 February 2000
Accepted 26 April 2000

PDB References: human deoxyhaemoglobin, 1a3n; Tyrα42His mutant, 1a3o.

The structures of deoxy human haemoglobin and an artificial mutant (Tyrα42$\rightarrow$His) have been solved at 120 K. While overall agreement between these structures and others in the PDB is very good, certain side chains are found to be shifted, absent from the electron-density map or in different rotamers. Non-crystallographic symmetry (NCS) is very well obeyed in the native protein, but not around the site of the changed residue in the mutant. NCS is also not obeyed by the water molecule invariably found in the α-chain haem pocket in room-temperature crystal structures of haemoglobin. At 120 K, this water molecule disappears from one α chain in the asymmetric unit but not the other.

1. Introduction

Haemoglobin is one of the best characterized proteins, both from a structural and functional view. Despite the enormous effort devoted to studies of the molecule, there still remain unanswered questions regarding the differences between the α and β subunits and the free-energy contribution of different interactions to the oxygen affinity of the molecule and the equilibrium between allosteric states. The crystal structure of human haemoglobin was first determined by Perutz and coworkers in the late 1950s and early 1960s. In the absence of oxygen (or other haem ligands), the protein was found to adopt a structure containing several salt bridges between the $\alpha_1\beta_1$ and $\alpha_2\beta_2$ dimers and the Fe atoms were displaced from the plane of the haem in each subunit. On binding oxygen, the Fe atoms moved into the haem plane, causing the subunits to change tertiary structure slightly but sufficiently to destabilize the deoxy quaternary structure. By 1970, the structures of both the deoxy and met forms of the protein were sufficiently well refined for Perutz to put forward his famous 'trigger hypothesis' in which the tense ('T') state, preferred by the deoxy protein, switches to the relaxed ('R') state on oxygenation by sensing the position of the Fe atoms.

Several predictions of the Perutz hypothesis were testable using naturally occurring human Hb (haemoglobin) mutants such as Hb Kansas (Asnβ102$\rightarrow$Thr), which has an unstable R state and a low oxygen affinity (Fermi & Perutz, 1984). Although more than 500 different mutants have been detected clinically, the great majority of these have little functional effect since the survival of a patient is clearly dependent on the molecule working adequately well. Natural mutants afforded little opportunity for experiments on the roles of many residues close to the haem or at subunit interfaces. The development of *Escherichia coli* expression systems for the production of engineered Hb mutants was therefore an

important step in the functional characterization of these residues. A number of expression systems have now been described. The first of these required expression of the individual subunits and subsequent refolding, which is highly labour intensive (Nagai *et al.*, 1985). Later systems produced large yields of folded haemoglobin but with altered N-termini (Looker *et al.*, 1992). By overexpressing methionine aminopeptidase at the same time as the α- and β-globin genes, Ho and coworkers have now succeeded in producing enormous yields of folded functional human Hb correctly processed at the N-termini from *E. coli* (Shen *et al.*, 1993). Detailed structural and functional analysis of Hb mutants is of great interest in the field of synthetic blood substitutes known as HBOCs (haemoglobin-based oxygen carriers; Bunn, 1993; Komiyama *et al.*, 1996). At present, however, the best refined structure of native deoxy human Hb is that of Fermi *et al.* (1984). Although the data used were of high resolution (1.74 Å), protein structure refinement methods were primitive compared with the graphical and computational programs now available. We have re-solved the structure of deoxy Hb using a cryo-cooled crystal and an image-plate X-ray detector, refining the structure using a maximum-likelihood method instead of real-space refinement. We have also solved the crystal structure of an engineered Hb mutant, Tyrα42$\rightarrow$His, treated in the same way. The differences between our structures and previous structures of Hb are discussed.

2. Methods

2.1. Crystallization

Human Hb was prepared from blood freshly drawn from one of the authors (JT).

Protein purification and crystallization were carried out according to the procedure of Perutz (1968), except that crystallization was carried out in 100 µl tubes and dithionite instead of ferrous citrate was used to remove oxygen. Conversion of the protein to the deoxy form was immediately visible as a colour change from bright red to purple. The oxygen affinity and reaction rate of the protein preclude the possibility of collecting X-ray data on deoxy Hb crystals under air at room temperature. All previous deoxy structures have therefore been mounted in capillaries, usually in a nitrogen glove-box. If sufficient dithionite is present in the mother liquor, however, it is possible to seal a deoxy Hb crystal in a capillary while working in an aerobic environment (Wireko & Abrahams, 1992). Capillary mounting is incompatible with cryocooling, so we took the logical step of adding dithionite to an appropriate mother liquor containing 25% glycerol in which deoxy crystals are stable for a few minutes and which can be cooled rapidly to 120 K without ice formation. A few grains of fresh sodium dithionite were added to 1 ml of this mother liquor in a glass dimple plate just prior to opening the tube of crystals. Immediately after opening the tube under air it was filled with a small volume of this mother liquor to prevent free oxygen reaching the crystals. It was then inverted over the remainder of the mother liquor to allow a few size-able crystals out. The crystal colour was examined under the microscope to check for signs of oxygenation or oxidation, but none were visible. After a brief (roughly 30 s) washing, a crystal was removed in a fibre loop and placed on the goniometer head in a cold stream of nitrogen at 120 K.

Table 1
Data-collection and refinement statistics.

Data set	DX	Y42αH
Resolution limits (Å)	15–1.8	15–1.8
$R_{\rm merge}$ (%, overall)	5.3	7.4
$R_{\rm merge}$ (%, 1.86–1.80 Å)	16.2	33.2
No. of observations	49661	49373
Completeness (%, overall)	95.8	96.0
Completeness (%, 1.86–1.80 Å)	76.1	76.4
Average multiplicity	2.4	3.2
Average $I/\sigma(I)$ (overall)	15.0	18.0
Average $I/\sigma(I)$ (1.86–1.80 Å)	4.0	2.9
No. of protein atoms	4542	4542
No. of water molecules	451	412
$R_{\rm cryst}$	0.171	0.180
Free $R_{\rm cryst}$	0.220	0.227
R.m.s. bond length (Å) (σ = 0.02)	0.012	0.013
R.m.s. bond angle (Å) (σ = 0.04)	0.028	0.031
R.m.s. planes (Å) (σ = 0.05)	0.034	0.036
Average temperature factors (Å^2)		
Main chain	14.0	19.1
Side chain	17.2	22.8
Solvent	24.9	31.3

† $R_{\rm merge} = \sum(|I_i - I_n|)/\sum I_n$, where I_i is an observed intensity and I_n is the average observed intensity for a given reflection. ‡ $R_{\rm cryst} = \sum_{hkl} ||F_o| - |F_c||/|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor amplitudes of a reflection hkl. 5% of reflections were excluded from refinement for the calculation of the free $R_{\rm cryst}$. No intensity cutoff was applied.

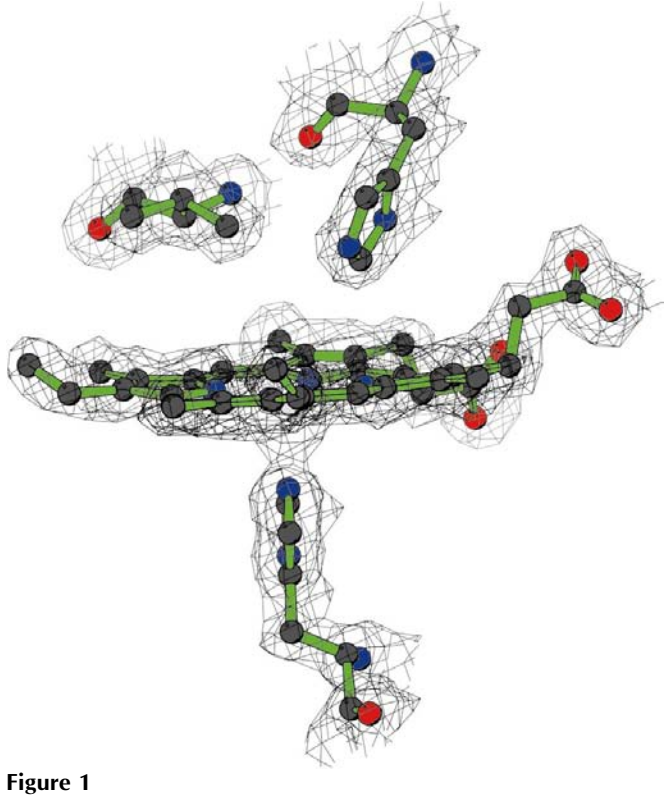

Figure 1
The α_1 haem. The $2F_o - F_c$ electron-density map is contoured at 1σ.

Table 2
R.m.s. deviations (Å) between DX and different deoxy-Hb models.

File	C^{α}	Main chain	All polypeptide atoms
Yα42H	0.273	0.273	0.557
2hhb	0.403	0.401	0.774
1hga	0.462	0.462	0.801
1hbb	0.444	0.438	0.731
1dxu	0.390	0.386	0.657

The mutant protein was produced using the α-globin expression system described by Tame *et al.* (1991). α-Globin was expressed alone in *E. coli* and then refolded with native human β-globin to form tetrameric Hb. The mutagenesis was carried out by Professor Kiyohiro Imai of Osaka University. Once synthesized and purified according to the published methods, the mutant protein was then handled in an identical fashion to the wild-type protein.

2.2. Data collection and refinement

X-ray data were collected using a Rigaku R-AXIS II image-plate system mounted on a copper rotating-anode generator. X-rays were focused by a mirror system (MSC), $K\beta$ radiation being removed by a nickel filter. 1° oscillation images were collected and the data were processed using *DENZO* (Otwinowski, 1990). All subsequent data handling was carried out using the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). Refinement was carried out using the structure of an artificial mutant Hb Valβ1→Met (PDB code 1dxu; Kavanaugh, Rogers & Arnone, 1992) as a starting model. Although the space group is the same and the unit-cell parameters are very similar, the *R* factor calculated with the data from frozen crystals and this model was very high. The model was therefore repositioned by molecular replacement using *AMoRe* (Navaza & Saludjian, 1997) at the start of refinement. 5% of reflections were randomly chosen to be excluded from refinement for the calculation of a free *R* factor (Brünger, 1992). After *AMoRe*, the initial *R* factor and free *R* factor were both 35%. Refinement was carried out using *REFMAC* (Murshudov *et al.*, 1997). Manual adjustment of the model was carried out using *QUANTA* (Molecular Biosym Inc.). Low-resolution terms to 15 Å were included in the refinement and bulk scaling was used to help model the solvent. No σ cutoff was applied at any stage of data handling or refinement. After the initial stages of refinement, individual *B* factors were refined for each atom and the water molecules were ordered by *B* factor once the structures were completed. Data-collection and refinement statistics are given in Table 1.

3. Results and discussion

3.1. Overall structure

Not surprisingly, the high-salt structure of deoxy Hb at 120 K shows few significant deviations from the crystal structures determined at room temperature and low-salt conditions. We have compared our structures with the previous refinement of high-salt deoxy Hb (PDB code 2hhb; Fermi *et al.*, 1984), the high-salt structure of Hb Valβ1→Met (1dxu; Kavanaugh, Rogers, Case *et al.*, 1992) and low-salt crystals grown from PEG, 1hbb (Kavanaugh, Rogers, Case *et al.*, 1992) and 1hga (Liddington *et al.*, 1992; Paoli *et al.*, 1996), all of which were solved at or near room temperature. We refer to our native structure as 'DX' and the mutant structure as 'Yα42H' throughout. R.m.s. deviations between these structures and the previous entries in the PDB are shown in Table 2. It can be seen that DX and Yα42H generally agree with each other rather better than with the other structures, despite the effects of the mutation (which are described below), showing that they reflect the treatment and handling of the crystals and X-ray data. Agreement between the structures is high, as would be expected given the resolution of the data sets (1.9 Å at worst) and in line with expected coordinate errors.

No electron density was observed for the N-terminal residue Valβ1 or the side chain of His2 in the maps calculated for DX and Yα42H, so these residues were omitted from the models. Since these residues have high temperature (*B*) factors in the other structures and there is also little agreement between them in position, it seems probable that they are disordered. The inclusion of these residues in other structures may reflect model bias, which is largely eliminated by the maximum-likelihood method.

3.1.1. Haem pockets. The haem pockets form the functional core of the protein, binding oxygen reversibly at the sixth coordination site of the haem Fe atoms. Communication

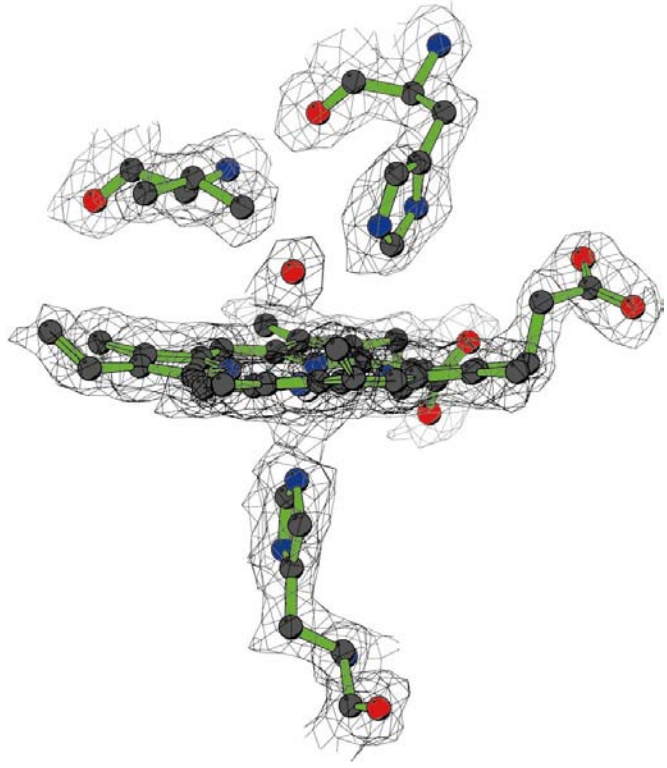

Figure 2
The α_2 haem. The $2F_o - F_c$ electron-density map is contoured at 1σ, clearly showing the haem-pocket water molecule found in room-temperature structures of deoxyhaemoglobin.

Table 3
The distances between the Fe atoms and the plane of porphyrin N atoms (Å) in different deoxy-Hb models.

File	α_1	α_2	β_1	β_2
DX	2.60	2.73	2.60	2.56
Yα42H	2.59	2.67	2.54	2.58
2hhb	2.59	2.49	2.47	2.39
1hga	2.50	2.57	2.51	2.65
1hbb	2.60	2.73	2.56	2.59
1dxu	2.70	2.63	2.52	2.52

between the haems is governed by the subunit interfaces. The connection between the geometry of these interfaces and the haems has been well documented (Perutz *et al.*, 1987), but further twists continue to emerge from X-ray structures of intermediate states which are unstable in solution (Paoli *et al.*, 1996; Wilson *et al.*, 1996). Figs. 1 and 2 show the $2F_o - F_c$ electron-density map over the α_1 and α_2 haems. The geometry of the residues around the haems is very similar in all the structures examined. The displacements of the Fe atom from the haem plane are given in Table 3.

The most significant feature of the haem pockets is the absence of any density for the water molecule in the α_1 subunit in both the DX and Yα42H structures. This water molecule, which hydrogen bonds to the distal histidine, is found in both α-chain haem pockets of all the room-temperature deoxy-Hb structures discussed in this paper. A water molecule found in the same position of sperm whale myoglobin appears to act as a kinetic barrier to ligand binding (Carver *et al.*, 1990). In contrast, there is insufficient space for a water molecule in the β subunits. Figs. 1 and 2 show clearly that there is no electron density for this water in the α_1 haem pocket, but very strong density appears in α_2. This difference is quite remarkable given that the protein atoms for the two α subunits can be overlapped very closely in this region. A similar situation is found in the structure Yα42H, but in this case the distal histidine side chain has been flipped 180° around χ_2 since a partially occupied water molecule has appeared above it in the haem pocket. Rotating the imidazole group brings this water within hydrogen-bonding distance of $N^{\varepsilon 2}$. The only structural explanation for the disappearance of the water molecule from the haem pocket of the α_1 subunit is a slight shift in the position of the haem relative to the distal residues. If the $\alpha_2\beta_2$ dimer is superimposed on the $\alpha_1\beta_1$ dimer by least-squares fitting the C^{α} positions, then it can be seen that the rotated and translated α_2 haem lies approximately 0.1 Å below the α_1 haem. It appears then that the haem pocket is compressed slightly by the treatment of the crystal: either the addition of glycerol, cryocooling or a combination of the two. This compression appears sufficient to expel the water molecule completely, despite only being on a scale comparable to the error in the coordinates of an atom in the structure.

Among the structures examined, Yα42H is unique in the position of the propionate groups of the β_2 haem. Although these groups point into solvent and seem to play little role in oxygen binding, their position has been found to be indicative of the oxidation state of the haem iron in the case of lupin leghaemoglobin. Harutyunyan and coworkers found that changing the haem from deoxy to met (but not to oxy) in this structure caused the propionate group attached to pyrrole A of the haem to move from a proximal position to the distal side of the haem (Harutyunyan *et al.*, 1995). However, the electron density of the Yα42H Hb mutant shows no evidence of oxidation, no ligand density being visible at very low contour levels and the position of the Fe atom relative to the haem plane being consistent with a completely deoxy haem.

3.2. Tyrα42→His

In native deoxy human Hb, it was found by Perutz (1970) that Tyrα42 forms a hydrogen bond with Asp$\beta_2$99 which is broken as the two $\alpha\beta$ dimers move relative to each other to reach the R state. This hydrogen bond gives rise to a resonance in the proton NMR spectrum of Hb which is taken as a signature of the T state (Fung & Ho, 1975). Although this resonance (at +9.4 p.p.m. relative to water) is not observed in

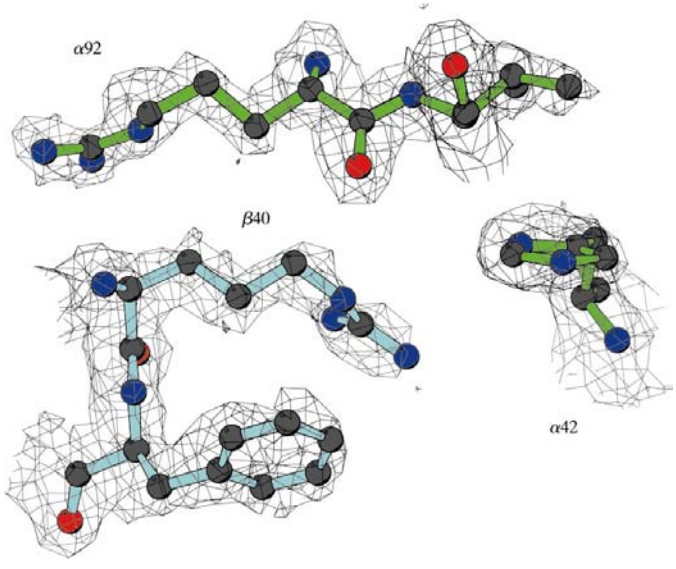

Figure 3
His$\alpha_1$42. The $2F_o - F_c$ electron-density map shows the residues neighbouring $\alpha_1$42 are undisturbed by the mutation from tyrosine.

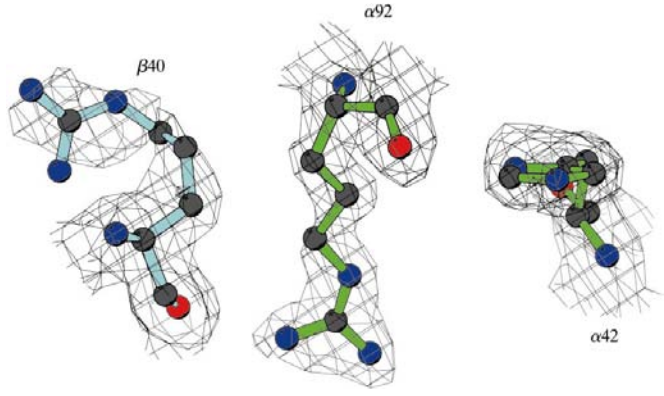

Figure 4
His$\alpha_2$42. In contrast to the NCS-related dimer within the same molecule, the $\alpha_2\beta_1$ contact shows large changes from the wild-type structure.

D_2O, indicating that the proton is exchangeable, the rate of exchange of the tyrosine proton and water protons appears to be slow on the NMR timescale (Russu *et al.*, 1987). This may be a consequence of the strength of the Tyr–Asp hydrogen bond, which is suggested by the relatively short distance between the donor and acceptor atoms, ~2.55 Å, and the fact that both side chains participate in a wider hydrogen-bonding network which will tend to stabilize the interaction. The equivalent residue in the β subunits, β41, is generally Phe. In about 10% of vertebrate species a Tyr is found at this position and has been implicated in lowered oxygen affinity (Tame *et al.*, 1996) and altered heterotropic ligand binding (Komiyama *et al.*, 1995). Mutating α42 would therefore be expected to weaken the T state and raise oxygen affinity. Functional analysis of two mutants, Tyrα42→His and Tyrα42→Phe, showed that this is the case (Imai *et al.*, 1991). Kim and coworkers have also demonstrated the importance of the α_1–β_2 hydrogen bond by making two partly compensatory mutations at these positions (Tyrα42→Asp, Aspβ99→Asn; Kim *et al.*, 1994). We were unable to obtain deoxy crystals of Hb(Tyrα42→Phe), probably as a result of the strong destabilization of the T state. The Phe side chain is unable to hydrogen bond to Aspβ99 at all, but a weak hydrogen bond can form with His at α42. The oxygen affinity of this mutant is not increased as much in the His mutant as the Phe mutant. Structural characterization of the Tyrα42→His mutant was carried out to determine whether or not the observed increase in oxygen affinity was simply a consequence of the loss of a T-state hydrogen bond. In the case of the $\alpha_1\beta_2$ contact this would appear to be the case, but there are marked differences around $\alpha_2$42 between the two copies of the α subunit found in the asymmetric unit. This region developed poor geometry and electron density early in refinement when NCS restraints were applied. On removing these it became clear that significant changes had occurred and the new positions of side chains were determined by omit maps. This region is compared in Figs. 3 and 4, which show clearly the structural rearrangement which has taken place in the $\alpha_2\beta_1$ contact region. The side chains of Arg$\alpha_2$92 and Arg$\beta_1$40 have swapped places, apparently as a result of electrostatic repulsion between His$\alpha_2$42 and Arg$\beta_1$40. This Arg residue swings its side chain towards Glu$\beta_1$43, which normally forms a salt bridge with Arg$\alpha_2$92. This residue takes up a position close to that usually occupied by $\beta_1$40, but with its guanidinium group further from $\alpha_2$42 and within 3.0 Å of the carbonyl O atom of $\beta_1$40. This rearrangement leads to a peptide flip of Arg$\alpha_2$92, moving it from the left-handed helical region of the Ramachandran plot to a position between the right-handed helical and β-sheet regions, a quite unusual departure from non-crystallographic symmetry. The *B* factors for the displaced Args are much higher, 20–45 Å^2, than those in the expected positions, 7–15 Å^2. Arg$\beta_2$40 is barely moved by the mutation, though its side-chain *B* factors are approximately double those in the wild-type structure (DX). Arg$\beta_1$40 shows more movement of its guanidinium group and its *B* factors are roughly tripled by the introduction of histidine at position $\alpha_2$42. The structure of Yα42H shows that the mutation raises oxygen affinity not only by weakening the α42–β99 bond but also by weakening other $\alpha_2\beta_1$ interactions. The disruption of NCS in the molecular tetramer indicates that the altered conformation is close in energy to the native conformation. This destabilization of the protein is accompanied by very minor changes in the structure of the $\alpha_1\beta_2$ contact, however, showing that protein structure does not always reflect stability. A related effect was observed among a group of Hb mutants with N-terminal mutations, the one with the structure closest to that of HbA having the most strongly altered oxygen affinity (Kavanaugh, Rogers & Arnone, 1992). In the case of the α42 mutant it seems that the destabilization of the T state not only arises because of the loss of a hydrogen bond, but also because of the greater flexibility around the altered residue.

Table 4
Distances between the side-chain O atoms of Aspβ_2 and hydrogen-bond donor atoms in T-state Hb.

Atom	Distance (Å)
Tyr$\alpha_1$42 O$^{\eta}$	2.51
Asn$\alpha_1$97 N$^{\delta}$	2.92
Glu$\beta_2$101 N	3.07
O (water)	2.67

Tyrα42 is of some interest not only because of its functional role but also because of the suggestion made on the basis of ultraviolet resonance Raman (UVRR) spectroscopy that it *accepts* a hydrogen bond from Asp$\beta_2$99 in the T state rather than donating one. This hypothesis was put forward on the basis of a slight upfield shift of a band (Y8a) attributable to the tyrosines in the structure (Rodgers *et al.*, 1992). Five of the six tyrosines found in the $\alpha\beta$ dimer have almost identical environments in both the R and T structures of Hb, the only exception being Tyrα42. *p*-Cresol was used as a model compound to test the shift of its UVRR band equivalent to the Y8a band of Hb. This shift was interpreted as indicative of hydrogen-bond donation from the aspartate to the tyrosine in the T state. It was claimed this is consistent with the crystallographic structure, which shows the Asp99 side chain to be 'in a solvent-inaccessible hydrophobic environment'. In fact, all of the T-state structures discussed in this paper include a surface water molecule hydrogen bonded to the Aspβ99 carboxyl group. Even if the side chain were buried, it is unlikely that its pK_a would be raised higher than that of the tyrosine hydroxyl group. Elevated Asp and Glu side-chain pK_as are observed in many enzymes where these groups are directly involved in catalysis, but even the close approximation of two carboxyl groups only raises the pK_a to about 7 (McIntosh *et al.*, 1996). The protonation of Asp$\beta_2$99 in T-state Hb, which is implied by its donation of a hydrogen bond, is made even less likely by the hydrogen bonds it forms with the main-chain N atom of Glu$\beta_2$101 and the side-chain N atom of Asnα97. The side chain of Asp$\beta_2$99 forms four strong hydrogen bonds in all (Table 4). Support for the Rodgers hypothesis has come from the group of Friedman, also using UVRR but in this case with a double mutant Hb, Tyrα42→Asp, Aspβ99→Asn (Huang *et al.*, 1997). This showed no change in the position of the Y8a band on ligation

Table 5
The water molecules in different deoxy-Hb models.

File	No. of waters present	Average B factor (Å^2)	$\sigma(B)$ (Å^2)
DX	440	25.1	7.2
Yα42H	404	31.5	8.6
2hhb	221	35.5	5.3
1hga	89	26.8	6.4
1hbb	235	32.1	9.1
1dxu	192	40.7	13.9

Table 6
Average distances between water molecules in DX and their equivalents in different deoxy-Hb models.

File	No. of waters used	Average distance (Å)
Yα42H	366	0.685
2hhb	184	0.891
1hga	66	0.909
1hbb	203	0.746
1dxu	158	0.900

of the Hb, indicating that the shift seen by Rodgers *et al.* was indeed because of Tyrα42. The Y8a and Y9a bands were also assigned directly by Nagai *et al.* (1996) using the Y42α mutant. However, the ability of histidine to substitute for tyrosine and retain a functional cooperative Hb shows that there is no need to invoke such an unusual hydrogen bond to make the protein work (Imai *et al.*, 1991). We therefore suggest that the proposal of this inverted hydrogen bond is wrong.

3.3. Water

The number of water molecules and their average B factors in the various structures are given in Table 5. It can be seen that the B factors are similar for the waters included in the different models, indicating that this reflects static disorder as much as thermal vibration. Quite stringent criteria have been applied in refining the model 1hga and only 89 waters are included, a relatively small number given the number of non-H protein atoms (4556). In order to examine the consistency of water positions among the various structures, the shortest distances from fully occupied waters in the different models to the equivalent water in DX (where one existed) were calculated. The structures were overlaid by least-squares minimization of the distances between the protein main-chain atoms before determining the inter-water distances. Water molecules were excluded if the nearest neighbour in the DX structure was found to be in a different asymmetric unit. An arbitrary cutoff of 3.5 Å was also applied to ensure that waters were in roughly equivalent positions, otherwise the comparison would be meaningless. The average distance of water molecules to their equivalents in the DX structure are shown in Table 6. The average displacement is lowest for Yα42H, despite the movement of water molecules near the site of the mutation, so the positions of the waters do seem to reflect the handling of the crystals and the data, which was identical in these two cases. The average distances are broadly in line with the error expected from the least-squares overlap and the errors in the coordinate positions.

The systematic differences between the water structure of deoxyhaemoglobin at room temperature and at 120 K clearly has implications for the inference of the biological role of water molecules in cryocooled crystal structures. Water molecules have a profound influence on the behaviour of proteins and are found to be buried in many protein–ligand complexes, where they have been shown to play an important role in controlling ligand affinity and specificity. The visualization of water molecules associated with protein structures by NMR is only possible for the small minority of waters which remain bound to the same protein molecule for relatively long periods. In contrast, X-ray crystallography readily reveals the pattern of hydration about biomacromolecules provided that the crystals diffract adequately. Cryocooling crystals has become a very popular method for extending the lifetime of protein crystals in an X-ray experiment and has the additional advantage of revealing many more water molecules in protein structures. The disappearance of the functionally important haem-pocket water molecule of the α_1 subunit at very low temperature, but not from the symmetry-related α_2 subunit, further illustrates the difficulties in correlating water structure to function.

3.4. Rotamers

Rotamer preferences of different residues are quite marked and distinct (Ponder & Richards, 1987). We have identified a number of residues which have unusual side-chain angles in 2hhb, 1hga, 1hbb and 1dxu, particularly leucine side chains. Even with data at better than 2 Å resolution it can be difficult to distinguish between a generally preferred rotamer of leucine and one in which χ_2 is rotated by 180°. These latter conformations are uncommon and are usually errors. 20 leucine residues were flipped in this way towards the end of refinement when the electron-density maps were clearest. One leucine residue (β32) does appear to be in an unusual conformation. Our models differ from 2hhb at three leucine side chains in each $\alpha\beta$ dimer. Although these changes may have little functional significance, they illustrate the point that comparison between models, by which the allosteric transition is mapped, is made more difficult when studying models refined under different conditions by different programs.

In the case of asparagine, glutamine and histidine residues, the last side-chain χ angle may easily be 180° in error since the shape of the group is preserved and N atoms are swapped with O or C atoms. This appears to have occurred for Hisα103, which has caused particular problems in NMR studies. This is a highly conserved residue at the $\alpha_1\beta_1$ contact, its side chain fitting into a shallow pocket formed by the β subunit. On the basis of the high-resolution deoxy structure (2hhb), a long hydrogen bond (~3.5 Å) was inferred between N$^{\varepsilon 2}$ of Hisα103 and the carbonyl O atom of Asnβ108 (Fermi *et al.*, 1984). This region of the molecule is of interest because a naturally occurring mutation (Asnβ108→Lys, found in Hb Presbyterian) has potential as a means of lowering the oxygen affinity in commercial HBOCs. The Hisα103 side chain is in the same position in 1hga, but is flipped in all the others. This brings the N$^{\varepsilon 2}$ atom within 2.8 Å of O$^{\varepsilon 1}$ of Glnβ131. In the structures

2hhb and 1hga there appears to be a steric clash between this O atom and Hisα103 $C^{\varepsilon 1}$, which suggests strongly that the histidine residue has been positioned incorrectly. This type of error is found fairly frequently in older PDB entries deposited before simple validation tools were available. Older entries are also found to contain a number of nomenclature errors which can cause problems when least-squares fitting different structures. For example, there are 45 nomenclature errors in 2hhb; the r.m.s. deviation of all polypeptide atoms between DX and 2hhb is 0.77 Å if these are ignored but 0.67 Å if they are corrected prior to the calculation.

The overall agreement between the two models described here shows that they are globally more closely related to each other than structures of Hb determined at room temperature, whether in the same space group or not, despite the local rearrangement caused by the mutation of Tyrα42. It has been noted before that different refinement programs leave their imprint on protein coordinate sets and it is generally possible to discover which one was used from characteristic features in the geometry of the model (Laskowski *et al.*, 1993). This makes the comparison of structures from the PDB rather more difficult. The error margins in crystallographic models are not usually a barrier to structural insights into molecular mechanism, but they may be significant in attempts to calculate the free energy of particular interactions, for example between a protein and its ligand. However, higher resolution structures are still unlikely to overcome the problems of free-energy estimation, so it appears there is some way to go before haemoglobin is truly understood. This is clearly illustrated by animal haemoglobins such as the fish haemoglobins whose extreme pH sensitivity remains unexplained after several decades of study.

We would like to thank Drs Kiyoshi Nagai and Julie Wilson, and Professors Kiyohiro Imai, Guy Dodson and Maurizio Brunori for commenting on the manuscript. JRHT is a Royal Society University Research Fellow.

References

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.

Bunn, H. F. (1993). *Am. J. Hematol.* **42**, 112–117.

Carver, T. E., Rohlfs, R. S., Olson, J. S., Gibson, Q. H., Blakmore, R. S., Springer, B. A. & Sligar, S. G. (1990). *J. Biol. Chem.* **265**, 20007–20020.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Fermi, G. & Perutz, M. F. (1984). *Haemoglobin and Myoglobin.* Oxford University Press.

Fermi, G., Perutz, M. F., Shaanan, B. & Fourme, R. (1984). *J. Mol. Biol.* **175**, 159–174.

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

Harutyunyan, E. H., Safonova, T. N., Kuranova, I. P., Popov, A. N., Teplyakov, A. V., Obmolova, G. V., Rusakov, A. A., Vainshtein, B. K., Dodson, G. G., Wilson, J. C. & Perutz, M. F. (1995). *J. Mol. Biol.* **251**, 104–115.

Huang, S., Peterson, E. S., Ho, C. & Friedman, J. M. (1997). *Biochemistry*, **36**, 6197–6206.

Imai, K., Fushitani, K., Miyazaki, G., Ishimori, K., Kitagawa, T., Wada, Y., Morimoto, H., Morishima, I., Shih, D. & Tame, J. (1991). *J. Mol. Biol.* **218**, 769–778.

Kavanaugh, J. S., Rogers, P. H. & Arnone, A. (1992). *Biochemistry*, **31**, 8640–8647.

Kavanaugh, J. S., Rogers, P. H., Case, D. A. & Arnone, A. (1992). *Biochemistry*, **31**, 4111–4121.

Kim, H. W., Shen, T. J., Sun, D. P., No, N. T., Madrid, M., Tam, M. F., Zou, M. F., Cottam, P. F. & Ho, C. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 11547–11551.

Komiyama, N. H., Miyazaki, G., Tame, J. & Nagai, K. (1995). *Nature (London)*, **373**, 244–246.

Komiyama, N., Tame, J. & Nagai, K. (1996). *Biol. Chem.* **377**, 543–548.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 1049–1067.

Liddington, R. C., Derewenda, Z., Dodson, E., Hubbard, R. & Dodson, G. (1992). *J. Mol. Biol.* **228**, 551–579.

Looker, D., Abbott-Brown, D., Cozart, P., Durfee, S., Hoffman, S., Mathews, A. J., Millerroehrich, J., Shoemaker, S., Trimble, S., Fermi, G., Komiyama, N. H., Nagai, K. & Stetler, G. L. (1992). *Nature (London)*, **356**, 258–260.

McIntosh, L. P, Hand, G., Johnson, P. E., Joshi, M. D., Korner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W. & Withers, S. G. (1996). *Biochemistry*, **35**, 9958–9966.

Murshudov, G., Vagin, A. & Dodson, E. (1997). *Acta Cryst.* D**53**, 240–255.

Nagai, M., Imai, K., Kaminaka, S., Mizutani, Y. & Kitagawa, T. (1996). *J. Mol. Struct.* **379**, 65–75.

Nagai, K., Poyart, C. & Perutz, M. F. (1985). *Proc. Natl Acad. Sci.* **82**, 7252–7255.

Navaza, J. & Saludjian, P. (1997). *Methods Enzymol.* **276**, 581–594.

Otwinowski, Z. (1990). *DENZO Data Processing Package.* Yale University, New Haven, CT, USA.

Paoli, M., Liddington, R., Tame, J., Wilkinson, A. & Dodson, G. (1996). *J. Mol. Biol.* **256**, 775–792.

Perutz, M. F. (1968). *J. Cryst. Growth*, **2**, 54–56.

Perutz, M. F. (1970). *Nature (London)*, **228**, 226–233.

Perutz, M. F., Germi, G., Luisi, B., Shaanan, B. & Liddington, R. C. (1987). *Acc. Chem. Res.* **20**, 309–321.

Ponder, J. W. & Richards, F. M. (1987). *J. Mol. Biol.* **193**, 775–791.

Rodgers, K. R., Su, C., Subramanian, S. & Spiro, T. G. (1992). *J. Am. Chem. Soc.* **114**, 3697–3709.

Russu, I. M., Ho, N. & Ho, C. (1987). *Biochem. Biophys. Acta*, **914**, 40–48.

Shen, T.-J., Ho, N. T., Simplaceau, V., Zou, M., Green, B. N., Tam, M. F. & Ho, C. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 8108–8112.

Tame, J., Shih, D. T., Pagnier, J., Fermi, G. & Nagai, K. (1991). *J. Mol. Biol.* **218**, 761–767.

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

Wilson, J., Phillips, K. & Luisi, B. (1996). *J. Mol. Biol.* **264**, 743–756.

Wireko, F. C. & Abrahams, D. J. (1992). *Protein Eng.* **5**, 3–5.