J. Mol. Biol. (1990) 213, 215–218

X-ray Crystal Structure of a Recombinant Human Myoglobin Mutant at 2·8 Å Resolution

Stevan R. Hubbard, Wayne A. Hendrickson

Howard Hughes Medical Institute
Department of Biochemistry and Molecular Biophysics
Columbia University
New York, NY 10032, U.S.A.

David G. Lambright and Steven G. Boxer

Department of Chemistry, Stanford University
Stanford, CA 94305, U.S.A.

(Received 23 September 1989; accepted 28 December 1989)

We have grown crystals in trigonal space group $P3_221$ of a mutant human myoglobin, aquomet form, in which lysine at position 45 has been replaced by arginine and cysteine at position 110 has been replaced by alanine. Suitable crystals of native recombinant human myoglobin have not been obtained. We have used the molecular replacement method to determine the X-ray crystal structure of the mutant at 2·8 Å resolution. At the present stage of refinement, the crystallographic R-value for the model, with tightly restrained stereochemistry, is 0·158 for 5·0 to 2·8 Å data. As expected, the overall structure is quite similar to the sperm whale myoglobin structure. Arginine 45 adopts a well-ordered conformation similar to that found in aquomet sperm whale myoglobin.

We began a crystallographic study of recombinant human myoglobin (Mb†) to provide a high-resolution structural basis for site-directed mutagenesis studies. The aim of such studies is to probe electrostatic interactions (Varadarajan *et al.*, 1989*a*) and ligand and proton exchange dynamics. The migration of ligands to the buried heme binding site is of particular interest. Crystallographic evidence suggests that Arg45 (CD3) in sperm whale Mb may be involved in ligand access (Kuriyan *et al.*, 1986). In human Mb, lysine is found at position 45. Whereas one would anticipate that the mutation Lys45 to Arg (K45R) would result in ligand binding properties more similar to those of sperm whale Mb, kinetics studies of carbon monoxide binding have shown that $K_{app}^{K45R} > K_{app}^{native\ human} > K_{app}^{sperm\ whale}$ (Lambright *et al.*, 1989).

We have not been able to grow suitable crystals of native recombinant human Mb, but we have succeeded in growing diffraction-quality crystals of a mutant, abbreviated K45R′, containing the replacements K45R and C110A. The C110A replacement was introduced because of the tendency for the cysteine to oxidize during recovery of the protein. In all studies to date, the C110A single mutant has behaved in the same way as native recombinant human Mb (Varadarajan *et al.*, 1989*b*). Using the molecular replacement method, we have solved the structure of the mutant (aquomet form) at 2·8 Å resolution (1 Å = 0·1 nm). A 1·6 Å resolution dataset has been collected on film at the National Synchrotron Light Source (Brookhaven) and will be used for high-resolution refinement.

The preparation of the site-specific mutant K45R′ has been described (Lambright *et al.*, 1989). Crystals of K45R′ were grown at 21 °C using the hanging drop method. In a typical crystallization procedure, 50 to 85 μl of solution was prepared consisting of

† Abbreviations used: Mb, myoglobin; K45R, arginine for lysine replacement at position 45; C110A, alanine for cysteine replacement at position 110; K45R′, recombinant human myoglobin with an arginine for lysine replacement at position 45 and an alanine for cysteine replacement at position 110; K_{app}, apparent binding constant; 2-D, 2-dimensional; 3-D, 3-dimensional; B-value, isotropic temperature parameter; R-value, standard crystallographic agreement factor; r.m.s., root-mean-square; c.p.u., central processor unit.

0022-2836/90/100215-04 $03.00/0

approximately 5 mg K45R′ protein/ml, 250 mM-Tris·HCl (pH 7·5), 75% saturated ammonium sulfate. A series of 5-μl drops of this solution were placed over 0·5 ml reservoirs containing from 85 to 92% saturated ammonium sulfate. After several days, the concentration of ammonium sulfate in the reservoirs was raised to between 88 and 94% saturation. Crystals, many per drop, would appear and reach full size within three days of raising the reservoir concentration. It was found that approaching the final salt concentration in two stages resulted in fewer highly twinned crystals. The crystals are rod-like, elongated along the c axis, with a triangular cross-section. The largest crystals that could be grown reproducibly have approximate dimensions 0·25 mm × 0·10 mm × 0·10 mm. The diffraction symmetry is consistent with the trigonal space group $P3_121$ or its enantiomorph $P3_221$, and the structure solution shows that the latter alternative is correct. The unit cell parameters are $a = b = 86{\cdot}2$ Å, $c = 35{\cdot}6$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$. There is one molecule in the asymmetric unit, and the solvent content of the crystals is 46% (v/v). The crystals are stable in the X-ray beam and diffract to Bragg spacings of at least 1·6 Å.

Data from four crystals were recorded at room temperature as limited-range ω-step scans on a Rigaku AFC5 four-circle diffractometer atop a Rigaku RU200 rotating anode operated at 5·0 kW. Reflections, including Friedel mates, in the range of Bragg spacings from 15·0 to 5·0 Å were collected from the first crystal, 15·0 to 6·0 Å data including Friedel mates from the second crystal, 5·0 to 3·0 Å data were collected from the third crystal, and 3·0 to 2·8 Å data from the fourth crystal. The reflections were fit with Gaussian profiles over smoothly varying backgrounds following a modification of the procedure of Hanson *et al.* (1979), and standard corrections were made for Lorentz and polarization factors. Absorption corrections were made according to North *et al.* (1968) and radiation damage corrections according to Hendrickson (1976).

Our initial approach to solving the structure was to use model-resolved anomalous phasing, combining phasing information from single-wavelength anomalous scattering (from the heme iron) and molecular replacement with the sperm whale structure. Because of the relatively small size of the crystals, the anomalous difference Patterson proved difficult to interpret. The highest peaks in the Harker section at $w = 1/3$ did, however, prove to be the peaks corresponding to the correct solution.

A molecular replacement solution was sought using the MERLOT package of programs (Fitzgerald, 1988). All of the atoms of the monoclinic sperm whale metMb structure of Takano (1977: Protein Data Bank entry 2MBN) were used as the search model "crystallized" in an orthogonal unit cell with $a = b = c = 130$ Å. K45R′ differs from sperm whale Mb at 22 of 153 residues. A Crowther (1972) rotation function map was calculated using 8·0 to 5·0 Å data, a Patterson cut-off radius of 29·9 Å, and a search grid of $\Delta\alpha = 1{\cdot}67°$, $\Delta\beta = 5{\cdot}0°$ and $\Delta\gamma = 5{\cdot}0°$. Figure 1(a) shows the section of the map containing the highest peak, which has a value of 5·6σ. The next highest peak in the map had a value of 3·5σ. A Crowther rotation function was then calculated with a finer β-spacing, $\Delta\beta = 1{\cdot}0°$, followed by a calculation of a Lattman (Lattman & Love, 1972) rotation function with $\Delta\alpha = 2{\cdot}0°$, $\Delta\beta = 1{\cdot}0°$ and $\Delta\gamma = 2{\cdot}0°$.

Upon rotation of the search molecule by the angles derived from the rotation function analyses, two 2-D Crowther–Blow (Crowther & Blow, 1967) translational searches, for $z = 1/3$ and $z = 2/3$, were performed for two molecules related by the 3-fold axis. The $z = 1/3$ map is shown in Figure 1(b). The highest peak has a value of 5·4σ, the next highest 3·4σ. The highest peak in the $z = 2/3$ map had a value of only 3·3σ. This established $P3_221$ as the correct enantiomorph. The peak in the $z = 1/3$ map yields three possible solutions for (x, y). Three 3-D translational searches for two molecules related by the 2-fold axis or by a combination of the 3-fold and 2-fold axes were computed to resolve the ambiguity in the (x, y) solution and, concomitantly, to determine z. The correct solution gave rise to peaks in the three 3-D maps that were consistently higher than those corresponding to the two incorrect solutions. In one search, though, the peak corresponding to the correct solution was only the fourth highest in the map. Our rigid-body least-squares program ROTLSQ was used to refine the rotation and translation parameters; the refined angles and translations differed from the starting values by less than 1° and less than 0·5 Å, respectively. The R-value at this stage was 0·38 for 10·0 to 3·0 Å data.

With the model-building program FRODO (Jones, 1982), the 22 side-chains differing from the sperm whale Mb sequence were changed to the correct side-chains for K45R′. After refitting several side-chains, 35 cycles of restrained least-squares refinement with PROLSQ (Hendrickson & Konnert, 1980) were run (~ 90 s c.p.u. time/cycle on a Convex C220), progressively adding higher-resolution data. Individual B-values, tightly restrained (Konnert & Hendrickson, 1980), were refined after the first 15 cycles. $2F_o - F_c$ and $F_o - F_c$ difference Fourier maps were then examined on an Evans & Sutherland PS390 graphics system; six side-chains were refitted and six water molecules were added. Fifteen more cycles of PROLSQ were then run. At this point, the R-value was 0·218 for 10·0 to 2·8 Å data, $F/\sigma_F \geqslant 5{\cdot}0$ (88% of the reflections), but the average B-value for all atoms in the model was unrealistically low. To obtain a more reasonable refinement, the B-values for all the atoms were fixed at 14·0 Å^2 (atomic co-ordinates unchanged from previous cycles) and 45 cycles of PROLSQ were run in which the scale factor between the observed and calculated structure factors was refined along with the atomic positions. The scale factor was then fixed and individual B-values were refined during the next 60 cycles. The R-value at this stage was 0·158 for 5·0 to 2·8 Å data, $F/\sigma_F \geqslant 4{\cdot}0$ (90% of the reflections), and the average

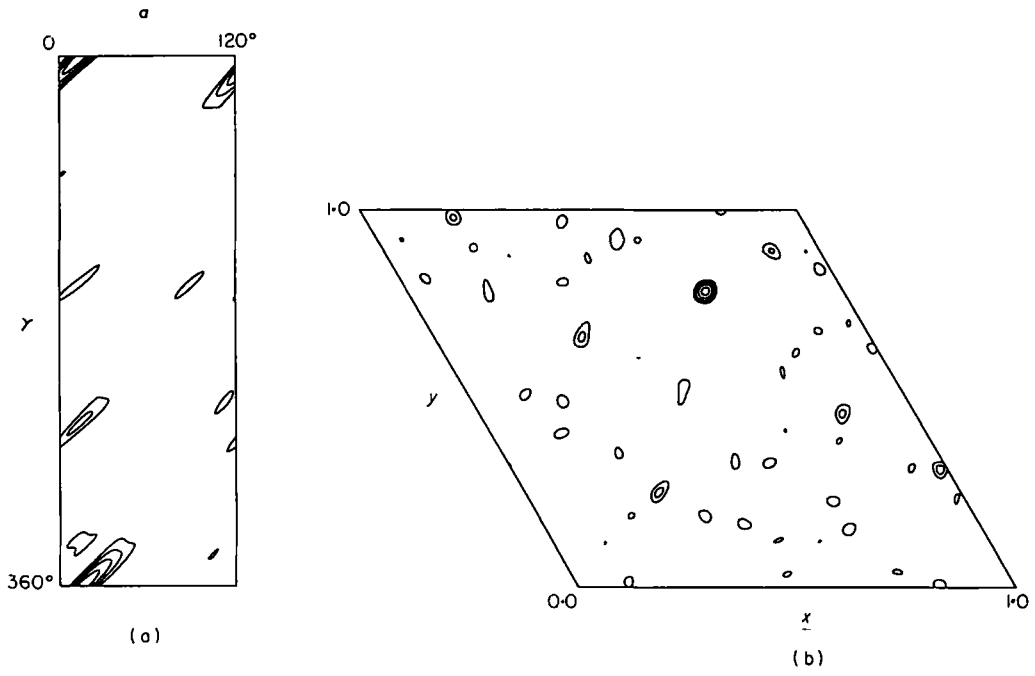

Figure 1. Molecular replacement solution of the K45R′ crystal structure. (a) The $\beta = 20°$ section of the Crowther rotation function map. Contours start at 2σ with steps of 1σ. The peak corresponding to the solution is at $\alpha = 15°$ and $\gamma = 0°$. (b) The Crowther–Blow translation function map for 2 molecules related by the 3-fold axis, $z = 1/3$. Contours start at 2σ with steps of 1σ. The peak corresponding to the solution is at $x = 0{\cdot}68$ and $y = 0{\cdot}78$.

B-value was 13·3 Å^2. The stereochemistry of this model is typified by an r.m.s. deviation from bond ideality of 0·014 Å and an r.m.s. difference of 0·55 Å^2 in *B*-values of bonded main-chain atoms. To confirm that the low *R*-value was not due to unreasonable excursions of the *B*-values, several cycles of refinement were run with the *B*-values for all atoms fixed at 13·3 Å^2. This gave an *R*-value of 0·175.

The good rigid-body fit ($R = 0{\cdot}38$) of the sperm whale Mb model to the diffraction data from human Mb had indicated that the 3-D structures of these molecules were very similar. A least-squares superposition (Hendrickson, 1979) of all backbone atoms of the currently refined 2·8 Å resolution model onto the sperm whale Mb model yields a r.m.s. displacement of only 0·5 Å. The largest displacements, of the order of 1 to 2 Å, occur at the N and C termini, and in the GH loop (residues 119 through 124). There are some significant differences in surface side-chain positions, but overall the similarities are striking. At present, the last five C-terminal residues do not lie in continuous density, although residues from this region in the current model are in contacting proximity with C-terminal residues from a symmetry-related molecule.

The environment of Arg45 is of particular interest, since this mutation from lysine affects the kinetics of carbon monoxide binding and also changes the crystallization behavior. Figure 2(a) shows the position of Arg45 within the K45R′ structure, and Figure 2(b) shows the near environment of Arg45. The conformation of Arg45 is very similar to that of sperm whale Mb in the aquomet state (Takano, 1977) and to one of the conformers of carbonmonoxy Mb (Kuriyan *et al.*, 1986). As in the sperm whale case, the guanidinium group is hydrogen-bonded to carboxylate groups from Asp60 (E3) and one of the heme propionates. However, water molecules appear to occupy sites close to the sulfate site found in aquomet sperm whale Mb. Attempts to refine models with either a sulfate group or a single water molecule in this position led to unsatisfactory results. Two water molecules restrained at hydrogen-bonding distance do refine well and interact appropriately with protein side-chains. In particular, water 160 is at good hydrogen-bonding distance to Arg45 and the distal histidine, His64 (E7). These interactions may affect ligand binding kinetics. From a superposition of the K45R′ and aquomet sperm whale structures, one finds that the sulfur of the sulfate group of sperm whale Mb is displaced from the center of the related K45R′ $2F_o - F_c$ density by approximately 1·1 Å. The proximity of His48 (CD6) from a symmetry-related molecule may be responsible for the apparent absence of a sulfate group in this region of the K45R′ structure. A difference in the crystallization behavior of the native human Mb and K45R′ led us to anticipate that Arg45 would be involved in a lattice contact. His48, the only symmetry-related residue in this vicinity, appears to exist in two conformers. The predominant conformer may make a long hydrogen bond with water 160, and the minor conformer is in van der Waals' contact with Arg45.

We hope to gain a better understanding of the structure and its implications for ligand binding upon refinement at 1·6 Å resolution. This crystal structure should also provide a basis for studies of other ligand states and other mutants. Atomic co-ordinates and diffraction data from the present work at 2·8 Å resolution have been deposited in the

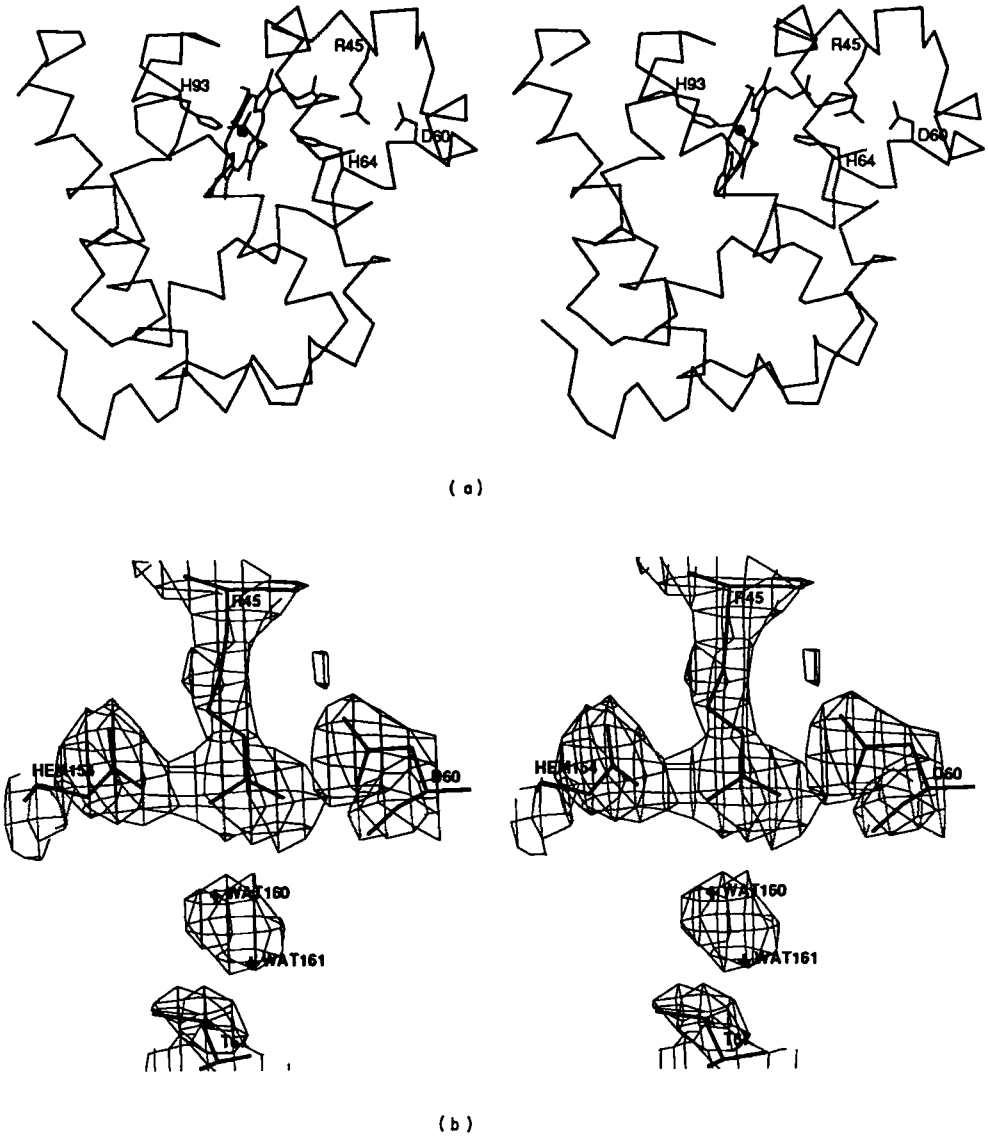

Figure 2. Stereo views of Arg45 in the K45R′ crystal structure. (a) C^{α} tracing of the K45R′ crystal structure showing the position of Arg45 within the molecule. (b) The $2F_o-F_c$ electron density in the vicinity of Arg45 contoured at 1σ.

Protein Data Bank (Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973, U.S.A.; co-ordinate entry 1MM1, structure factor entry R1MM1SF).

We thank Bill Royer for helpful discussions. This work was supported in part by NIH grant GM27738 to S.G.B.

References

Crowther, R. A. (1972). In *The Molecular Replacement Method* (Rossmann, M. G., ed.), pp. 172–178, Gordon and Breach, New York.

Crowther, R. A. & Blow, D. M. (1967). *Acta Crystallogr.* **23**, 544–548.

Fitzgerald, P. M. D. (1988). *J. Appl. Crystallogr.* **21**, 273–278.

Hanson, J. C., Watenpaugh, K. D., Sieker, L. C. & Jensen, L. H. (1979). *Acta Crystallogr. sect. A*, **35**, 616–621.

Hendrickson, W. A. (1976). *J. Mol. Biol.* **106**, 889–893.

Hendrickson, W. A. (1979). *Acta Crystallogr. sect. A*, **35**, 158–163.

Hendrickson, W. A. & Konnert, J. H. (1980). In *Computing in Crystallography* (Diamond, R., Ramaseshan, S. & Vankatesan, K., eds), pp. 13.01–13.23, Indian Academy of Sciences, Bangalore.

Jones, T. A. (1982). In *Computational Crystallography* (Sayre, D., ed.), pp. 3303–3317, Clarendon Press, Oxford.

Konnert, J. H. & Hendrickson, W. A. (1980). *Acta Crystallogr. sect. A*, **36**, 344–350.

Kuriyan, J., Wilz, S., Karplus, M. & Petsko, G. A. (1986). *J. Mol. Biol.* **192**, 133–154.

Lambright, D. G., Balasubramanian, S. & Boxer, S. G. (1989). *J. Mol. Biol.* **207**, 289–299.

Lattman, E. E. & Love, W. E. (1972). *Acta Crystallogr. sect. B*, **26**, 1854–1857.

North, A. C. T., Phillips, D. C. & Mathews, F. S. (1968). *Acta Crystallogr. sect. A*, **24**, 351–359.

Takano, T. (1977). *J. Mol. Biol.* **110**, 537–568.

Varadarajan, R., Zewert, T. E., Gray, H. B. & Boxer, S. G. (1989*a*). *Science*, **243**, 69–72.

Varadarajan, R., Lambright, D. G. & Boxer, S. G. (1989*b*). *Biochemistry*, **28**, 3771–3781.

Edited by R. Huber