THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 264, No. 35, Issue of December 15, pp. 21052-21061, 1989
Printed in U.S.A.

The 2.4-Å Crystal Structure of *Scapharca* Dimeric Hemoglobin

COOPERATIVITY BASED ON DIRECTLY COMMUNICATING HEMES AT A NOVEL SUBUNIT INTERFACE*

(Received for publication, June 28, 1989)

William E. Royer, Jr. and Wayne A. Hendrickson

From the Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032

Emilia Chiancone

From the Consiglio Nazionale delle Ricerche Center of Molecular Biology, Department of Biochemical Sciences, University "La Sapienza," 00185 Rome, Italy

The crystal structure of the cooperative dimeric hemoglobin from the arcid clam, *Scapharca inaequivalvis*, has been determined in the carbonmonoxy state. The phase problem was solved for reflections with Bragg spacings greater than 3 Å using anomalous scattering from the porphyrin iron atoms measured at a single wavelength in combination with molecular averaging. The model built into this electron density map has been refined at 2.4 Å resolution by means of stereochemically restrained least squares minimization to a conventional *R*-value of 0.156. The root mean square deviation from ideal bond lengths and angles are 0.013 Å and 1.7°, respectively. In addition to the 2336 hemoglobin atoms, 214 water molecules have been incorporated into the model.

This structure reveals the details of an assemblage of two identical myoglobin-like subunits that is radically different from vertebrate hemoglobins. The subunit interface is formed by direct apposition of the E and F helices, whereas these surfaces are external in vertebrate hemoglobins. The interface has both hydrophobic and hydrophilic character. Two symmetrically related hydrophobic regions are formed between subunits. Six residues are involved in each of these regions that pack tightly enough to exclude water but have only a few atoms in close van der Waals contact. A number of ordered water molecules line the interface and form bridging hydrogen bonds between subunits. Four intersubunit ionic interactions are formed, two of which involve negatively charged propionate groups of the porphyrin. In contrast to cooperative vertebrate hemoglobins, a hydrogen bond network provides a direct route for communication between the two heme groups.

Nature has, on a number of occasions, found that the efficiency of protein molecules can be improved by allowing communication to occur between functional units. A number of enzyme systems (*e.g.* aspartate transcarbamoylase; Schachman, 1988) have been found to exhibit such cooperative functioning. Probably the most widely studied example of a cooperative, allosteric molecule is the mammalian hemoglobin tetramer (see Perutz *et al.*, 1987). In the case of hemoglobin, cooperativity can be defined as the increasing affinity of the molecule for oxygen as oxygen binding proceeds.

A simpler model system for investigating the cooperative function of protein molecules can be found in the hemoglobins from the clams of the arcid family, otherwise known as blood clams. Intracellular dimeric and tetrameric hemoglobins have been found in a number of arcid clams (Ohnoki *et al.*, 1973; Furuta *et al.*, 1977; Como and Thompson, 1980; Chiancone *et al.*, 1981; San George and Nagel, 1985; Borgese *et al.*, 1987). The hemoglobins from *Scapharca inaequivalvis* are formed from three distinct polypeptide chains, two of which associate to form a 66-kDa tetramer (HbII) and a third that associates to form a 33-kDa homodimer (HbI). Both molecules bind oxygen cooperatively, with HbI exhibiting a maximum Hill coefficient of 1.5 and HbII showing a maximum Hill coefficient of 2.0 (Chiancone *et al.*, 1981). The amino acid sequence of HbI has been determined by Petruzzelli *et al.* (1985). With two identical binding sites and the absence of heterotropic ligand effects, HbI is an ideal model system for investigating allostery. We have undertaken a crystallographic study of Scapharca HbI to learn the structural bases for its cooperative oxygen binding.

In an earlier low resolution study on carbon monoxide liganded-HbI and HbII, Royer *et al.* (1985) found that the *Scapharca* hemoglobin subunits share a similar tertiary structure with other hemoglobins, but that the assemblage of subunits to form dimers and tetramers is radically different from mammalian hemoglobins. The dimer is assembled by associating subunits across the surface formed by the E and F helices. The tetrameric assemblage, although formed from different subunits, is essentially that of two dimers. In *Scapharca* hemoglobins the E and F helices are on the inside of the molecule and the G and H helices are on the outside, whereas the reverse is true of mammalian hemoglobins.

We report here the extension of the resolution of the carbon monoxide-liganded HbI structure to 2.4 Å. This structure reveals the details of this cooperative molecule's assemblage and provides clues for the structural basis for cooperative oxygen binding.

MATERIALS AND METHODS

Crystallization—Crystals of *Scapharca* HbI were grown at room temperature under an atmosphere of carbon monoxide in Corning culture tubes closed with Critocaps and sealed with DeKhotinsky cement (Thomas Scientific). A typical crystallization experiment consisted of mixing 15 μl of CO-liganded HbI (28 mg/ml) with 4.3 M phosphate buffer at pH 7.5. Crystals grew from solutions whose final

* This work was supported in part by a grant from the National Science Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

phosphate concentration was 2.1–2.4 M. Seeding these solutions with smaller crystals was essential for reproducibly growing large (0.2 × 0.2 × 0.5 mm) crystals. These crystals show the symmetry of space group C2 with cell constants a = 93.25 Å, b = 43.98 Å, and c = 83.50 Å, β = 122.03°. There is one dimer per asymmetric unit.

Data Collection—Diffraction data were collected with an AFC5 diffractometer mounted on a Rigaku RU 200 rotating anode generator (Molecular Structure Corporation, The Woodlands, TX) using peak-top ω step-scans. Data reduction included fitting peak profiles to Gaussian distributions (Hanson *et al.*, 1979), a semi-empirical absorption correction (North *et al.*, 1968), radiation damage correction (Hendrickson, 1976), and correction for Lorentz and polarization effects. The expected average signal for the anomalous scattering from the iron atoms in HbI was calculated to be about 1.9% of $|F|$ for 1.54178 Å wavelength x-rays. Thus, the data must be measured very accurately, and, in particular, care must be taken to minimize systematic errors between Friedel pairs. For the anomalous scattering experiment, we measured hkl reflections at $(2\theta, \omega, \phi, \chi)$ and their Friedel mates −h,−k,−l at $(-2\theta, -\omega, \phi, \chi)$ in groups of 20 as suggested by Smith and Hendrickson (1982). Data were first collected from one crystal for all reflections and their Friedel mates with Bragg spacings between 30 and 3.0 Å and put on an approximate absolute scale using intensity statistics (Wilson, 1942). An anomalous difference Patterson (Rossmann, 1961) calculated from these data had peaks for the four independent vectors expected from two iron atoms in the asymmetric unit of space group C2, but the map was very noisy. To improve the phasing power by collecting more accurate data, we recollected these data from two more crystals, one from 30 to 3.7 Å and the other from 3.75 to 3.0 Å. These data were more accurate than the diffraction data from the first crystal because twice the time was spent in counting each reflection and because more care was taken to bathe the crystals fully in the x-ray beam. The intensity data from these two crystals were put on an approximate absolute scale by scaling to the data from the first crystal. Diffraction data from a fourth crystal were collected for the unique data (no Friedel pairs) corresponding to Bragg spacings between 3.05 and 2.4 Å. Only half the time was spent counting each reflection from this crystal compared with the previous two crystals.

Anomalous Scattering Analysis—Iron positions (consistent with the low resolution results of Royer *et al.*, 1985) were obtained from a difference Patterson map (Rossmann, 1961). The iron coordinates were refined against the strongest 41% of anomalous differences to an R-value of 0.325. Phase probability coefficients (Hendrickson and Lattman, 1970) were then calculated from the anomalous differences.

Molecular Replacement and Averaging—Royer *et al.* (1985) determined the crystal structure of Scapharca HbII using multiple heavy atom isomorphous replacement and molecular averaging about a non-crystallographic diad relating two halves of the tetramer. They then excised the 5.5-Å electron density corresponding to a half tetramer, averaged this density about a local diad relating the two subunits, and used it in molecular replacement calculations against diffraction data collected from HbI crystals. These calculations demonstrated the similarity of HbI to HbII at low resolution and established the orientation and position of the molecule in the dimer crystals. This 5.5-Å crystal structure of HbI was used as the starting point for resolving the phase ambiguities from our anomalous scattering experiment. Molecular averaging was used to improve and extend the phasing as discussed under "Results." The programs used for the molecular averaging were originally written by J. L. Smith (Purdue University) and W. A. Hendrickson and modified for this problem.

Modeling and Refinement—All fitting of protein models to electron density maps were performed on a Silicon Graphics IRIS 3020 system using a version of FRODO (Jones, 1985) modified for this system by C. M. Cambillau (Marseille, France).

Cycles of stereochemically restrained least squares refinement (Konnert, 1976) were performed using the PROTIN/PROLSQ package (Hendrickson and Konnert, 1980; Hendrickson, 1985). We used three types of Fourier coefficients for rebuilding the protein model at various stages of refinement: 1) $(2|F_o| - |F_c|)\exp(i\alpha_c)$; 2) $(|F_o| - |F_c|)\exp(i\alpha_c)$ "difference" maps; and 3) $(|F_o| - |F_c|)\exp(i\alpha_c)$ "fragment" maps in which the structure factor (including phase) calculation did not include the region being refit. An overall isotropic B-value of 14 Å^2 (consistent with intensity statistics) was used for the first 15 cycles of refinement, after which individual isotropic atomic B-values were refined. X-ray reflections corresponding to Bragg spacings between 5.0 and 3.0 Å were used for the first 24 cycles; at cycle 25 reflections to 2.6 Å were added, and at cycle 30 reflections to 2.4 Å were included. After 50 cycles of refinement, by which time the R-value had reached 0.223, we began adding solvent molecules to the model and also began including the reflections in the 10–5-Å shell in the refinement. Water molecules were added using FRODO at positions where the electron density in difference maps was greater than 0.2e/Å^3 if the position was chemically reasonable (hydrogen bonding potential and no interpenetration with other atoms). Occupancy and B-factors were refined in alternate cycles for these water molecules. If the occupancy for a water site fell below 0.3, it was removed from the model. Non-crystallographic symmetry restraints were applied during all but the last ten refinement cycles. It was obvious early in the procedure that, due to lattice contacts, the pre-A helices had somewhat different conformations in the two subunits. For this reason, only very loose symmetry restraints (σ = 5 Å) were applied to residues 1–9. Tight restraints (σ = 0.05 Å) were applied to main chain atoms in residues 10–146 and medium restraints (σ = 0.3 Å) were applied to side chains 10–146 and the heme groups.

The program WATERPATH, written by S. Sheriff (Squibb Institute for Medical Research), was used to analyze the hydrogen bonding of water molecules in the molecular interface.

RESULTS

Phasing and Refinement—The extension of phases from the low resolution molecular replacement solution (Royer *et*

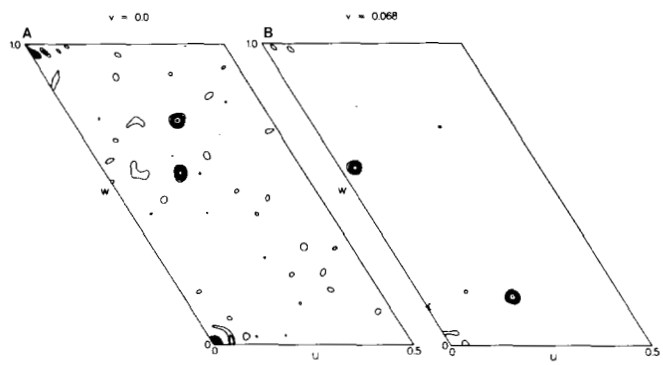


FIG. 1. **Anomalous difference Patterson map from *Scapharca* HbI CO crystals.** This 3.0-Å map is contoured every 2σ (starting at 2σ) where σ is the standard deviation of the entire Patterson map. This map could be readily interpreted in terms of two anomalously scattering atoms per asymmetric unit. *A*, Harker section at v = 0 has peaks for the vectors between the crystallographic symmetry related positions (space group C2) for the two iron atoms in the asymmetric unit. *B*, section at v = 0.068 shows peaks for the two independent cross-vectors between non-crystallographically related iron positions.

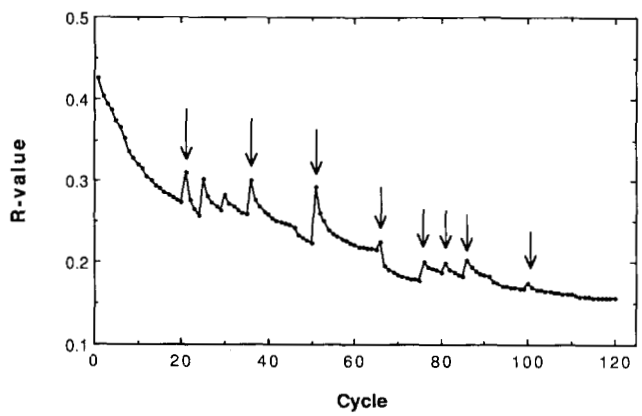


FIG. 2. **Progress of refinement.** The R-value at the beginning of each refinement cycle is plotted for reflections with $I > 2.5\sigma_I$. Refitting of the model to electron density maps was performed after cycles 20, 35, 50, 65, 75, 80, 85, and 99 as indicated by the *arrows*. Water molecules were added after cycles 50 and 65. Initially only reflections with Bragg spacings between 5.0 and 3.0 were included in the R-value calculation. Reflections with Bragg spacings between 3.0 and 2.6 Å were added in after cycle 24, those between 2.6 and 2.4 Å were included after cycle 29, while those between 10 and 5 Å were included after cycle 50.

TABLE I

Refinement statistics: Scapharca CO HbI

The stereochemical parameters are defined by Hendrickson and Konnert (1980) and Hendrickson (1985). The target σ shown for each parameter is that used for the final 20 cycles of refinement. At various earlier times during the refinement, slightly different target σ values were used for non-bonded contacts, non-crystallographic symmetry, and isotropic thermal factor restraints. Early in the refinement, significantly higher target σ values were used on the structure factors in order to lower the weight on the structure factor terms and preserve satisfactory model geometry. ($R = \Sigma| |F_o| - |F_c| |/\Sigma |F_o|$).

Atomic model	
Hemoglobin atoms:	2336
Solvent molecules:	214
Diffraction data:	
Reflections:	10,423 (10–2.4 Å, I> $2.5\sigma_I$
R-value:	0.156

	Stereochemical	
	Root mean square deviations from ideality	Target σ
Bond distances	0.013 Å	0.02
Angle distances	0.032 Å	0.03
1–4 distances	0.038 Å	0.05
Planarity	0.010 Å	0.02
Chiral volume	0.157 Å^3	0.15
Non-bonded contacts		
Single torsion	0.187 Å	0.50
Multiple torsion	0.236 Å	0.50
Possible H-bond	0.277 Å	0.50
Conformation torsion angles		
Planar (0, 180°)	1.8°	3.0
Staggered (+/−60, 120°)	23.0°	15.0
Orthonormal (+/−90°)	29.2°	20.0
Non-crystallographic symmetry		
Main chain (10–146)	0.117 Å	0.05
Side chains (10–146)	0.418 Å	0.30
Residues 1–9	1.143 Å	5.00
Isotropic thermal factors		
Main chain bond	0.582 Å^2	1.0
Main chain angle	0.987 Å^2	1.5
Side chain bond	1.216 Å^2	1.5
Side chain angle	1.929 Å^2	2.0
Average $\|F_o - F_c\|$	89.44	(50.0–110.0* ($\sin\theta/\lambda - 1/6$))

al., 1985) to 3.0 Å was based on anomalous scattering measurements from a single wavelength (CuKα, 1.54178 Å). The most significant contribution to the anomalous scattering signal is made by the porphyrin iron atoms. Their positions could be accurately defined from the anomalous difference Patterson shown in Fig. 1. Just as is the case for a single heavy atom derivative, anomalous scattering phasing using measurements at a single wavelength leads, in general, to a phase ambiguity. In the case of crambin (Hendrickson and Teeter, 1981), trimeric hemerythrin (Smith *et al.*, 1983) and the Cs-gramicidin complex (Wallace and Ravikumar, 1988) the phases from the partial structure of the anomalous scattering atoms were used to break the phase ambiguities. We deemed the contribution of the iron atoms in hemoglobin to the overall diffraction pattern to be too small to reliably break the phase ambiguities from our anomalous scattering measurements. Instead we used the phases from the molecular replacement solution to resolve the phase ambiguities to 5.5 Å and followed this by molecular averaging of the electron density of the two subunits to improve and extend the phasing. The resultant 5.5 Å map after 16 cycles of symmetry averaging was superior to that obtained for the *Scapharca* tetrameric hemoglobin (Royer *et al.*, 1985).

The use of molecular averaging to extend phases has been very successfully applied to arthropod hemocyanin (Gaykema *et al.*, 1984) and to a number of spherical virus structures (Rossmann *et al.*, 1985; Hogle *et al.*, 1985; Luo *et al.*, 1987; Acharya *et al.*, 1989). The presence of 5–60-fold redundancy in these cases was powerful enough to use low resolution maps to derive phases at higher resolution, provided that significantly small steps of increasing diffraction angle were used. Arnold and Rossmann (1986) have shown that the power of this method is proportional to the square root of the redundancy. With one dimer in the asymmetric unit of the *Scapharca* HbI crystals, only 2-fold non-crystallographic redundancy was available. We reasoned that although this would not be adequate for extending the phases in the absence of other phase information, perhaps it could be powerful enough to break the phase ambiguities from the anomalous data. We used very small phase extension steps (½ lattice point along a^*) followed by 6 cycles of molecular averaging before calculating phases for the next reciprocal lattice shell. There were 28 such steps to include all reflections with Bragg spacings greater than 3.0 Å. The final 3.0-Å anomalous scattering/molecular averaged phase set (5709 reflections) comprised the following: 3653 reflections phases for which the molecular averaged phase was just used to resolve the anomalous scattering phase ambiguity, 382 reflections with phasing based

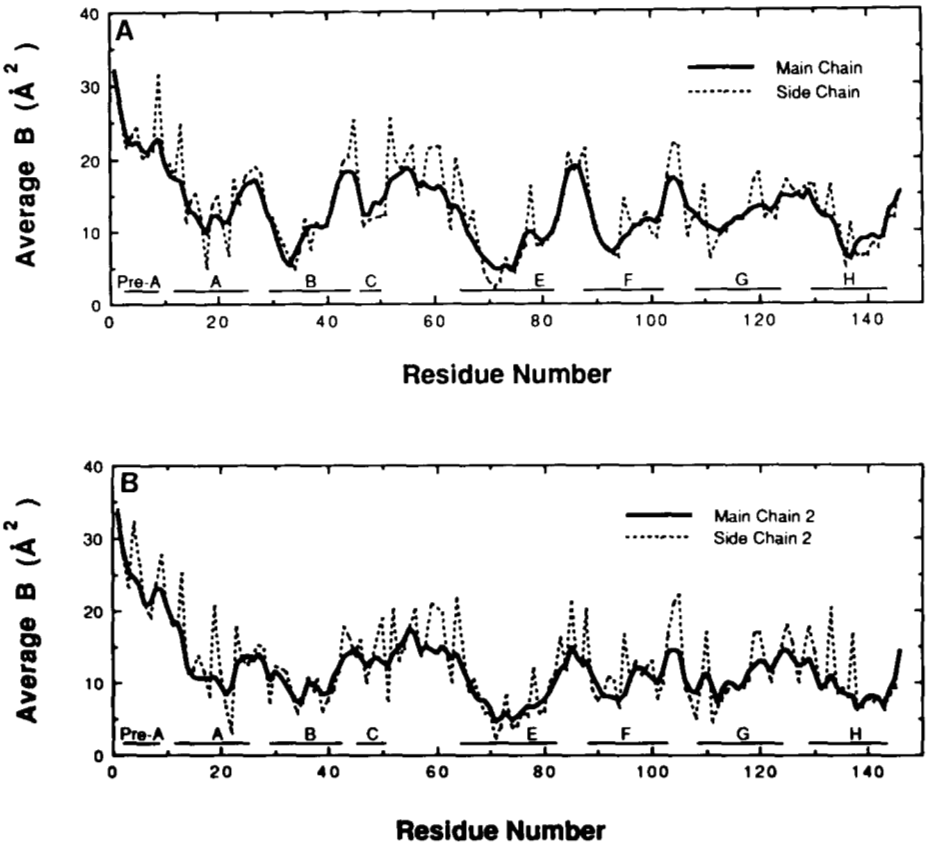


FIG. 3. **Average thermal (B factor) parameters for each residue.** Plotted in *solid curves* are the averages for the main chain atoms while *dashed curves* show the average side chain B factors. The helical segments are indicated at the bottom of each graph. B factors are highest at the amino termini and lowest for the interface regions in the E and F helices and the beginning of the B helix. *A*, subunit 1; *B*, subunit 2.

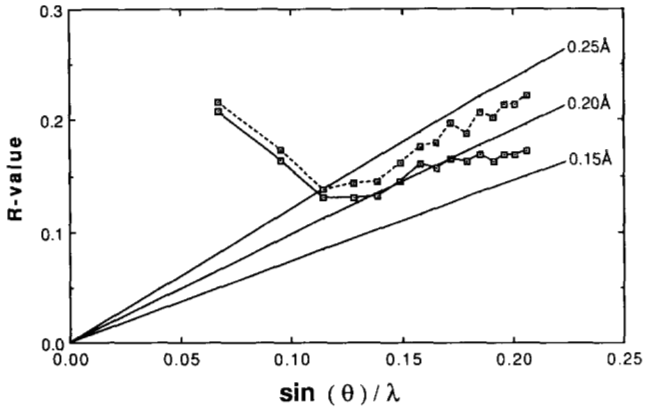


FIG. 4. **Plot of variation of *R*-value with scattering angle.** *Solid curve* shows *R*-value including those reflections with $I > 2.5\sigma_I$ in the 10–2.4-Å shell, and *dashed curve* shows *R*-value calculated using all reflections in the 10–2.4-Å shell. Also plotted are expected *R*-values given root mean square coordinate errors of 0.15, 0.20, and 0.25 Å according to Luzzatti (1952).

solely on unimodal anomalous scattering phases, 1119 reflections phased using phase combination of molecular averaging and anomalous scattering, and 555 centric reflections based solely on the molecular averaged map. This electron density map, although somewhat noisy, contained enough detail to allow building a preliminary model of *Scapharca* HbI. The heme planes were readily apparent and many helical segments were clear. As well, side chain density was clear for the distal histidine (69), proximal histidine (101), tryptophans 22 and 135, and phenylalanines 41 and 51. A model of *Aplysia* myoglobin based on a 2.0-Å refinement (Bolognesi *et al.*, 1985) was used as an aid in the interpretation of this map, but significant differences were apparent in the subunit tertiary structures even between these two molluscan hemoglobins. The ability to locate a side chain confidently in each of the A, B, C, E, F, and H helices increased the accuracy of the helical segments in this model. Other than the residues cited above, this model was polyalanine.

While it might have been possible to obtain a better fit to this map, the presence of a model, even with significant errors, allows one to use structure factors calculated from this model to resolve the anomalous phase ambiguity. This model-resolved phasing technique has been successfully applied to myohemerythrin (Smith and Hendrickson, 1982; Sheriff *et al.*, 1987). Use of model-resolved anomalous phases along with 16 cycles of symmetry averaging dramatically improved the *Scapharca* HbI electron density map. Many poorly fit main chain regions were improved, and reasonable density appeared for over a third of the side chains (53/146 residues in a subunit) allowing atomic modeling to be extended to these groups. This procedure of refitting followed by model-resolved anomalous phasing was repeated three more times. In the final 3.0-Å map, reasonable density was apparent for 95 side chains (80% of the non-glycine, non-alanine residues). All fitting was done to a single subunit with a second subunit generated by the non-crystallographic diad axis. Calculation of structure factors from this model yielded an *R*-value of 0.425.

The model derived from the 3.0-Å electron density map was considered sufficiently accurate to begin stereochemically restrained least squares refinement (Hendrickson and Konnert, 1980; Hendrickson, 1985). The progress of refinement is shown in Fig. 2. After 20 cycles of refinement using reflections with Bragg spacings greater than 3.0 Å, diffraction data corresponding to Bragg spacings between 3.0 and 2.4 Å were collected from another crystal and included in the refinement. After 120 cycles of least squares refinement along with eight

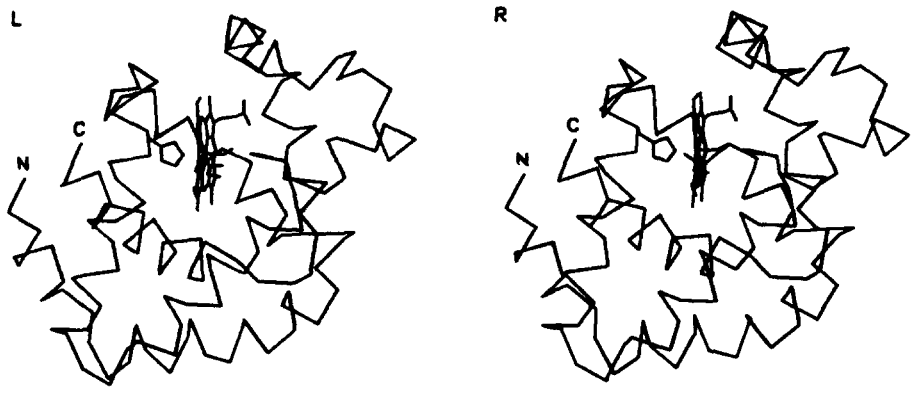


FIG. 5. **Stereo α-carbon plot of one subunit of *Scapharca* HbI.** The amino and carboxyl termini are labeled *N* and *C*, respectively. Note the pre-A helix running nearly vertically at the left edge of the diagram. The characteristic "V" formed by the E and F helices can be seen to the front of this diagram. The heme group is embedded between these helices and ligated to the proximal histidine (His-101) in the F helix (shown on the *left side* of the heme plane). The distal histidine (His-69) in the E helix can be seen to the *right side* of the heme plane along with the carbonyl ligand. Note how the propionate side groups of the heme point out from this surface where the dimer interface is formed.

TABLE II
Helical assignments in Scapharca HbI

Helix	Residue span	Subunit 1	Subunit 2
		avg. φ/ψ	
Pre-A	3–9	−65.3/−34.4	−61.9/−42.4
A	12–25	−64.7/−42.3	−65.4/−42.2
B	29–43	−66.4/−42.8	−65.0/−43.7
C	45–50	−65.6/−23.7	−59.8/−29.0
E	64–82	−64.1/−40.7	−65.2/−39.7
F	88–103	−66.4/−39.7	−65.8/−39.9
G	108–124	−63.2/−36.4	−61.6/−37.6
H	129–144	−60.8/−40.4	−63.3/−37.3

sessions of model rebuilding, the R-value is 0.156 for the 10,423 reflections ($I < 2.5\sigma_I$) that were used in the refinement. The *R*-value for all 11,269 reflections with Bragg spacings between 10 and 2.4 Å is 0.178. Detailed characteristics of the refined model are given in Table I. The low root mean square deviations from ideality demonstrate the good molecular geometry in this model. A plot of the average *B*-values is given in Fig. 3. It is noteworthy that the lowest *B*-values occur in regions of contact between the two subunits, namely in the E and F helices and the beginning of the B helix. Fig. 4 shows a plot of the *R*-value *versus* the scattering angle based on the method of Luzzati (1952). This would put an upper limit of approximately 0.2 Å on the root mean square error in the atomic coordinates.[1]

Tertiary Structure of the Subunits—The overall fold of one subunit of *Scapharca* dimeric hemoglobin can be seen from the α-carbon plot in Fig. 5. The most notable difference between this hemoglobin and other hemoglobins and myoglobins is the presence of an additional helix at the amino terminus. As before (Royer *et al.*, 1985) we designate this helix as pre-A. The two subunits have nearly identical structures and, other than the pre-A helix, they have rather standard myoglobin folds. The helical segments, as judged from the φ/ψ angles, are shown in Table II. Based on these designations, 75% (110/146) of the residues are in helical conformations.

After 2 non-helical residues at the amino termini, the next 7 residues of each protomer fold into the pre-A helix. The first turn of this helix shows α-helical hydrogen bonding pattern, whereas the second turn adopts a 3_{10} conformation. Both pre-A helices are involved in crystal contacts, albeit different ones. As a result, they adopt somewhat different conformations in the two subunits. The pre-A helix of subunit

[1] The atomic coordinates and structure factors (codes 1SDH and R1SDHSF, respectively) have been deposited in the Protein Data Bank (Bernstein *et al.*, 1977).

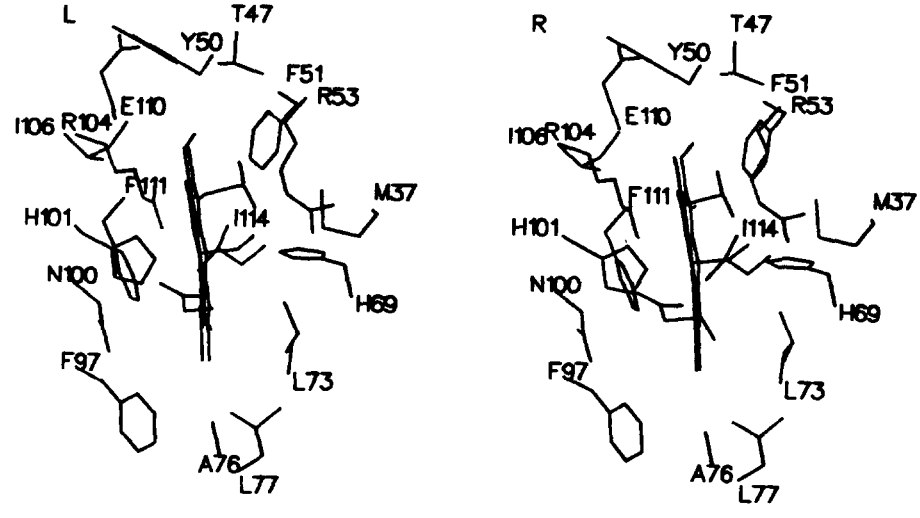


FIG. 6. **Stereo diagram of the porphyrin binding pocket in one subunit of *Scapharca* HbI.** Included with the porphyrin are all side chains that have one or more atom within 4.0 Å of a heme atom or the carbonyl ligand. In addition to the non-polar side groups lining the binding pocket, one can see the distal (His-69) and proximal (His-101) histidines and the two arginines (Arg-53 and -104) that interact with the two negatively charged propionate groups of the heme.

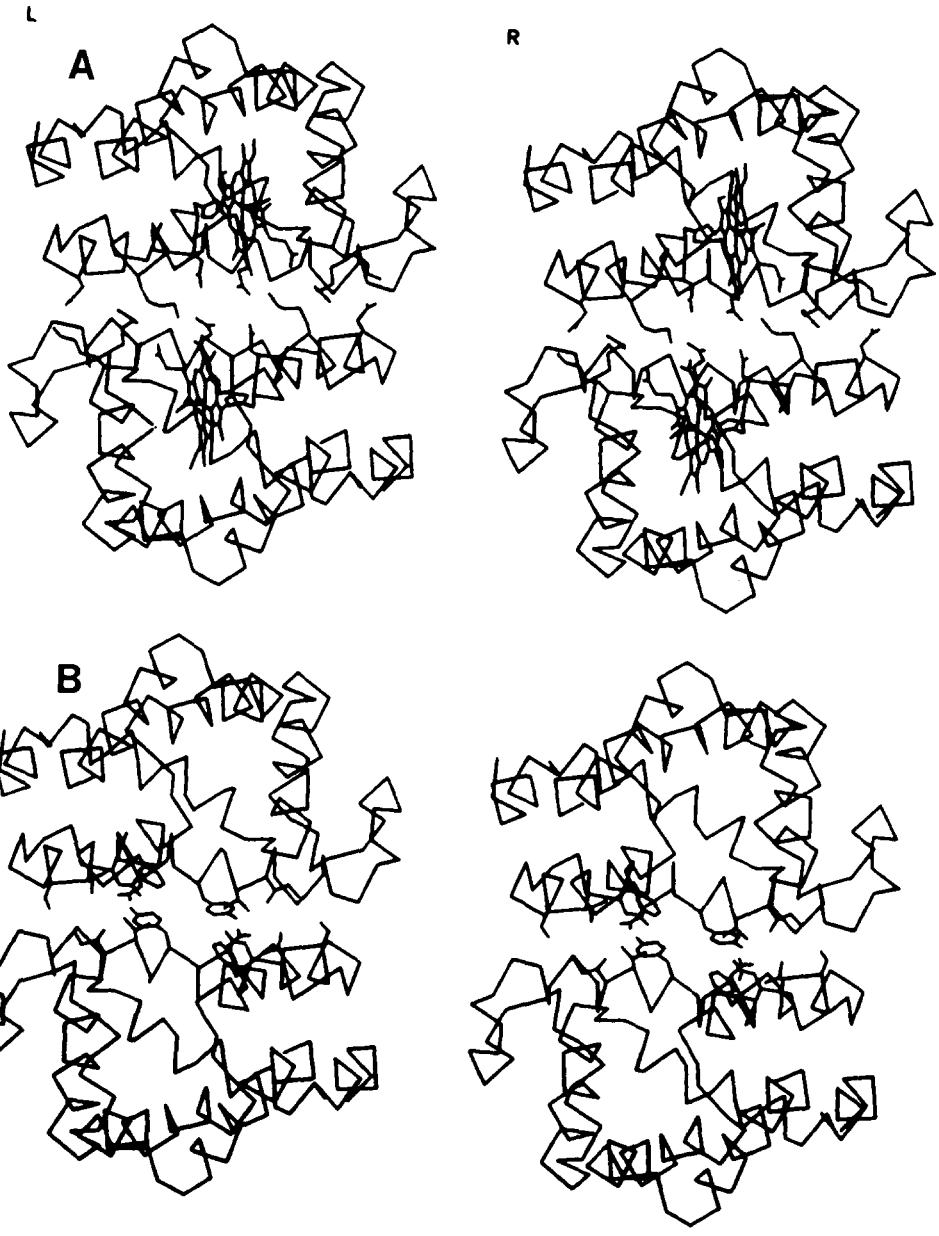


FIG. 7. **Stereo diagrams of *Scapharca* HbI plotted down the molecular diad.** In each diagram an α-carbon plot of the dimer is shown. *A*, the hydrophilic interactions of the Heme-F and B–F regions. Near the *right* and *left ends* of the diagram are the salt bridges formed between lysine 30 (at the beginning of the B helix) and aspartate 89′ (at the beginning of the F helix). Near the *center*, the interactions of the propionate groups from the hemes are shown including intrasubunit salt bridges with arginines 53 and 104 and interactions across the interface with lysine 96′ and aspargine 100′. Also shown are the side chains of the proximal histidines. *B*, the largely hydrophobic interactions of the E–E and E–F regions. For each subunit, the side chains for tyrosine 75, asparagine 79′, and aspartate 82′ in the E–E region and side chains for isoleucine 71, threonine 72, cysteine 92′, valine 93′, and phenylalanine 97′ in the E–F regions are shown. The heme groups have been omitted from this plot for clarity. Two hydrophilic interactions are the hydrogen bonds between tyrosines 75 and 75′ and aspartates 82′ and 82. Note the closer approach of the atoms in the E–E region near the center of the plot to that in the neighboring E–F regions.

1 shows an average ψ angle about halfway between that expected for an α-helix and a 3_{10} helix, while the pre-A helix of subunit 2 shows a slightly larger average ψ angle (see Table II).

The myoglobin fold begins properly with the A helix, which has the hydrogen bonding of an α helix except for a missing hydrogen bond between aspartate 20 and lysine 24 where the helix bends toward the B helix. The B helix is a standard α helix for its length of 15 residues. The C helix has a 3_{10} conformation except for the carbonyl oxygen of glutamate 46 which is pointing toward the solvent rather than the nitrogen of glycine 49. The CD loop begins with the strictly conserved phenylalanine 51 and extends for 13 residues with main chain atoms forming hydrogen bonds characteristic of a 3_{10} helix five times (involving the carbonyls of residues 51, 52, 56, 57, 60). Like the α chains in mammalian hemoglobins, there is no D helix. The E helix with 19 residues is the longest helix in this structure. For most of its length its residues adopt a standard α helix, except for a single turn of 3_{10} near its center which is involved in the intersubunit contact (see below). The

TABLE III

Residue pairs within 4.2 Å across dimer interface

Distance of closest atom pair between residues is given.

Subunit 1	Distance	Subunit 2
	Å	
Lys-30 (Nζ)	3.9	Asn-86 (Nδ)
Lys-30 (Nζ)	2.9	Asp-89 (Oδ1)
Arg-53 (Nη2)	3.8	Lys-96 (Cγ)
Asp-64 (O)	3.4	Cys-92 (Cβ)
Arg-67 (Nη1)	3.6	Asp-88 (O)
Arg-67 (Cδ)	3.5	Asp-89 (Oδ2)
Arg-67 (Nη1)	3.6	Cys-92 (Sγ)
Gly-68 (Cα)	3.6	Cys-92 (Sγ)
Gly-68 (O)	4.2	Lys-96 (Cδ)
Ile-71 (Cγ2)	3.6	Asn-79 (Nδ)
Ile-71 (Cγ2)	4.2	Gln-83 (Nε)
Ile-71 (Cγ2)	4.1	Val-93 (Cγ2)
Thr-72 (Oγ)	3.5	Lys-96 (Cδ)
Thr-72 (Cγ)	3.5	Phe-97 (Cζ)
Thr-72 (Oγ)	3.5	Asn-79 (Oδ)
Tyr-75 (Cε2)	3.6	Gln-78 (Cγ)
Tyr-75 (Cε1)	3.3	Asn-79 (Nδ)
Tyr-75 (Oη)	2.3	Asp-82 (Oδ2)
Tyr-75 (Oη)	4.0	Gln-83 (Nε)
Tyr-75 (Cδ2)	3.9	Tyr-75 (O)
Gln-78 (Cγ)	3.7	Tyr-75 (Cε2)
Asn-79 (Oδ)	3.3	Tyr-75 (Cε1)
Asn-79 (Oδ)	3.4	Ile-71 (Cγ2)
Asn-79 (Nδ)	3.3	Thr-72 (Oγ)
Asp-82 (Oδ2)	2.7	Tyr-75 (Oη)
Gln-83 (Nε)	3.6	Ile-71 (Cγ2)
Gln-83 (Nε)	3.5	Tyr-75 (Oη)
Asn-86 (Nδ)	4.2	Lys-30 (Nζ)
Asp-88 (O)	3.3	Arg-67 (Nη1)
Asp-89 (Oδ1)	2.7	Lys-30 (Nζ)
Asp-89 (Oδ2)	3.3	Arg-67 (Cδ)
Cys-92 (Cβ)	3.3	Asp-64 (O)
Cys-92 (Sγ)	3.8	Arg-67 (Nη1)
Cys-92 (Sγ)	3.5	Gly-68 (N)
Val-93 (Cγ2)	4.1	Gly-68 (Cα)
Val-93 (Cγ2)	3.8	Ile-71 (Cγ2)
Lys-96 (Nζ)	2.5	Heme (OC21)
Lys-96 (Cδ)	4.1	Thr-72 (Cγ)
Lys-96 (Cγ)	3.9	Gly-68 (O)
Phe-97 (Cζ)	3.2	Thr-72 (Cγ)
Asn-100 (Nδ)	2.9	Heme (OC21)
Heme (OC21)	2.4	Lys-96 (Nζ)
Heme (OC21)	3.0	Asn-100 (Nδ)
Heme (OC22)	4.0	Heme (OC22)

F helix is an α helix with several of the hydrogen bonds elongated toward the outside of the subunit, leading to a gentle curve of the helix (see Fig. 7). The G helix has a kink caused by the inability of proline 117 to contribute to hydrogen bonding. The H helix is again primarily an α helix, but the last turn has 3_{10} hydrogen bonding. The carboxyl terminus is thus brought to within 8 Å from the amino terminus in each subunit.

The porphyrin binding pocket is shown in Fig. 6. Seventeen residues have one or more atoms within 4.0 Å of the porphyrin or carbonyl ligand. The proximal histidine is ligated through Nε to the porphyrin iron (2.2 Å in one subunit and 2.15 Å in the second subunit), and the distal histidine has its Nε in a position to allow hydrogen bonding to the ligand (3.1 Å, 3.0 Å). In addition, the sulfur atom of methionine 37 is only 3.6 Å (3.4 Å in subunit 2) from the bound carbonyl oxygen. The strictly conserved phenylalanine CD1 (51) stacks against the porphyrin in a manner similar to that seen in other globins. Other hydrophobic side chains with two or more atoms within 4.0 Å of heme atoms are tyrosine 50, leucine 73, phenylalanines 97 and 111, and isoleucines 106 and 114.

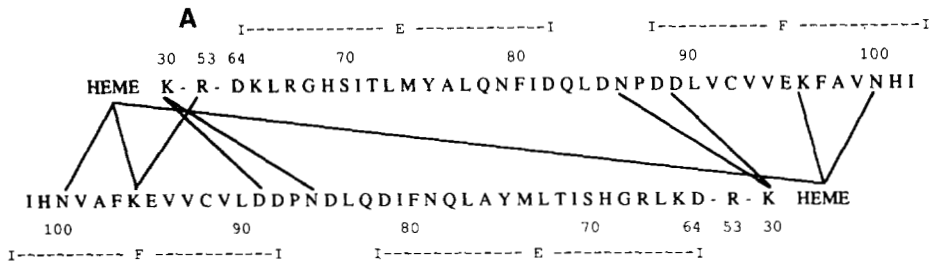
In addition to its hydrophobic region, protoporphyrin IX has two negatively charged groups: the carboxylates at the end of the propionate side chains which point out from the subunits. In *Scapharca* HbI, these groups interact with the positively charged ends of arginine 53 and arginine 104 as can be seen in Fig. 6. A carboxylate oxygen from one propionate group is 3.0 Å (2.9 Å in the second subunit) from arginine 53 Nε, and a carboxylate oxygen from the other propionate is 2.6 Å (2.4 Å) from arginine 104 Nη2.

Interface between Subunits—At low resolution Royer *et al.* (1985) identified the dimer formed by the non-crystallographic diad (used for symmetry averaging in this study) as being the functional dimer. This dimer was chosen because it appeared to have the most extensive interface and because such an assemblage would explain earlier sequence and functional data. In our present model this dimer has 123 atom pairs in contact (within 4.0 Å) between subunits. Other axes that could conceivably relate two halves of the functional dimer in these crystals are the two distinct crystallographic diads along *b*. One of these passes through a solvent channel. The other relates two dimers that have significant contacts formed between three different monomer pairs. One pair of monomers related to one another directly by this diad has 35 atom pairs in contact. An additional 16 atom pairs make contact in each of two equivalent monomer pairs involved in heterologous (non-diad) interfaces. This brings a total of 67 atom pairs into contact between dimers. As described at low resolution, the packing of dimers at this interface is similar to the packing between halves of tetrameric *Scapharca* HbII. It is interesting to note that only slightly over one-half as many interactions occur between the two dimers at this contact as occur within each dimer interface. Clearly, the earlier choice of the functional dimer is the only tenable choice.

An overall view of the dimer assemblage can be seen in the α-carbon plots shown in Fig. 7. Both hydrophilic and hydrophobic regions participate in the contact. The contact can be thought of as involving four unique regions. The "Heme-F" region is highly hydrophilic and is formed from both heme groups and the last two turns of both F helices. The "E–E" region forms where the center of the E helices in both subunits cross each other. Two symmetrically related "E–F" hydrophobic patches form where the E helix of one subunit lies across the F helix of the other. There are also two symmetrically related "B–F" salt bridges. Table III and Fig. 8 indicate those pairs of residues with one or more atoms within 4.2 Å across the interface.

The hydrophilic, locally charged surface around the heme propionate groups is exposed to solvent in monomeric hemoglobins and mammalian tetrameric hemoglobins. However, in *Scapharca* HbI this region is directly involved in the contact between subunits. The involvement of this region necessitates extensive hydrophilic character in the Heme-F region of the interface. This region is illustrated in Fig. 7*A*. The propionate carboxylates are brought in a position to form ionic interactions (2.5/2.4 Å) with lysine 96′ and hydrogen bonds (2.9/3.0 Å) with asparagine 100′ from the last two turns of the F helix (primes indicate symmetrically related subunit).

The Heme-F region is highly hydrated with a number of water molecules involved in bridged hydrogen bonds across the interface. One of these bridges passes through a water molecule that is hydrogen-bonded simultaneously to propionate oxygen atoms from both hemes. In our model, including the whole interface, six water molecules are directly hydrogen-bonded to residues from both subunits. These 6 are involved in 11 bridges across the interface, 5 of which terminate at propionate oxygen atoms. If paths through 2 water molecules



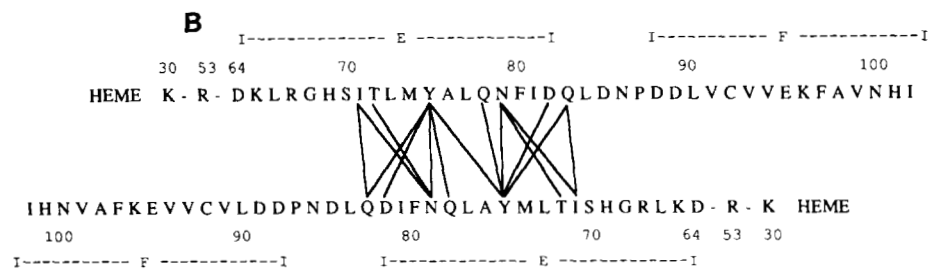
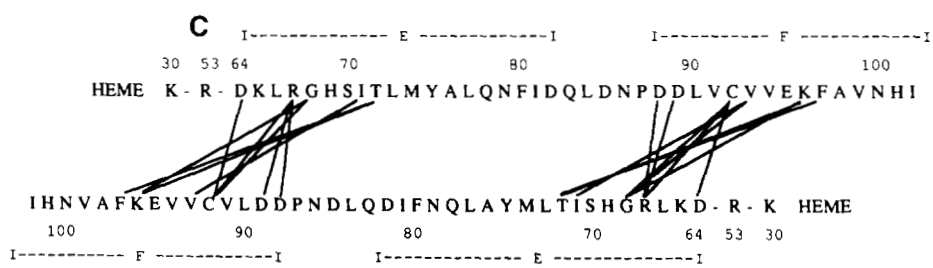


FIG. 8. **Residues within 4.2 Å across the dimer interface.** The amino acid sequence (Petruzzelli *et al.*, 1985) for each subunit is shown using the single-letter code with the two sequences running in opposite directions. The residue number and helical designations are shown above and below the sequences for subunit 1 and 2, respectively. *A*, Heme-F and B–F interactions. *B*, E–E interactions. *C*, E–F interactions.



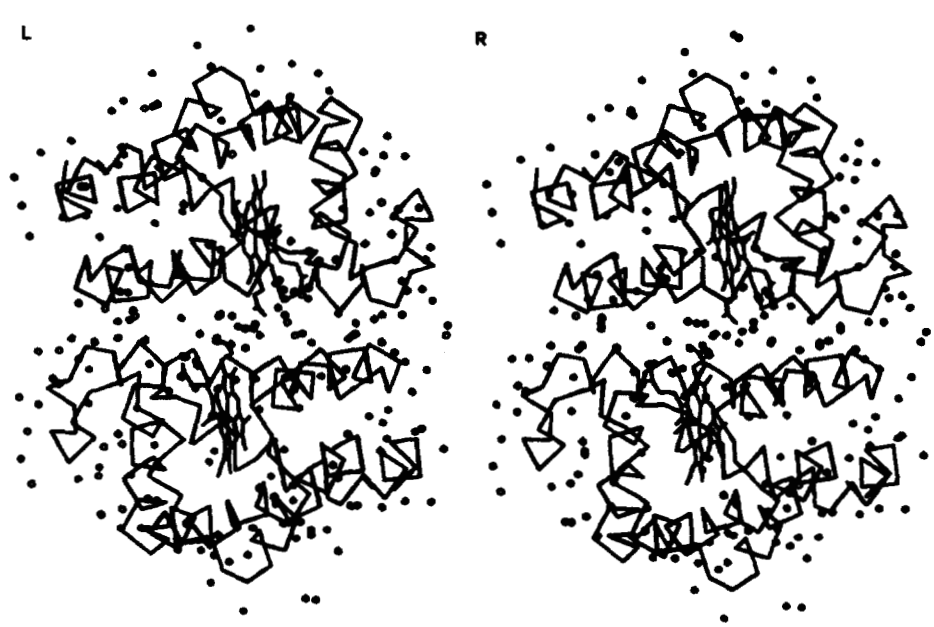


FIG. 9. **Stereo diagram of 214 ordered water molecules in *Scapharca* HbI.** *Model*, water molecules are plotted as *circles* along with the α-carbon and heme plot for HbI. Note particularly the large number of water molecules in the vicinity of the propionate groups in the Heme-F region of the contact.

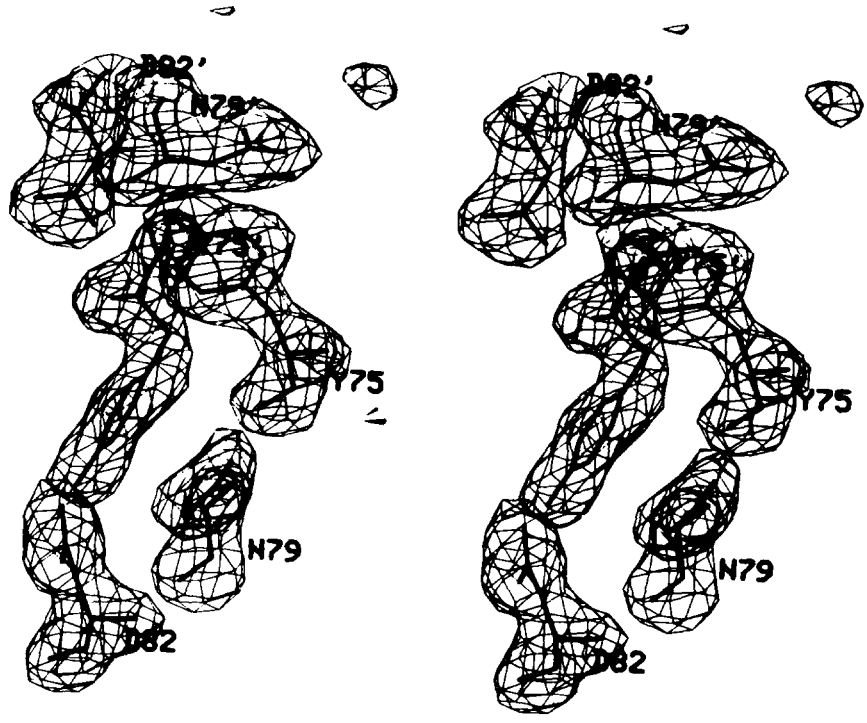


FIG. 10. **Fragment difference electron density map showing the E–E region of the interface.** The atoms for residues 75, 79, and 82 (from each subunit) were not used in the calculation of the phases for this electron density map. The map is contoured at 0.3 e/Å^3. Only those bonds between atoms that were omitted in the phase calculation are shown. Residues are labeled near their carbonyl carbon with unprimed residue numbers corresponding to one subunit and primed residue numbers corresponding to the second subunit. Note the hydrogen bonds between the hydroxyl groups of tyrosines 75 and 75′ with the carboxylate oxygens of aspartates 82′ and 82 and the stacking of the aromatic rings of 75 and 75′ with the Cα, Cβ, and Cγ of asparagines 79′ and 79. The map has a sphere of density in the *upper right* that could be due to a partially ordered water molecule that has yet to be modeled.

are considered, 18 bridges are present. In all, 22 water molecules are used to form these 29 hydrogen-bonded bridges across the interface. Fig. 9 shows the positions of all 214 water molecules in our model.

The E–E patch near the center of the contact is formed where the side chain of tyrosine 75 from each subunit points into the interface as shown in Figs. 7*B* and 10. The closest approach between these side chains is 4.6 Å (although the carbonyl oxygen of 75 is 3.9 Å from 75′ Cδ2). Significant interactions occur between tyrosine 75 and asparagine 79′. Part of the aromatic ring stacks up against the Cα, Cβ, and Cγ of the asparagine group with 8 pairs of atoms within 4.0 Å. The hydroxyl group of tyrosine 75 is in a position to form a hydrogen bond across the interface with aspartate 82′. Residues 73–75 adopt a 3_{10} conformation in the middle of the otherwise α-helical E helix in each subunit. This conformation appears to be necessary to accommodate the packing of tyrosine 75 in the interface. Tyrosine 75 and aspargine 79′ also have atoms within 4.2 Å of atoms from glutamine 78′, glutamine 83′, isoleucine 71, and threonine 72 (see Fig. 8 and Table III).

The E–F patches include the most extensive hydrophobic regions of the interface. Involved in these regions are glycine 68, isoleucine 71, threonine 72, cysteine 92′, valine 93′, and phenylalanine 97′ as shown in Fig. 7*B*. The methyl carbon of threonine 72 is 3.5 Å (3.2 Å) from Cζ of phenylalanine 97′ and is the closest approach of atoms in this region. The presence of a glycine at position 68 is essential to accommodate cysteine 92′ whose sulfur is 3.6 Å (3.6 Å) from the glycine Cα atom. The residues in this region are packed tightly enough to exclude solvent, but not many atoms are brought into close van der Waals contact. Additional residues peripherally in the E–F region are aspartate 64, arginine 67, aspartate 88′, aspartate 89′, and lysine 96′ (see Fig. 8 and Table III).

Two symmetrical salt bridges are formed between lysine 30 at the beginning of the B helix and aspartate 89′ in the F helix to form the B–F region as can be seen in Fig. 7*A*. The distances between lysine 30Nζ and aspartate 89′Oδ1 are 2.9 and 2.7 Å. A comparison of amino acid sequences of arcid hemoglobins reveals that in the known dimeric sequences the residues at position 30 and 89 are identical (Petruzzelli *et al.*, 1985). Curiously, however, in tetrameric arcid hemoglobins the analogous position for 30 is occupied by a glutamate and that for 89 is occupied by an arginine (Gilbert and Thompson, 1985). This would allow a similar salt bridge to form, but with a sign reversal. This may play a role in the inability of tetrameric subunits to hybridize with dimeric subunits.

DISCUSSION

A primary goal of our crystallographic studies on *Scapharca* HbI is to gain insight into the basis for the cooperative oxygen binding expressed in this simple molecule. Given that HbI has a lower affinity for oxygen than monomeric hemoglobins, the fully deoxygenated dimer presumably holds each heme group in a low affinity state. Upon binding of oxygen to one subunit, a structural change must occur and be transmitted to the oxygen binding regions of the second subunit. While a description of the structural principles for cooperativity requires a deoxy structure, the liganded structure that we have determined offers several clues as to how cooperative oxygen binding might occur.

Communication between heme groups requires a pathway for information to be transferred from one heme to the other. In mammalian hemoglobins, this pathway is through the $\alpha_1\beta_2$ interface involving the FG corner and the C helix (Baldwin and Chothia, 1979). In *Scapharca* dimeric hemoglobin, a shorter pathway is available through the Heme-F region between iron atoms that are 18.4 Å apart. The two heme groups are almost in direct contact with each other; propionate oxygen atoms are 4.0 Å apart and a water molecule is in a position for simultaneous hydrogen bonding to both oxygen atoms. More extensive intersubunit interactions of the propionate carboxylates are with lysine 96 and asparagine 100. Both of these are within 1-½ helical turns from the proximal histidine 101 in the F helix. Comparisons of liganded and unliganded structures of hemoglobins and myoglobins show significant intrasubunit differences in the F helix, presumably initiated by movement of the proximal histidine upon binding

of the ligand at the distal side of the heme (Baldwin and Chothia, 1979; Phillips, 1980). Since similar movements may occur in arcid clam hemoglobins, information about the ligand state of one heme could readily be communicated to the second heme through the Heme-F region.

Studies of the salt dependence of stability and oxygen binding properties of *Scapharca* hemoglobins has led to the conclusion that hydrophobic interactions play a primary role (Chiancone *et al.*, 1981; Gattoni *et al.*, 1983). Our model has definite hydrophobic regions in the intersubunit contact. One such region is the interaction between the aromatic ring of tyrosine 75 and the Cα, Cβ, and Cγ of asparagine 79 of the neighboring subunit. More extensive regions are found in the two symmetrically related E-F patches. These regions are tight enough to exclude solvent but do not bring many atoms into close contact. If the side chains were brought into a tighter packing arrangement, a more stable assemblage could be formed by increasing van der Waals interactions. One might expect the deoxy assemblage to be more stable than the liganded assemblage to provide the energy needed to clamp the oxygen binding region in a low affinity state. Indeed for the case of human hemoglobin, Ackers (1980) has reviewed thermodynamic experiments from his laboratory which show that, under their conditions, the deoxygenated tetramer is more stable than the oxygenated tetramer by 6.3 kcal/mol. Evidence that the E-F regions have a different packing arrangement in the deoxy and oxy states comes from experiments on the rate of the reaction of cysteine 92 with thiol reagents which was found to depend upon the state of oxygenation (Furuta *et al.*, 1980).[2] A difference in reactivity would be expected if the packing arrangement of residues near cysteine 92 were different.

In addition to cooperative oxygen binding, communication between subunits is evident in reactions that are cooperative in *Scapharca* HbI, but not in vertebrate hemoglobins. Thus, the cleavage of the bond between the proximal histidine Nϵ atom and the heme iron at low pH cannot be described by a simple protonation curve in *Scapharca* HbI in either the NO derivative (Spagnuolo *et al.*, 1986) or in the deoxy state.[3] The cleavage of the bond between the proximal histidine and the heme would certainly be accompanied by a displacement of the F helix and information about this alteration could be transmitted to the second subunit through the Heme-F region. Moreover, the binding of *p*-chloromercuribenzoate to cysteine 92 has also been found to be cooperative in HbI (Boffi *et al.*, 1987). The binding to one cysteine must alter the interface so much that cysteine 92′ in the symmetrically related E-F region becomes much more accessible. In addition, this alteration at the interface results in an increased oxygen affinity and marked decrease in cooperative oxygen binding. These phenomena demonstrate the far reaching effects of alterations in a small portion of the contact which may result from the relative rigidity of the E and F helices as indicated by their low *B*-values (see Fig. 3).

The dimeric hemoglobins from the arcid clams are elegantly simple systems for investigating allosteric interactions. Our model of the carbon monoxide-liganded structure of *Scapharca* HbI provides a basis for understanding the structural details about the subunit interactions leading to cooperative oxygen binding. However, to fully understand the structural basis for allostery, a structure of the molecule in the fully deoxygenated state is essential. We have recently grown crystals of *Scapharca* HbI under deoxygenated conditions. These crystals are not isomorphous to the liganded crystals used in this study and we are pursuing their structure.

Acknowledgments—We thank Dr. W. E. Love for his interest in this project, Dr. M. Bolognesi for sharing coordinates of *Aplysia* myoglobin prior to their deposition in the Protein Data Bank and Drs. A. Pahler, J. L. Smith, S. Sheriff, and C. M. Cambillau for computer programs.

[2] D. Verzili, G. Colotti, M. Coletta, and E. Chiancone, unpublished data.

[3] M. Coletta, A. Boffi, P. Ascenzi, M. Brunori, and E. Chiancone, submitted for publication.

REFERENCES

Acharya, R., Fry, E., Stuart, D., Fox, G., Rowlands, D., and Brown, F. (1989) *Nature* **337,** 709–716
Ackers, G. K. (1980) *Biophys. J.* **32,** 331–346
Arnold, E., and Rossmann, M. G. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 5489–5493
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) *J. Mol. Biol.* **112,** 535–542
Baldwin, J., and Chothia, C. (1979) *J. Mol. Biol.* **129,** 175–220
Boffi, A., Gattoni, M., Santucci, R., Vecchini, P., Ascoli, F., and Chiancone, E. (1987) *Biochem. J.* **241,** 499–504
Bolognesi, M., Coda, A., Gatti, G., Ascenzi, P., and Brunori, M. (1985) *J. Mol. Biol.* **183,** 113–115
Borgese, T. A., Harrington, J. P., Hoffman, D., San George, R. C., and Nagel, R. L. (1987) *Comp. Biochem. Physiol.* **86B,** 155–165
Chiancone, E., Vecchini, P., Verzili, D., Ascoli, F., and Antonini, E. (1981) *J. Mol. Biol.* **152,** 577–592
Como, P. F., and Thompson, E. O. P. (1980) *Aust. J. Biol. Sci.* **33,** 643–652
Furuta, H., Ohe, M., and Kajita, A. (1977) *J. Biochem. (Tokyo)* **82,** 1723–1730
Furuta, H., Ohe, M., and Kajita, A. (1980) *Biochim. Biophys. Acta* **625,** 318–327
Gattoni, M., Verzili, D., Chiancone, E., and Antonini, E. (1983) *Biochim. Biophys. Acta* **743,** 180–185
Gaykema, W. P. J., Hol, W. G. J., Vereijken, J. M., Soeter, N. M., Bak, H. J., and Beintema, J. J. (1984) *Nature* **309,** 23–29
Gilbert, A. T., and Thompson, E. O. P. (1985) *Aust. J. Biol. Sci.* **38,** 221–236
Hanson, J. C., Watenpaugh, K. D., Sieker, L., and Jensen, L. H. (1979) *Acta Crystallogr.* **A35,** 616–621
Hendrickson, W. A. (1976) *J. Mol. Biol.* **106,** 889–893
Hendrickson, W. A. (1985) *Methods Enzymol.* **115,** 252–270
Hendrickson, W. A., and Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond, R., Ramaseshan, S., and Venkatean, K., eds) pp. 13.01–13.23, Indian Academy of Sciences, Bangalore
Hendrickson, W. A., and Lattman, E. E. (1970) *Acta Crystallogr.* **B26,** 136–143
Hendrickson, W. A., and Teeter, M. M. (1981) *Nature* **290,** 107–113
Hogle, J. M., Chow, M., and Filman, D. J. (1985) *Science* **229,** 1358–1365
Jones, T. A. (1985) *Methods Enzymol.* **115,** 157–171
Konnert, J. (1976) *Acta Crystallogr.* **A32,** 614–617
Luo, M., Vriend, G., Kamer, G., Minor, I., Arnold, E., Rossmann, M. G., Boege, U., Scraba, D. G., Duke, G. M., and Palmenberg, A. C. (1987) *Science* **235,** 182–191
Luzzati, P. V. (1952) *Acta Crystallogr.* **5,** 802–810
North, A. C. T., Phillips, D. C., and Mathews, F. S. (1968) *Acta Crystallogr.* **A24,** 351–359
Ohnoki, S., Mitomi, Y., Hata, R., and Satake, K. (1973) *J. Biochem.* **73,** 717–725
Perutz, M. F., Fermi, G., Luisi, B., Shaanan, B., and Liddington, R. C. (1987) *Acc. Chem. Res.* **20,** 309–321
Petruzzelli, R., Goffredo, B. M., Barra, D., Bossa, F., Boffi, A., Verzili, D., Ascoli, F., and Chiancone, E. (1985) *FEBS Lett.* **184,** 328–332
Phillips, S. E. V. (1980) *J. Mol. Biol.* **142,** 531–554
Rossmann, M. G. (1961) *Acta Crystallogr.* **14,** 383–388
Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H.-J., Johnson, J. E., Kamer, G., Luo, M., Mosser, A. G., Ruekert, R. R., Sherry, B., and Vriend, G. (1985) *Nature* **317,** 145–153
Royer, W. E., Jr., Love, W. E., and Fenderson, F. F. (1985) *Nature* **316,** 277–280
San George, R. C., and Nagel, R. L. (1985) *J. Biol. Chem.* **260,** 4331–4337
Schachman, H. K. (1988) *J. Biol. Chem.* **263,** 18583–18586
Sheriff, S., Hendrickson, W. A., and Smith, J. L. (1987) *J. Mol. Biol.* **197,** 273–296
Smith, J. L., and Hendrickson, W. A. (1982) in *Computational Crystallography* (Sayre, D., ed) pp. 209–222, Oxford, New York
Smith, J. L., Hendrickson, W. A., and Addison, A. W. (1983) *Nature* **303,** 86–88
Spagnuolo, C., Desideri, A., Chiancone, E., and Ascoli, F. (1986) in *Invertebrate Oxygen Carriers* (Linzen, B., ed) pp. 101–105, Springer-Verlag, Berlin
Wallace, B. A., and Ravikumar, K. (1988) *Science* **241,** 182–187
Wilson, A. J. C. (1942) *Nature* **150,** 152