Protein Science (1993), *2*, 597-619. Cambridge University Press. Printed in the USA.

Crystal structure of deoxygenated *Limulus polyphemus* subunit II hemocyanin at 2.18 Å resolution: Clues for a mechanism for allosteric regulation

BART HAZES,[1] KAREN A. MAGNUS,[2] CELIA BONAVENTURA,[3]
JOSEPH BONAVENTURA,[3] ZBIGNIEW DAUTER,[4]
KOR H. KALK,[1] AND WIM G.J. HOL[1]

[1] BIOSON Research Institute, Department of Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
[2] Department of Biochemistry, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106
[3] Marine Biomedical Center, Duke University Marine Laboratory, Beaufort, North Carolina 28516
[4] EMBL % DESY, Notkestrasse 85, 2000 Hamburg 52, Germany

(RECEIVED September 8, 1992; REVISED MANUSCRIPT RECEIVED December 14, 1992)

Abstract

The crystal structure of *Limulus polyphemus* subunit type II hemocyanin in the deoxygenated state has been determined to a resolution of 2.18 Å. Phase information for this first structure of a cheliceratan hemocyanin was obtained by molecular replacement using the crustacean hemocyanin structure of *Panulirus interruptus*. The most striking observation in the *Limulus* structure is the unexpectedly large distance of 4.6 Å between both copper ions in the oxygen-binding site. Each copper has approximate trigonal planar coordination by three histidine N^{ϵ} atoms. No bridging ligand between the copper ions could be detected. Other important new discoveries are (1) the presence of a cis-peptide bond between Glu 309 and Ser 310, with the carbonyl oxygen of the peptide plane hydrogen bonded to the N^{δ} atom of the copper B ligand His 324; (2) localization of a chloride-binding site in the interface between the first and second domain; (3) localization of a putative calcium-binding site in the third domain. Furthermore, comparison of *Limulus* versus *Panulirus* hemocyanin revealed considerable tertiary and quaternary rigid body movements, although the overall folds are similar. Within the subunit, the first domain is rotated by about 7.5° with respect to the other two domains, whereas within the hexamer the major movement is a 3.1° rotation of the trimers with respect to each other. The rigid body rotation of the first domain suggests a structural mechanism for the allosteric regulation by chloride ions and probably causes the cooperative transition of the hexamer between low and high oxygen affinity states. In this postulated mechanism, the fully conserved Phe 49 is the key residue that couples conformational changes of the dinuclear copper site into movements of the first domain.

Keywords: allosteric regulation; cooperativity; dinuclear copper site; hemocyanin; oxygen transport; X-ray structure

Hemocyanins are the oxygen transport proteins in many molluscs and arthropods. Their oxygen-binding sites are formed by two copper ions liganded directly to the protein. Hemocyanins do not form a homogeneous class of proteins, since there are large differences between arthropodan and molluscan hemocyanins. The sequence homology between both classes is actually limited to only one region, which is known to bind the second copper ion, Cu B, in the arthropodan hemocyanins (Volbeda & Hol, 1989b; Lang & van Holde, 1991). Also the quaternary structures are very different, as is shown by electron microscopy (van Bruggen et al., 1982; van Holde & Miller, 1982). The molluscan hemocyanins form huge cylindrical complexes, with molecular masses of up to 9 MDa. Each cylinder is built from 10 or 20 identical subunits containing 7 or 8 oxygen-binding centers per subunit (van Holde & Miller, 1982). The arthropodan hemocyanins, in contrast, form hexameric or multi-hexameric complexes, with 1, 2, 4, 6,

Reprint requests to: Wim G.J. Hol, Biological Structure, SM-20, School of Medicine, University of Washington, Seattle, Washington 98195.

Abbreviations: EXAFS, extended X-ray absorption fine structure spectroscopy; EPR, electron paramagnetic resonance; rms, root mean square; MD, molecular dynamics.

or 8 hexamers (van Holde & Miller, 1982; Gaykema et al., 1984). Within such a complex, up to eight different, although homologous, subunit types may be found. Each subunit type has a molecular weight of about 75 kDa and reversibly binds one molecule of oxygen.

In this article the structure of a hemocyanin subunit of the arthropod *Limulus polyphemus* (a horseshoe crab) is described. *Limulus polyphemus* has the most complex arthropodan hemocyanin known. It consists of eight hexamers and has eight immunologically distinct subunit types, named I, II, IIA, IIIA, IIIB, IV, V, and VI (Brenowitz et al., 1981). Each subunit type is functionally different (Sullivan et al., 1974; Brenowitz et al., 1984) and occupies a specific position in the 48-meric complex (Lamy et al., 1983). The oxygen binding of the whole complex is cooperative and is further regulated by several allosteric effectors, of which the most important are protons, calcium, and chloride ions (Brouwer et al., 1977, 1982, 1983; Brouwer & Serigstad, 1989).

Much more simple single hexameric *Limulus* hemocyanin complexes can also be obtained from several purified subunit types. Purified subunit type II hemocyanin (abbreviated as *Limulus* II), which is the object of study here, can form such a single-hexameric structure (Brenowitz et al., 1984). These *Limulus* II homohexamers retain several of the oxygen-binding and regulatory properties of the native complex, e.g., chloride reduces oxygen affinity, protons increase oxygen affinity, and oxygen can bind cooperatively (Sullivan et al., 1974; Brenowitz et al., 1984).

In the oxygen-binding site of hemocyanin, oxygen is thought to bind as a bridging peroxide between two copper ions, whose valence state then changes from Cu(I) to Cu(II) (Freedman et al., 1976). Extensive spectroscopic investigations have indicated that in oxygenated hemocyanin the cupric ions are separated by 3.6 Å (Brown et al., 1980; Woolery et al., 1984) and that the two oxygen atoms of the bridging peroxide reside in a very similar environment (Thamann et al., 1977). In addition, the cupric ions must be strongly antiferromagnetically coupled by the bound peroxide, since oxygenated hemocyanins are EPR silent (Dooley et al., 1978). Based on these data several models for the oxygenated dinuclear copper site in hemocyanin have been proposed (Eickman et al., 1979; Maddaluno & Giessner-Prettre, 1991; Ross & Solomon, 1991).

Because cuprous ions are practically transparent to spectroscopic techniques, less information on the dinuclear copper site in the deoxygenated state is available. Some insight concerning the immediate environment of the coppers has, however, been obtained by EXAFS measurements. These studies indicated that in deoxygenated hemocyanin, both copper ions are most likely separated by 3.4 Å, although in one study no copper–copper interactions within 4.0 Å could be detected (Brown et al., 1980; Co & Hodgson, 1981; Woolery et al., 1984).

Arthropodan hemocyanins have been found in the Crustacea (crabs, lobsters, and shrimps), the Chelicerata (spiders, scorpions, and horseshoe crabs), and the Myriapoda (centipedes). In the other major group of arthropods, the insects, no hemocyanin has been found so far. However, in some insects hexameric protein complexes dissolved in the hemolymph have been observed. Amino acid sequence information for this class of proteins, generally referred to as insect storage proteins, clearly shows a relation with the hemocyanins (for references see the legend of Fig. 1). These molecules do not transport oxygen, but function as amino acid stores rich in aromatic residues or methionine. Accordingly, the class of insect storage proteins provides us with an elegant method for investigating whether a residue is conserved for structural

Fig. 1. Sequence alignment based on structural superposition of *Limulus* II and *Panulirus* A-chain hemocyanin. The sequences of these two proteins are printed in bold type. The other sequences are aligned manually. Sequences 1–4 represent crustacean hemocyanins, sequences 5–10 represent cheliceratan hemocyanins, and sequences 11–17 represent insect storage proteins. Sequence positions that are fully conserved in one of the three classes are boxed and stippled. The copper ligands—His 173, His 177, His 204, His 324, His 328, and His 364—and the functionally important Phe 49 and Glu 309 are printed in white on a black background. Line headings: SN PINT A = secondary structure nomenclature according to Gaykema et al. (1984); SS PINT A = position of helices (A) and β-strands (B) in the *Panulirus* structure; NR PINT A = sequence numbering for *Panulirus* hemocyanin; PINT C = *Panulirus interruptus* C-chain (Neuteboom et al., 1992); ALEP B = *Astacus leptodactylus* B-chain (Schneider et al., 1986); PVUL B = *Palinurus vulgaris* B-chain (Neuteboom et al., 1989); PINT A = *Panulirus interruptus* A-chain (Bak & Beintema, 1987); EQUIVALENT = positions that are structurally equivalent in the *Limulus* II and the *Panulirus* structure are indicated by "="; LIM II = *Limulus polyphemus* subunit II (Nakashima et al., 1986); EURY A = *Eurypelma californicum* A-chain (Voit & Feldmaier-Fuchs, 1990); EURY B = *Eurypelma californicum* B-chain (Sonner et al., 1990); EURY D = *Eurypelma californicum* D-chain (Linzen et al., 1985); EURY E = *Eurypelma californicum* E-chain (Voit & Feldmaier-Fuchs, 1990); TACHY A = *Tachypleus tridentatus* A-chain (Nemoto & Takagi, 1983); MSEXTA A = *Manduca sexta* A-chain (Willott et al., 1989); MSEXTA B = *Manduca sexta* B-chain (Willott et al., 1989); BMORI SP1 = *Bombyx mori* SP1-chain (Sakurai et al., 1988); BMORI SP2 = *Bombyx mori* SP2-chain (Fujii et al., 1989); TNI AJSP1 = *Trichoplusia ni* AJSP1-chain (Jones et al., 1990); CVIC = *Calliphora vicina* (Naumann & Scheller, 1991); DMEL SP2 = *Drosophila melanogaster* SP2-chain (Mousseron, 1986); NR LIM II = residue numbering for the *Limulus* II sequence; SS LIM II = position of helices (A) and β-strands (B) in the *Limulus* II structure. Special characters: "?" = not determined or ambiguous; "-" = deletion; "|" = domain boundary; "x" = structure of *Limulus* II unknown at this position. (*Figure appears on the three pages that follow.*)

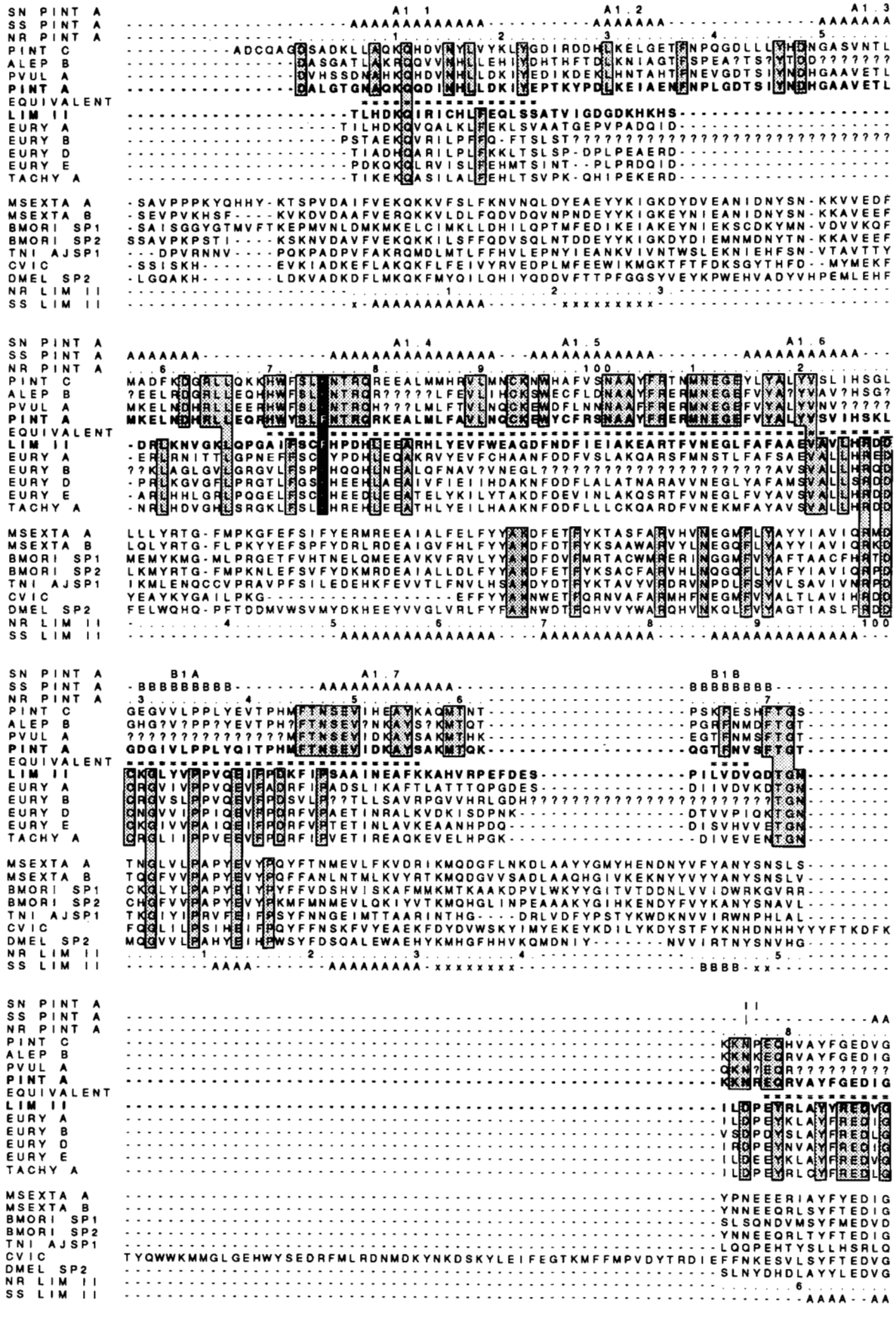

Fig. 1. *See caption on facing page.*

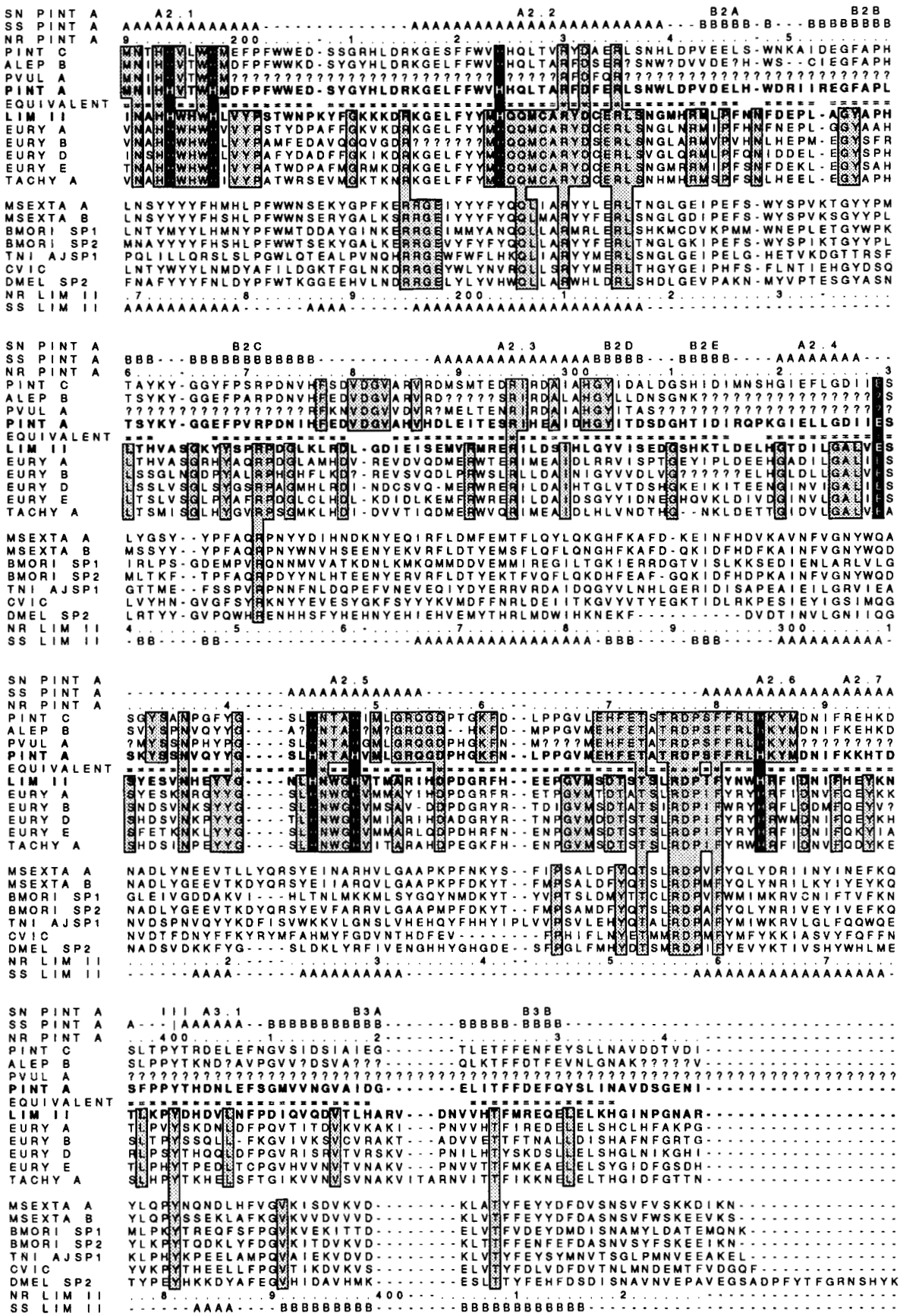

Fig. 1. *Continued.*

Fig. 1. *Continued.*

or functional reasons. Structurally important residues are likely to be conserved in all sequences, whereas functionally important residues will be conserved within the hemocyanins only. The alignment of 10 arthropodan hemocyanin sequences and 7 insect storage protein sequences is given in Figure 1.

At this moment two arthropodan hemocyanin structures are known, that of deoxygenated *Panulirus interruptus* hemocyanin (Gaykema et al., 1984, 1986; Volbeda & Hol, 1989a) and the deoxygenated *Limulus* II structure presented in this paper. These structures represent a crustacean and a cheliceratan hemocyanin, respectively. Hemocyanins from both classes show sequence identities around 33% and their subunits form similar hexameric complexes (Magnus et al., 1991). Electron microscopy has, however, shown that aggregation of quaternary structures beyond the hexameric level is basically different (de Haas et al., 1991). On the sequence level a striking difference is a 21-residue deletion near the N-terminus in the cheliceratan hemocyanins. It has been suggested that this 21-residue deletion is one of the factors causing the difference in the multi-hexameric quaternary structure (Soeter et al., 1987).

The *Limulus* II crystals give specific information on cheliceratan hemocyanins, but in addition they offer a number of advantages over the *Panulirus* crystals. First of all the crystals diffract to a much higher resolution (2.18 Å, compared to 3.2 Å for *Panulirus*). Secondly, the crystals grow at a pH of about 6.5–7.0. This is only slightly below the physiological pH value, which varies between 7.1 and 7.55 (Johansen & Petersen, 1975). *Panulirus* crystals in contrast were grown at pH 3.8. Accordingly, the *Limulus* II structure allows us for the first time to study an arthropodan hemocyanin in detail under near physiological conditions.

Results

Quality of the model

The *R*-value for the final model, including 345 water molecules, 2 copper ions, 1 chloride ion, and 1 calcium ion, is 17.4% for all reflections between 10.0 and 2.18 Å. The rms deviations for bond lengths and bond angles are 0.014 Å and 2.9°, respectively, and based on the PROCHECK program (Morris et al., 1992) the geometry of the protein is significantly better then the average PDB structure with this resolution. The average *B*-values for protein and water atoms are 20.0 Å^2 and 31.5 Å^2, respectively and the estimated coordinate error as given by SIGMAA (Read, 1986) is 0.19 Å.

There are, however, a few problematic regions. These regions are probably highly flexible and contain residues: 1, 21–29, 132–139, 148–149, 527–530, and 569–572. (All residue numbers in this article refer to the *Limulus* II sequence numbering. Figure 1 can be used to obtain the corresponding residue numbers in the *Panulirus* structure.) All these stretches are located at the surface of the protein. Because it was not possible to trace the chain unambiguously in these regions, we have not included these parts in the model. Also a number of side chains were not included in the model because there was no density for them. The residues with missing side chains are Thr 20, Glu 75, Glu 140, Ser 141, Pro 142, Asp 146, Val 147, Lys 187, Glu 412, Arg 427, Leu 505, Lys 550, Ser 573, His 627, and Glu 628. All missing atoms, except the Lys 187 side chain, are located in domains one and three. These two domains also have a higher average *B*-value as can be seen in Figure 2.

The phi-psi plot, given in Figure 3, shows a few residues outside the allowed regions (Ramakrishnan & Ramachandran, 1965). The major outliers, Ser 31 and His 522, are close to flexible regions in the structure and their conformation is likely not completely correct. Asp 390 is located in a sharp turn at the surface of the protein and has only weak density for the C_α and C_β atoms. The density for the neighboring residues is however very clear.

From the residues that are only slightly outside the allowed phi-psi regions, Phe 49, Phe 83, and Pro 591 are well centered in density and probably are an accurate representation of the protein structure. In this respect it is interesting to note that we propose Phe 49 to be a key residue in regulation of oxygen affinity and cooperativity as will be discussed below. The other residues with slightly unfavorable phi-psi values are less well centered in density, but inspection did not indicate significant errors.

Based on the electron density, we have changed the amino acid sequence, obtained by protein sequencing (Nakashima et al., 1986), at two positions:

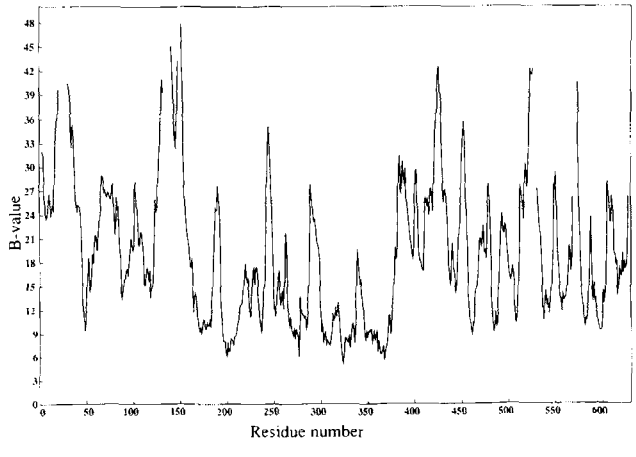

Fig. 2. Average main-chain *B*-value per residue. The gaps in the curve correspond to amino acid residues that were not determined in the final structure. Note that the average *B*-value for domain 2 (155–380) is significantly lower than in domains 1 (1–154) and 3 (381–628). (Domains 1, 2, and 3 have average-main chain *B*-values of 25.0, 12.8, and 20.7 Å^2, respectively.)

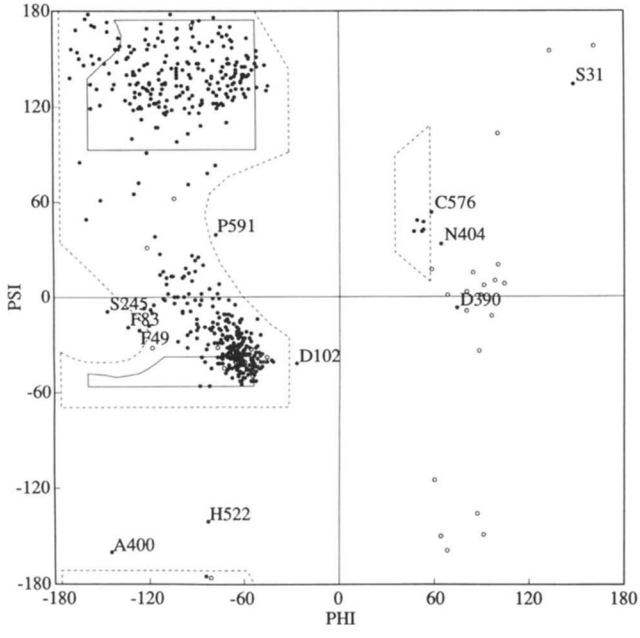

Fig. 3. Phi–psi plot of the final model. Allowed regions are indicated by broken lines. Open circles represent glycine residues. Filled circles represent the non-glycine residues. The labeled residues are discussed in the text.

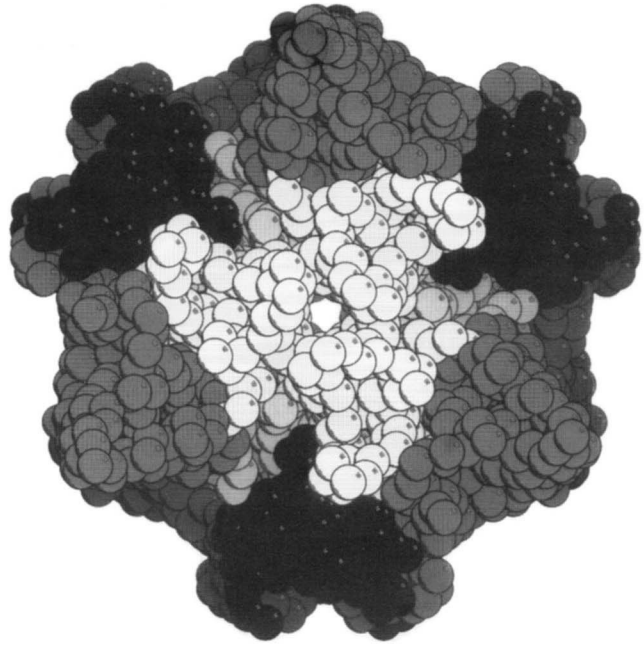

Fig. 4. CPK model of the hexamer of *Limulus* II hemocyanin. Each residue is represented by a sphere with a radius of 3.5 Å centered at the C_α position. Domains 1, 2, and 3 are depicted in black, pale gray, and dark gray, respectively. The three types of intersubunit contacts are more clearly shown in Figure 14. This figure and Figures 5, 13, and 14 were prepared with MOLSCRIPT (Kraulis, 1991).

1. Val 9 → Ile 9: Strong additional density connected to one of the C^γ atoms of Val 9 was observed. The density was located in a hydrophobic cavity, excluding the possible presence of a buried ion or water molecule. The density could however be fitted reasonably with an isoleucine residue. The *B*-value of the added C^δ atom refined to a final value of 28 Å^2.
2. Phe 408–Thr 409 → Thr 408–Phe 409: The refinement of residues 400–427 was problematic until it was discovered that the sequence alignment had to be modified by introducing a four-residue insertion after position 400 and a three-residue deletion after residue 427 (compare Linzen et al., 1985, and Fig. 1). In the new alignment the density clearly indicated that the sequence of residues 408 and 409 had to be reversed. The sequence modification is also supported by the new sequence alignment, where Thr 408–Phe/Tyr 409 is now strictly conserved (see Fig. 1).

Description of the structure

The *Limulus* II crystals contain one hexamer in the rhombohedral unit cell. The hexamer is built from two trimers and has exact 32-point group symmetry. A CPK representation of the hexamer is shown in Figure 4. Within the hexamer each monomer is in contact with four other monomers, namely the monomers in the same trimer and two of the three monomers in the other trimer. The interfaces between these subunits are indicated as trimer, tight dimer, and loose dimer interfaces, analogous to the nomenclature used for the *Panulirus* structure (Volbeda & Hol, 1989a). Table 1 shows the amino acid contacts involved in the three different types of intersubunit interfaces of *Limulus* II and *Panulirus* hemocyanin. These interfaces will be described in more detail in the section discussing cooperativity.

The *Limulus* II subunit can be divided into three domains, which, analogous to the *Panulirus* domain definition, comprise residues 1–154, 155–380, and 381–628, respectively. This agrees reasonably well with the known gene structure of *Eurypelma californicum* subunit e hemocyanin (Voll & Voit, 1990), where introns were found after residues 65, **141**, 210, 342, **385**, 430, 492, and 568 (numbers in boldface are close to domain divisions). Schematic diagrams of the *Limulus* II subunit and the three separate domains are given in Figure 5A–D and Kinemages 1–4.

Domain 1

The first domain, shown in Figure 5B and Kinemage 2, is mainly α-helical, with one β-strand (B1B) at its C-terminal end. B1B interacts with β-strand B3C of the third domain. The loops just before and after B1B (132–139 and 148–149) showed uninterpretable density. These loops were therefore not included in the model and are probably

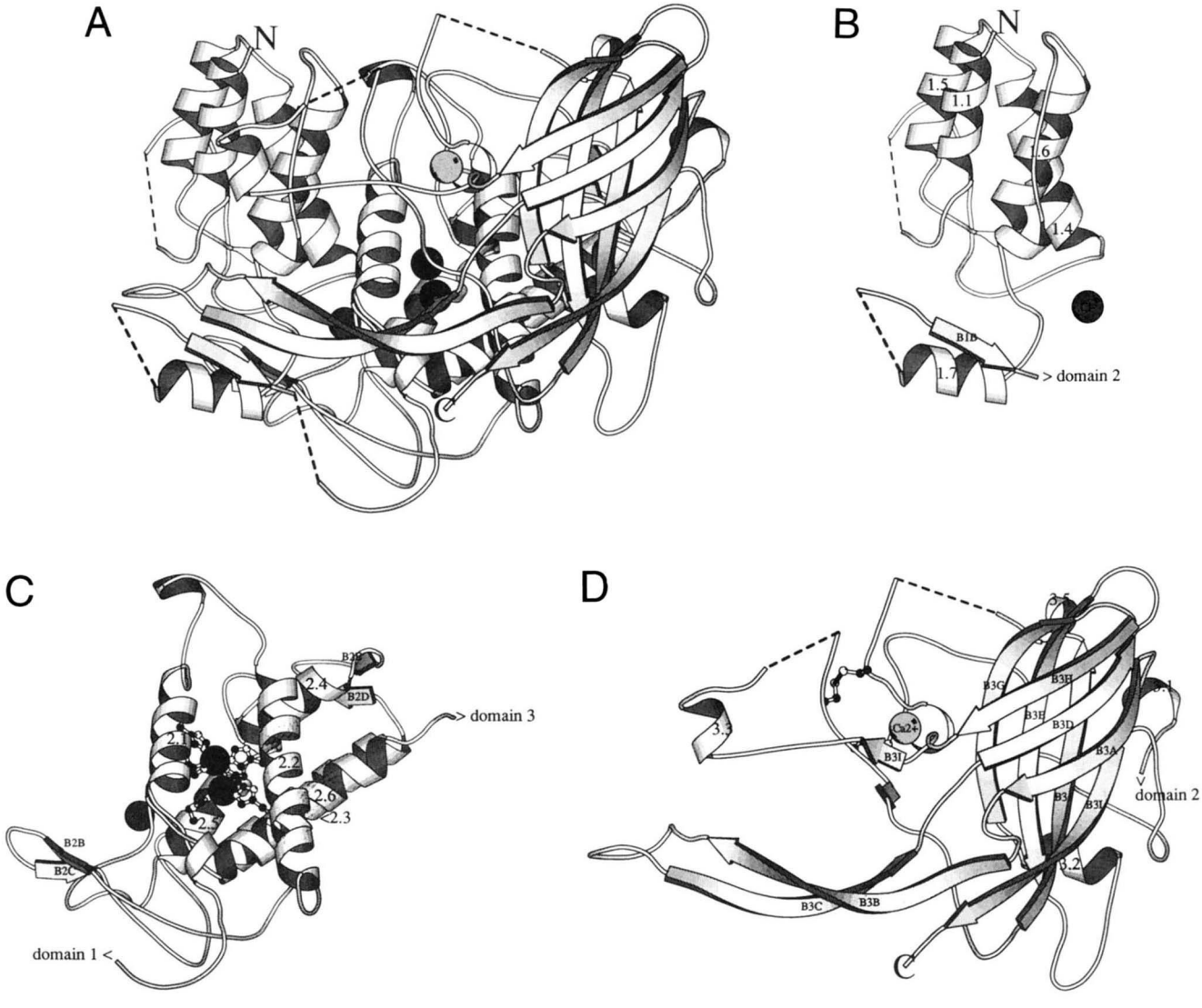

Fig. 5. Schematic diagrams of the *Limulus* II subunit and its domains. **A:** The whole subunit. **B:** Domain 1. **C:** Domain 2. **D:** Domain 3. All diagrams show the same view. Residues absent in the final structure are represented by dashed lines. The black spheres represent the coppers, the pale gray sphere represents the putative calcium ion, and the darker gray sphere represents the chloride ion. In the domain diagrams the secondary structure nomenclature according to Gaykema et al. (1984) has been included, and in domain 3 the disulfide bridges are indicated, although the disulfide bridge Cys 536–Cys 583 is partly hidden by β-strand B3I.

mobile. In the *Panulirus* structure the same region also showed weak density, indicating that flexibility of this region may be a conserved feature.

After helix 1.1, a 21-residue deletion has occurred in the cheliceratan hemocyanin sequences when compared to the crustacean hemocyanins and the insect storage proteins (see Fig. 1). As a result of this deletion, helix 1.2, helix 1.3, and the loop connecting these two helices as observed in the *Panulirus* structure are absent in *Limulus* II hemocyanin. Instead, helix 1.1 seems to be connected directly to helix 1.4 by a long loop containing 37 residues. Part of this loop is probably also very flexible, since residues 21–29 did not show interpretable density. These observations are in very good agreement with what had been predicted based on sequence comparisons (Soeter et al., 1987).

Domain 2

The second domain, shown in Figure 5C and Kinemage 4, is also mostly α-helical and is packed in between the first and the third domains. It contains the dinuclear copper site that is completely buried in its core. Each copper ion is liganded by three histidine side chains provided by a pair of helices. The ligands and ligand-donating helices for Cu A and Cu B are related by a pseudo-twofold axis, as has been described for *Panulirus* hemocyanin (Volbeda

Table 1. *Intersubunit contacts shorter than 3.5 Å*

Limulus II	*Panulirus*[a]
Trimer interface	
Glu 274 OE1-Arg 271 NE	Arg 250 NH-Asp 279 OD
Glu 274 OE2-Arg 271 NH1	Asp 301 O-Gln 338 NE
Asp 278 OD2-Tyr 319 OH	Asp 301 CB-Arg 295 NH
His 281 CD2-Glu 318 OE1	Asp 301 CG-Arg 295 NH
His 281 ND1-Arg 275 NH1	Asp 301 OD-Arg 295 NH
His 281 ND1-Arg 275 NH2	His 302 NE-Tyr 339 CE
Leu 282 CG-Tyr 319 CE1	His 302 NE-Tyr 339 CZ
Leu 282 CD2-Tyr 319 CD1	Arg 316 NH-Gln 338 OE
	Arg 634 NE-Arg 64 NE
	Arg 634 CZ-Asp 62 OD
Tight dimer interface	
Asn 126 OD1-Ser 245 O	Tyr 155 CB-Met 159 CE
Ala 152 O-His 342 CE1	Tyr 155 CG-Met 159 CE
Ala 152 O-His 342 NE2	Tyr 155 CD1-Met 159 CE
Leu 154 CD1-Asp 338 O	Tyr 155 CD2-Met 159 CE
Ala 237 CB-Glu 343 OE1	Ser 156 OG-Asp 438 OD1
His 239 CD2-Glu 343 OE1	Met 159 CB-Asp 438 O
His 239 CE1-Phe 341 O	Thr 160 CG-Gly 440 N
His 239 NE2-Phe 341 O	Arg 177 CZ-Phe 361 CE1
His 239 NE2-Glu 343 OE1	Arg 177 NH1-Phe 361 CE1
	Glu 254 OE-Lys 360 NZ
	Gly 255 O-Phe 361 CZ
	Glu 267 CB-Glu 267 OE
	Asp 273 OD2-Phe 361 CE2
	Asp 273 OD2-Phe 361 CZ
Loose dimer interface	
Ala 152 CA-Asn 472 OD1	Lys 175 CB-Asn 489 O
Ile 153 N-Asn 472 OD1	Lys 175 O-Ile 491 CD
Ile 153 O-Leu 470 CD1	
Ile 153 O-Asn 472 ND2	
Arg 159 NH2-Glu 469 O	

[a] Intersubunit contacts for *Panulirus* hemocyanin were taken from Table 7 of Volbeda and Hol (1989a). Only contacts present in at least four of the six noncrystallographically related subunit interfaces were included.

& Hol, 1989b). Besides the obvious functional role in oxygen binding, domain 2 is also important from a structural point of view, because it dominates the intersubunit contacts within the hexamer. The binding site for the functionally important chloride ion is located on the interface between domains 1 and 2 as will be described below. The proposed disulfide bridge between Cys 208 and Cys 213 (Eyerle & Schartau, 1985) was not observed.

Domain 3

The third domain, shown in Figure 5D and Kinemage 3, has a seven-stranded Greek key β-barrel topology, from which long loops extend. Two of these loops embrace domains 1 and 2 (see Fig. 5A,D). One loop forms a huge β-hairpin (strands B3B and B3C), which interacts with β-strand B1B from domain 1 as mentioned before. The other loop interacts with domain 1 via helix 3.3. Domain 3 contains two disulfide bridges (Cys 534–Cys 576 and Cys 536–Cys 583) and a putative calcium-binding site to be described later. Both the disulfide bridges and the calcium site are near two flexible regions (527–530 and 569–572) and are likely required for stability.

The β-barrel shows a strong structural relationship with the β-barrels observed in the immunoglobulins, Cu,Zn-superoxide dismutase, and the group of chromophore-carrying antitumor proteins. This structural similarity without a known functional or evolutionary relationship has led to a novel hypothesis regarding the protein-folding pathway of Greek key β-barrel proteins (Hazes & Hol, 1992).

Cu–Cu distance

The copper ions were inserted into our model after the *R*-value of all data to 2.2 Å dropped below 30%. By this time the density maxima for the copper ions had become the strongest features in the $F_o - F_c$ map. The distance between both copper ions was 4.6 Å. During refinement the Cu–Cu distance proved to be well determined by the density, as it remained in the range of 4.56–4.65 Å without any constraints other than the crystallographic data (see Materials and methods). The 4.6-Å distance is much longer than was expected from other studies. Based on EXAFS data (Brown et al., 1980; Woolery et al., 1984) the Cu–Cu distance in deoxygenated hemocyanin was expected to be 3.4 Å. Also in the reduced ascorbate oxidase structure (Messerschmidt et al., 1989), which has a copper site with spectral features similar to hemocyanin, a Cu–Cu distance of 3.4 ± 0.2 Å was observed. In the *Panulirus* structure (Volbeda & Hol, 1989a) the Cu–Cu distance was estimated to be 3.5 ± 0.3 Å. Again, this was significantly shorter than the 4.6 Å observed in the *Limulus* II structure.

Copper coordination

In Figure 6A and Kinemages 4 and 6 the dinuclear copper site is depicted viewed perpendicular to the Cu–Cu axis. This view shows that each copper is nearly coplanar with its three liganding N^{ε} atoms (see also Table 2). In Figure 6B the view parallel to the Cu–Cu axis is represented, showing an approximately antiprismatic orientation of the histidine ligands of Cu A with respect to the ligands of Cu B. This is in contrast to the dinuclear copper site of ascorbate oxidase where a prismatic orientation of the liganding histidines was observed (Messerschmidt et al., 1989). The coordination of Cu A and Cu B can be best described as distorted trigonal planar. Copper A most closely approaches an ideal trigonal coordination, with coordination angles in the range of 99–131° and ligand distances between 2.1 and 1.9 Å. For copper B we see a similar coordination, although it is somewhat more distorted. The coordination angles now range from 97 to 142°, with ligand distances between 2.2 and 1.9 Å.

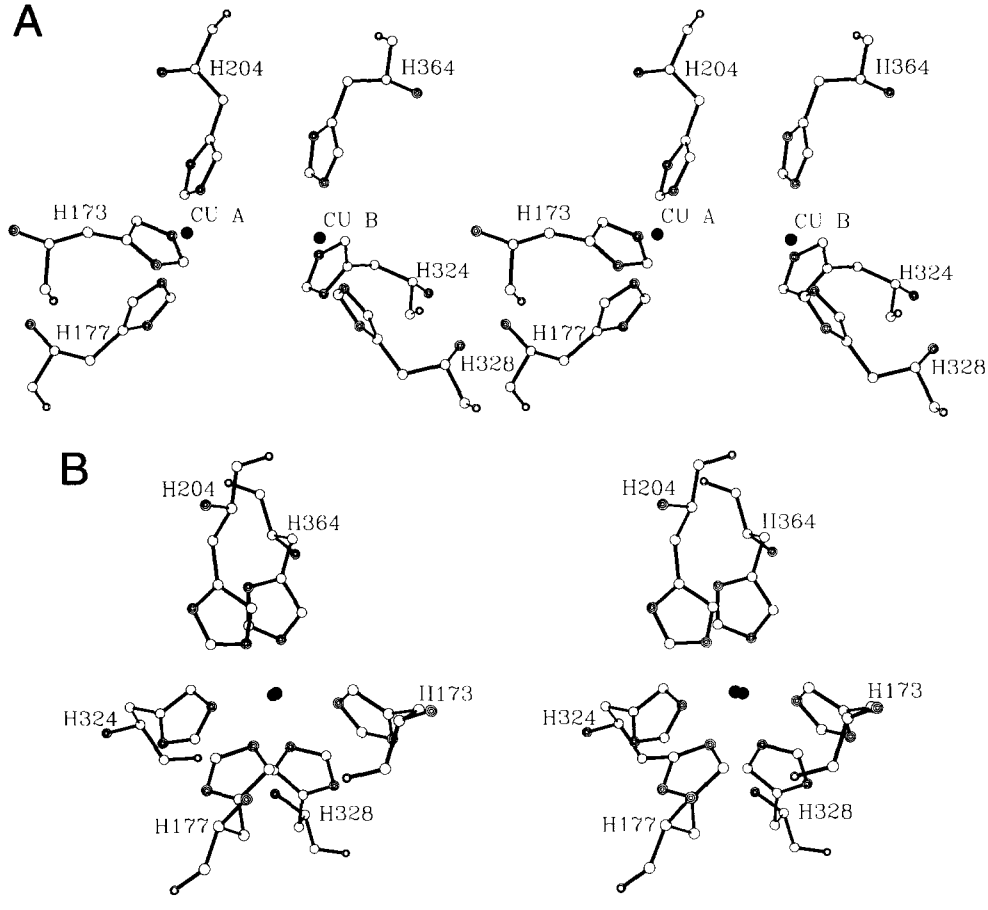

Fig. 6. Diagrams of the dinuclear copper site. **A:** View perpendicular to Cu-Cu axis. **B:** View along Cu-Cu axis. These figures and Figures 8, 10, and 11 were made by the ARTPLOT program (Lesk & Hardmann, 1985). Carbon, oxygen, and nitrogen atoms are represented by one, two, and three concentric circles, respectively. The copper ions are completely filled.

Absence of bridging ligand

Apart from the three histidine ligands, it has been speculated that the hemocyanin active site might contain a fourth ligand, bridging both coppers in the deoxygenated state (McKee et al., 1984; Volbeda & Hol, 1989a). The omit map density contoured at 3σ (shown in Fig. 7) gives no indication for such a bridging ligand. Even at a low contour level of 1σ (not shown), no density for a bridging ligand could be observed. Moreover, the nearly planar trigonal coordination of both coppers and the long Cu-Cu distance suggest that no bridging ligand is present.

Stabilization of histidine side-chain orientations

The properties of the hemocyanin dinuclear copper site depend on the geometry in which the coppers are coordinated by the protein. Therefore the interactions by which the protein orients the liganding histidine side chains are of great importance. In the *Limulus* II structure we see that the histidine side-chain conformations are largely controlled by steric interactions and by hydrogen-bond interactions with their imidazole N^{δ} atoms. In Figure 8 and Kinemages 4 and 6, the most important steric interactions, made with three phenylalanines, Phe 49, Phe 200, and Phe 360, are shown. These three phenylalanines are completely conserved in all known arthropodan hemocyanins, but not in the insect storage proteins, suggesting a functional rather than a structural role.

Figure 9 shows a schematic representation of the dinuclear copper site including the hydrogen bonds made by the histidine N^{δ} atoms (see also Kinemage 5). Each N^{δ} atom forms at least one hydrogen bond with a water or main-chain carbonyl oxygen atom. In addition His 324 makes a hydrogen bond to the side-chain oxygen of Asn 325. Asn 325 is conserved in all hemocyanins, but not in the insect storage proteins.

Cis-peptides

The final *Limulus* II model contains two cis-peptide bonds. A cis-proline was observed between Phe 598 and Pro 599. This cis-peptide bond is probably required for structural reasons, since it is located in a region that is well

Table 2. *Coordination of the dinuclear copper site*

Distance (Å)			
CuA-CuB	4.61		
CuA-NE173	2.10	CuB-NE324	2.16
CuA-NE177	2.01	CuB-NE328	2.08
CuA-NE204	1.94	CuB-NE364	1.92
Angle (degrees)			
NE173-CuA-NE177	99	NE324-CuB-NE328	97
NE173-CuA-NE204	131	NE324-CuB-NE364	108
NE177-CuA-NE204	126	NE328-CuB-NE364	142
CE173-NE173-CuA	128	CE324-NE324-CuB	126
CE177-NE177-CuA	122	CE328-NE328-CuB	113
CE204-NE204-CuA	110	CE364-NE364-CuB	103
CD173-NE173-CuA	117	CD324-NE324-CuB	119
CD177-NE177-CuA	128	CD328-NE328-CuB	134
CD204-NE204-CuA	140	CD364-NE364-CuB	146
Copper out of plane distance (Å)			
NE173-NE177-NE204	0.23	NE324-NE328-NE364	0.39
Imidazole 173	0.73	Imidazole 324	0.79
Imidazole 177	0.25	Imidazole 328	0.24
Imidazole 204	0.21	Imidazole 364	0.22
Histidine chi angle (degrees)			
	$\chi 1$ $\chi 2$		$\chi 1$ $\chi 2$
His 173	−174 −119	His 324	154 −89
His 177	−85 −76	His 328	−75 −95
His 204	−72 93	His 364	−85 −80
Dihedral angle (degrees)			
NE173-CuA-CuB-NE324	161	NE177-CuA-CuB-NE364	163
NE173-CuA-CuB-NE328	62	NE204-CuA-CuB-NE324	−69
NE173-CuA-CuB-NE364	−91	NE204-CuA-CuB-NE328	−168
NE177-CuA-CuB-NE324	55	NE204-CuA-CuB-NE364	39
NE177-CuA-CuB-NE328	−44		

conserved in both the hemocyanin and insect storage protein sequences (see Fig. 1).

A second cis-peptide was observed between Glu 309 and Ser 310 (see Kinemage 5). Non-proline cis-peptides are relatively rare since they are energetically more unfavorable. The cis-conformation for this peptide bond was however clearly indicated in an $F_o - F_c$ map, where a negative peak showed up at the position of the main-chain carbonyl oxygen in the trans-conformation, and a positive peak was found at the position for the oxygen in the cis-conformation. Furthermore an unallowed phi-psi combination of Ser 310 was removed by introducing the cis-peptide bond.

Pseudo-twofold symmetry in the dinuclear copper site

In the *Panulirus* structure a pseudo-twofold symmetry was observed between the helical pair donating the histidine ligands to Cu A and the helical pair donating the histidine ligands to Cu B (Volbeda & Hol, 1989b). This pseudo-twofold symmetry is also present in the *Limulus* II structure relating the same 31 residues (171–179 and 197–218 of Cu A with 322–330 and 357–378 of Cu B; see Kinemage 4) with an rms deviation between C_α atoms of 1.5 Å. In the *Limulus* II structure the symmetry even extends to the coordination of the two coppers. This can be seen in Table 2, where the coordination distances and angles for homologous histidines show the same trend. Also, the histidine dihedral angles have been conserved except for $\chi 2$ of His 204/His 364. In spite of the extended symmetry, the sequence identity of the twofold-related residues in the *Limulus* II structure is only 16%, whereas it was 26% in the *Panulirus* structure. It is, however, interesting to note that in addition to the histidine ligands, the functionally important Phe 200 and Phe 360, discussed above and depicted in Figure 8, are also related by the pseudo-twofold symmetry.

These two phenylalanines form the sequence pattern "Phe x x x His" in the copper A and B sites of all arthropodan hemocyanins sequenced so far, with His being the most C-terminal histidine ligand for each copper (histidines 204 and 364; see Fig. 1 and Kinemage 4). The separation of the phenylalanine and histidine by three residues on an α-helix results in their side chains pointing in the same direction from subsequent turns of the helix. In this way Phe 200 interacts tightly with His 204 and Phe 360

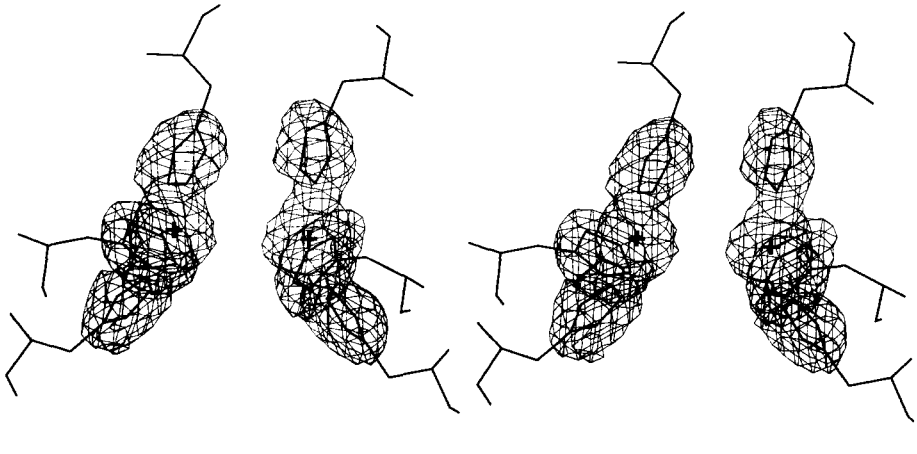

Fig. 7. Omit map of the dinuclear copper site. Both copper ions and the liganding histidine side chains beyond C_β were left out of the model. Before calculating the omit map an energy minimization run was carried out to reduce model bias. The density is contoured at a 3σ level. The copper density is about 33σ above background. Residue numbers for the histidines can be obtained from Figure 6A, which has the same orientation.

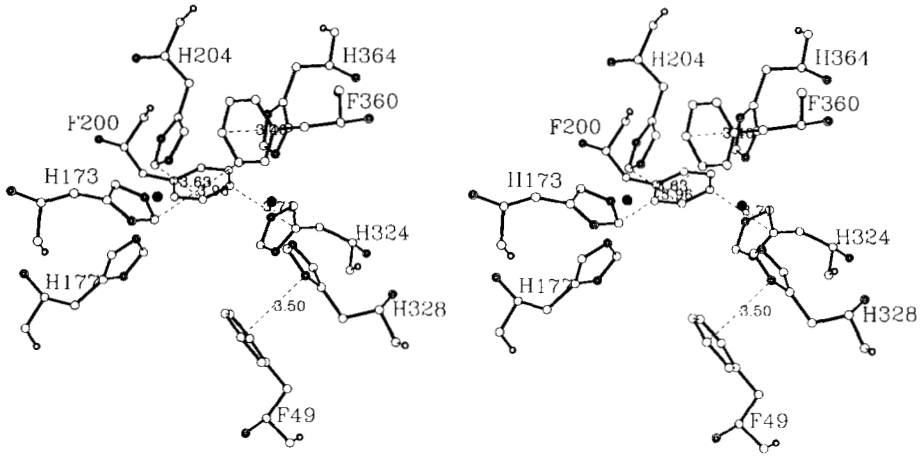

Fig. 8. Dinuclear copper site contacts with Phe 49, Phe 200, and Phe 360. The dashed lines give the shortest contact between a phenylalanine and a liganding histidine that is within 4 Å distance.

interacts tightly with His 364 (shortest contact distances are 3.6 Å and 3.4 Å, respectively). Surprisingly, we found that the "Phe x x x His" pattern is not only conserved in the arthropodan hemocyanins, but also in both the copper A and copper B sites of all known sequences of molluscan hemocyanins and tyrosinases (Lang et al., 1991). For the copper B site this is not unexpected, because there is a significant sequence homology with the arthropodan copper B site. For the copper A site, in contrast, no sequence homology with the arthropodan copper A site was reported so far. The finding of a conserved "Phe x x x His" pattern suggests, however, that the Cu A site in molluscan hemocyanins and tyrosinases has its most C-terminal histidine ligand donated by an α-helix. However, it cannot yet be decided whether this conserved feature is the result of convergent or divergent evolution.

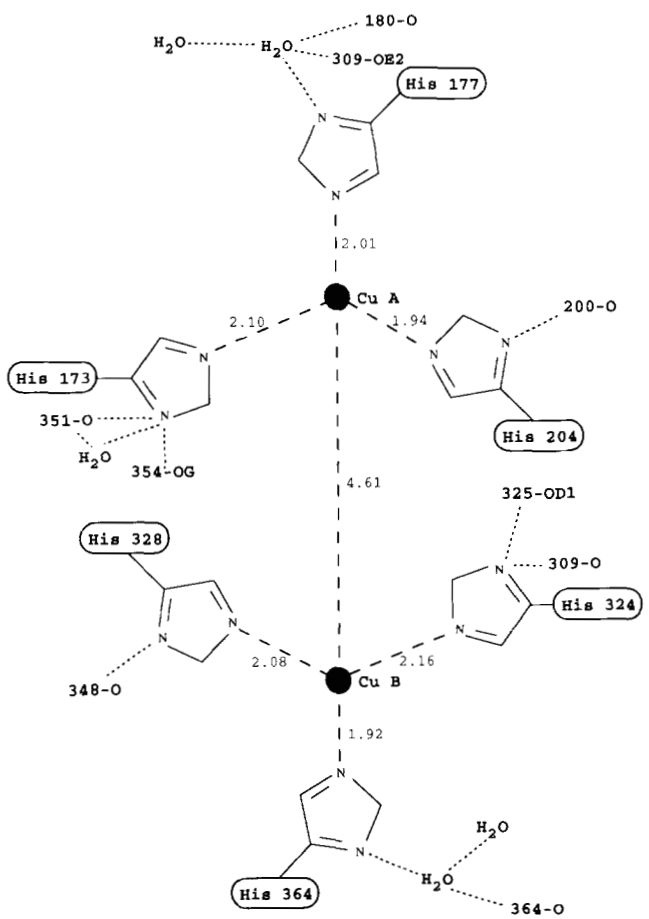

Fig. 9. Schematic diagram of the dinuclear copper site of *Limulus* II hemocyanin. The dashed lines show copper coordination distances. The dotted lines show the hydrogen bonds formed by the N^{δ} atoms of the liganding histidine residues.

Chloride-binding site

In the interface between domains 1 and 2 a strong density feature was observed. This density could not be fitted satisfactorily by a water molecule, because the *B*-value refined to the lower limit of 2 Å^2 set in X-PLOR and still gave a strong peak in the $F_o - F_c$ map. Based on the strength and the spherical shape of the density peak a heavier ion needed to be present. In view of the high (>0.5 M) NaCl concentration present during crystallization, the known allosteric regulation of *Limulus* II hemocyanin by chloride ions, and the nature of the ligands (including a positive arginine side chain), the peak was attributed to a chloride ion. In Figure 10 and Kinemage 2 the chloride-binding site is shown. The position of the chloride ion in the overall structure is given in Figure 5A.

Metal-binding site

During refinement, a density peak was observed in the third domain that could not be a water molecule in view of the many short contacts with neighboring atoms. Instead, the near octahedral coordination by six oxygen atoms, including a negatively charged aspartate side-chain oxygen (see the legend of Fig. 11), suggested the presence of a metal ion. Because the only metal ion abundant during crystallization was sodium, we expect that the metal-binding site is occupied by a sodium ion in our structure.

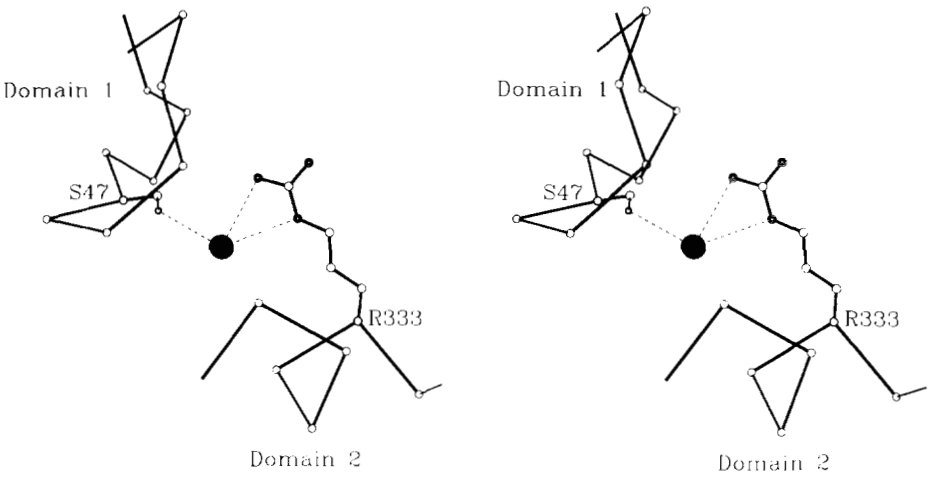

Fig. 10. Coordination of the chloride-binding site showing Ser 47 donated by a loop of the first domain and Arg 333 donated by helix 2.5 of the second domain. Coordination distances are: Ser 47 OG-Cl^- = 2.94 Å; Arg 333 NE-Cl^- = 3.40 Å; Arg 333 NH2-Cl^- = 3.38 Å.

It is, however, tempting to speculate that in vivo this metal-binding site is occupied by a calcium ion, because (1) calcium ions are known to be important for both the structural and functional properties of *Limulus* hemocyanin (Brouwer et al., 1983; Brenowitz et al., 1984), and (2) the nature and coordination geometry of the ligands correspond well with those observed in calcium ion-binding sites (Strynadka & James, 1991). We will therefore refer to this binding site as the putative calcium-binding site. In Figure 11 and Kinemage 3 this site is shown in detail. The position in the overall structure is depicted in Figure 5A.

Comparison of deoxygenated Limulus II versus deoxygenated Panulirus interruptus hemocyanin

The overall folds of the deoxygenated *Limulus* II and deoxygenated *Panulirus interruptus* hemocyanin structures are similar. However, local and global differences do exist. In this section we will focus on the global changes represented by rigid body motions. As a first analysis, we have superimposed parts of the hemocyanin structures, ranging in complexity from single domains to the full hexamer. The results of these superpositions are given in the top half of Table 3. The domain superpositions show that domain 2 superimposes better than domain 3 and that domain 1 superimposes least well. This is nicely correlated with the sequence homology in the three domains, as is shown in Table 3.

The monomer superimposes less well than could be expected based on the results for the domains separately. This indicates that the relative positions of the domains within the subunit of the two hemocyanin structures is different. To quantify these differences we have superimposed the second domains of both structures by least-squares fitting of the C_α atoms (Rao & Rossmann, 1973). Subsequently we analyzed the additional transformation needed to optimally superimpose both domains 1 and 3. The results, shown in the lower half of Table 3, indicate that the major difference is the orientation of the first domain, which is rotated by 7.5° with respect to the second

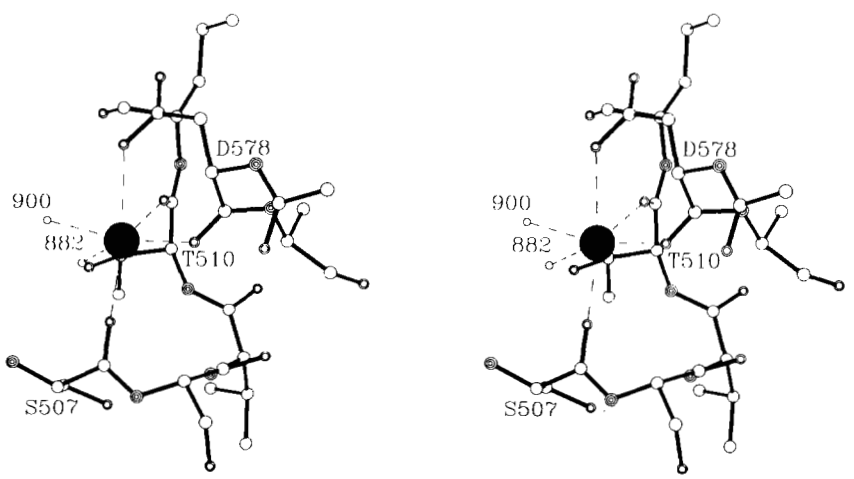

Fig. 11. Coordination of the putative calcium site. Coordination distances are: Ser 507 O-Ca^{2+} = 2.29 Å; Thr 510 O-Ca^{2+} = 2.35 Å; Asp 578 O-Ca^{2+} = 2.31 Å; Asp 578 OD2-Ca^{2+} = 2.79 Å; H_2O 900-Ca^{2+} = 2.43 Å; H_2O 882-Ca^{2+} = 3.02 Å.

Table 3. *Analysis of global differences in the quaternary and tertiary structures of Limulus II and Panulirus hemocyanin*

Structural unit	Number of atoms[a]	rms deviation (Å)	Sequence identity (%)[b]
rms deviation of equivalent C_α atoms			
Domain 1	108	1.25	25.5
Domain 2	210	0.87	39.9
Domain 3	188	1.00	34.7
Monomers	506	1.24	33.1
Tight dimers	1,012	1.47	33.1
Loose dimers	1,012	1.59	33.1
Trimers	1,518	1.26	33.1
Hexamers	3,036	1.70	33.1

Structural unit A	Structural unit B	Kappa (°)	Phi (°)	Psi (°)	Δt (Å)[d]
Relative movements[c]					
Domain 2	Domain 1	7.5	144	132	0.2
Domain 2	Domain 3	0.8	109	34	−0.5
Domain 3	Domain 1	7.7	147	137	0.0
First monomer of tight dimer	Second monomer of tight dimer	2.7	288	100	−0.3
First monomer of loose dimer	Second monomer of loose dimer	2.5	270	94	−0.5
First monomer of trimer	Second monomer of trimer	0.9	0	70	0.4
First trimer	Second trimer	3.1	270	90	−0.9

[a] The number of residues that were superimposed. Only the residues that are structurally homologous in the *Panulirus* and *Limulus* II structures were used in the superpositioning. These residues were selected on a graphics device and are indicated by "=" in Figure 1.

[b] All residues that are aligned in Figure 1 were used.

[c] Relative movements between two structural units were obtained by first superimposing structural unit A of the *Limulus* II and the *Panulirus* structures. The given data then describe the additional rotation and translation that is required to subsequently superimpose structural unit B of both hemocyanins.

[d] Translation along the rotation axis.

domain. The second and third domains seem to have relatively similar orientations in both structures.

When the hexamers of the two hemocyanins are superimposed, the rms deviation is 1.70 Å. This is considerably higher than the 1.24 Å for the superimposed monomers. However, when two trimers are superimposed, the rms deviation drops to only 1.26 Å. The loose and tight dimers (for a definition see Volbeda & Hol, 1989a) give intermediate results, with rms deviations of 1.59 Å and 1.47 Å, respectively. These results indicate that the organization of subunits in the hexamer, tight dimer, and loose dimer is dissimilar in the *Panulirus* and *Limulus* hemocyanin structures, whereas the quaternary structure of their trimers is very similar. Hence, the major difference between both structures must be a different position of the two trimers with respect to each other.

To analyze the difference in quaternary structure we first superimposed one subunit of a structural unit (a trimer, tight dimer, or loose dimer) of *Panulirus* and *Limulus* hemocyanin. Subsequently we determined the additional transformation that gave the best fit for a second subunit of the same type of structural unit. The results, given in the lower half of Table 3, show that the difference in subunit orientation within the tight and loose dimers is approximately three times as large as for the subunits in the trimer (2.7°, 2.5°, and 0.9°, respectively). In the same manner we found that, once a trimer of *Panulirus* is superimposed onto a trimer of *Limulus* II, a rotation of 3.1° about the threefold axis and a translation of 0.9 Å along the axis is required to optimally superimpose the other trimers of the two hemocyanins (in the *Panulirus* structure the trimers are further apart).

Discussion

Cis-peptide Glu 309–Ser 310

The presence of cis-peptides or other forms of strain in protein structures is often linked to functionally important residues (Herzberg & Moult, 1991). The same is observed in the *Limulus* II structure, where a cis-peptide linkage occurs between Glu 309 and Ser 310. The carbonyl oxygen of this cis-peptide plane is directly hydrogen bonded to the N^δ atom of the copper ligand His 324 (see Kinemage 5) and may therefore be required for the correct positioning of the His 324 side chain. Alternatively, the cis-peptide plane may be required in relation to a special function of Glu 309 as discussed below.

Oxygen entrance pathway

The dinuclear copper site is completely buried in the core of domain 2. Some structural change is therefore required before oxygen can reach the coppers. Inspection showed that there is a solvent tunnel that extends from the protein surface to the dinuclear copper site. This tunnel is obstructed by the carboxyl side-chain atoms of Glu 309 (see Fig. 12). Glu 309 could therefore act as a "shutter" that controls the entrance of oxygen to the coppers by switching between alternative conformations. The negative charge of the glutamate side chain seems to be important for its function, because a glutamate is conserved in all hemocyanins, whereas in six of the seven insect storage protein sequences a noncharged residue is present (5 × Gln, 1 × Leu; see Fig. 1). Three other residues close to Glu 309 have been conserved in all hemocyanin sequences: (1) Asn 325: the side chain of this residue is hydrogen bonded to the Glu 309 side chain via a water molecule and makes a direct hydrogen bond to the Glu 309 carbonyl oxygen in the cis-peptide; (2) Lys 196: this residue forms a salt bridge to Glu 309 at a distance of 4.7 Å; and (3) Ser 311: this residue might become a hydrogen-bond partner of Glu 309 in an alternative conformation.

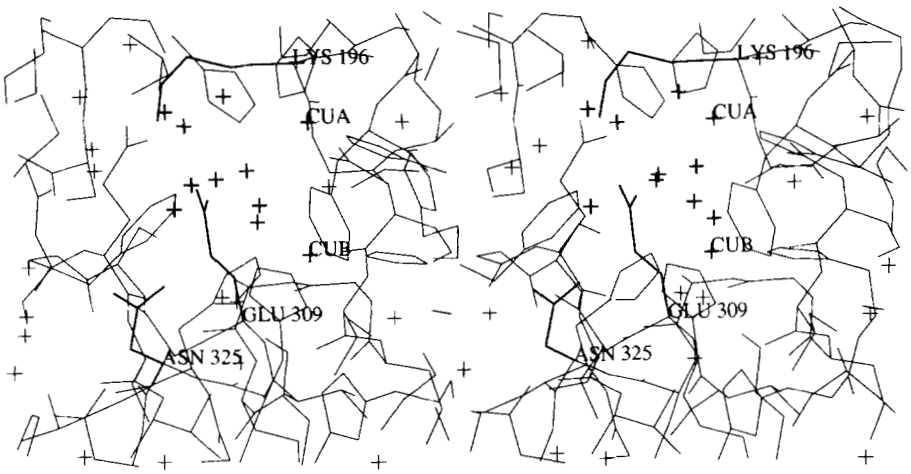

Fig. 12. Picture of the putative oxygen entrance pathway. The copper ions and the water molecules that fill the solvent tunnel are represented by bold pluses. The conserved residues Glu 309, Lys 196, Ser 311, and Asn 325 are also printed in bold lines and are discussed in the text.

Surprisingly, we found that the molluscan hemocyanins also have completely conserved Glu 309 as well as Asn 325 (residues 159 and 175 in Lang et al., 1991). There is, however, no other obvious sequence conservation in the Glu 309 region between the two classes of hemocyanins. It therefore remains unclear if the two classes of hemocyanins have conserved a glutamate at this position for the same functional role.

Carbon monoxide binding

Carbon monoxide is a ligand for the reduced dinuclear copper site of molluscan and arthropodan hemocyanins, as well as tyrosinases (Yen Fager & Alben, 1972; Bonaventura et al., 1974). In contrast to oxygen binding, only one copper ion is involved in the binding of carbon monoxide. Sorrel and Jameson (1982, 1983) suggested that this may be explained by a linear coordination of one of the copper ions in the dinuclear copper site, because model compound studies showed that linearly coordinated copper ions bind carbon monoxide very weakly. Based on the *Limulus* II structure, neither of the copper ions has a linear coordination. The Cu B site has, however, distortions that tend toward a linear coordination, with a His 328–Cu B–His 364 bond angle of 142° and the third histidine ligand, His 324, in a more distal position at 2.2 Å. (For the Cu A site the distortions are smaller. The largest bond angle, between His 173–Cu A–His 204, is now 131°. Moreover, the third histidine ligand, His 177, is not at a distal position, but it is tightly bound at 2.0 Å [see Table 2].) Based on the slightly more linear character of Cu B, it might be that carbon monoxide binds more tightly to Cu A than to Cu B. It is, however, uncertain if this is sufficient to explain why hemocyanins bind only one carbon monoxide molecule. The mechanism that prevents the binding of a second carbon monoxide molecule could also involve steric restrictions or conformational changes resulting from binding the first carbon monoxide molecule. It is however interesting to note that in tyrosinase, carbon monoxide has been experimentally shown to bind to Cu A (Jackman et al., 1991).

Regulation of oxygen binding

The hemocyanins are a class of proteins having well-developed allosteric interactions. Their allosteric regulation seems to provide the tissues with sufficient oxygen under all physiological conditions. The homotropic effect of oxygen is manifest in the cooperativity between dinuclear copper sites. Important heterotropic allosteric regulators of *Limulus* hemocyanins are Cl^-, Ca^{2+}, and protons (Brouwer et al., 1977, 1982, 1983; Brouwer & Serigstad, 1989). It is interesting to note that the effect of protons on oxygen binding by whole *Limulus* hemocyanin, the Bohr effect, is "reverse" to that seen in most other hemocyanin and hemoglobin systems.

The influence of chloride ions

The effect of chloride ions is not general for all hemocyanins. Even for *Limulus* hemocyanin the chloride effect is observed only for subunit types II, IIA, IIIA, and IIIB. Detailed information on the regulation by chloride ions on both the native complex as well as on separate subunits can be found in Sullivan et al. (1974), Brouwer et al. (1982), and Brenowitz et al. (1984). From these studies it becomes clear that the function of chloride ions is to reduce the oxygen affinity by more tightly binding to hemocyanin in the low-affinity state (T-state) than to the high-affinity state (R-state). It was further found that chloride ions reduce the oxygen affinity of the monomer as well as of the hexamer. The chloride effect on monomeric *Limulus* II hemocyanin indicates that the chloride-binding site undergoes a structural change at the tertiary level when the subunit goes from the T- to the R-state. In agreement with this, the chloride-binding site in our crystal structure is not observed in an intersubunit interface. Instead it is located in the interface between domains 1

and 2, which both donate one ligand, Ser 47 and Arg 333, respectively (see Figs. 5A, 10; Kinemage 2).

An indication that a change in the domain 1–domain 2 interface might indeed occur comes from the comparison of the *Panulirus* and *Limulus* II structures. The position of domain 1 with respect to domain 2 in the two structures was found to differ by a rotation of 7.5° (see Table 3, Fig. 13A). A crucial question is whether the deoxygenated *Limulus* II hemocyanin structure reported here and the published *Panulirus* hemocyanin structure (Volbeda & Hol, 1989a) represent the T- and the R-state conformation, respectively. In order to investigate the conformational states of *Limulus* II and *Panulirus* hemocyanin in the crystal structures, we have performed oxygen-binding experiments in solutions that mimic crystallization conditions (data not shown). These studies show that under such conditions *Limulus* II hemocyanin binds oxygen noncooperatively and with a T-like affinity. For *Panulirus* hemocyanin (at pH 3.8) no oxygen binding could be detected. Since the oxygen affinity in the latter case does not resemble either the T-state or the R-state affinity, we cannot assess the conformational state of the *Panulirus* crystal structure. We will therefore refer to the *Panulirus* and *Limulus* II structures as the "open" and "closed" states, respectively, reflecting the nature of the domain 1–domain 2 interface shown in Figure 13.

Figure 13B shows the chloride-binding site in *Limulus* II hemocyanin (thin lines). In the same figure the comparable region of the *Panulirus* structure is shown (bold lines) after superpositioning the second domains of the two structures. Let us assume that the "open" *Panulirus* structure resembles the *Limulus* II structure in the R-state. The reduced chloride affinity of the *Limulus* II R-state would then be explained, because the outward movement of domain 1 and its associated chloride ligand Ser 47 would disrupt the chloride-binding site.

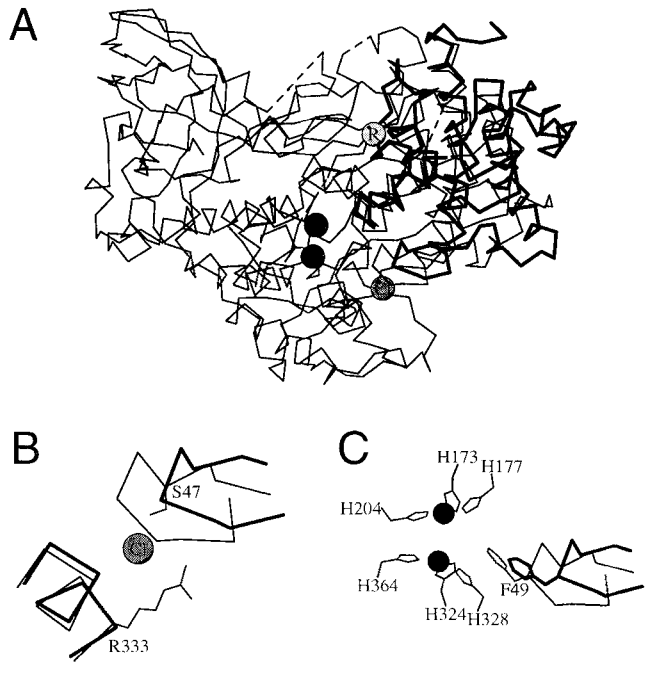

Fig. 13. Structural comparison of *Limulus* II and *Panulirus* hemocyanin subunits. **A:** The C_α trace of the full *Limulus* II subunit is shown (thin lines) plus the copper ions (black spheres), the chloride ion (dark gray sphere), and the position of the domain 1 rotation axis (pale gray sphere with the label "R"). All diagrams have the same view with the domain 1 rotation axis perpendicular to the plane of the paper. For clarity, only domain 1 from the *Panulirus* structure is shown (bold lines). The position of domain 1 of the *Panulirus* structure was obtained after superpositioning the second domains of both hemocyanins, illustrating the 7.5° rotation of domain 1 relative to domains 2 and 3. **B:** Detail of the chloride-binding site, showing the effect of the domain 1 rotation on the loop containing Ser 47. **C:** Detail of the dinuclear copper site, showing the effect of the domain 1 rotation on the position of Phe 49. Note the tight packing of the Phe 49 and His 328 side chains in the *Limulus* structure (thin lines).

Structural mechanism of oxygen affinity regulation

The hypothesis that the open and closed states resemble the R- and T-states, respectively, is further supported by the observation that the rotation of the first domain has a direct effect on the dinuclear copper site environment. In Figure 13C, this effect is shown to be mediated by Phe 49. This crucial residue is close to the coppers and changes its position dramatically in the transition from the closed to the open state. An important functional role for Phe 49 is also suggested by its conservation in all known hemocyanin sequences, but not in the insect storage proteins.

In the closed state, shown in thin lines in Figure 13C, the aromatic ring of Phe 49 packs tightly onto the imidazole ring of His 328. Both rings are nearly parallel and at a mutual distance of approximately 3.5 Å (shortest contact distance). Due to this interaction in the closed state, Phe 49 restricts the conformational freedom of His 328. When oxygen binds to the dinuclear copper site, the Cu–Cu distance has to shorten from 4.6 to 3.6 Å. This movement must be associated with a reorientation of the copper ligands. We propose that the position of Phe 49 in the closed state prevents His 328 from moving into the position that is energetically most favorable for the oxygenated dinuclear copper site. This energetic penalty does not exist in the open state, because the removal of Phe 49 now allows His 328 to reach the lower energy position. Accordingly, the open state has a higher oxygen affinity than the closed state.

Cu–Cu distance

The Cu–Cu distance of 4.6 Å observed in the *Limulus* II structure differs significantly from the expected 3.4 Å. This expectation was based on EXAFS studies and X-ray structures of *Panulirus* hemocyanin and ascorbate oxidase (Brown et al., 1980; Woolery et al., 1984; Messerschmidt

et al., 1989; Volbeda & Hol, 1989a). One explanation for this discrepancy is that the Cu–Cu distance does not only depend on the ligation state, but that it is also directly influenced by the conformational state of the protein. In the open state a Cu–Cu distance of 3.4 Å would then be most stable (in this state the cuprous ions may possibly be linked by a hydroxyl bridging ligand). In the closed state, however, the dinuclear copper site with a Cu–Cu distance of 3.4 Å would be destabilized because His 328 is forced in a less favorable position by Phe 49, as described above. Consequently, a Cu–Cu distance of 4.6 Å may become the most stable conformation.

It is interesting to note that the distance between the His 328 side chain in the open state, bold lines in Figure 13C, and the Phe 49 side chain in the closed state, thin lines in Figure 13C, is 2.6 Å (shortest contact distance). The position of the His 328 side chain observed in the open state is thus incompatible with the position of Phe 49 in the closed state. Another interesting observation is that in the closed state (*Limulus* II structure) the phi–psi angle of Phe 49 is outside the allowed region although it fits nicely in density, whereas in the open state (*Panulirus* structure) the corresponding phenylalanine falls well within the allowed region. This local strain in the molecule in the closed state may be a consequence of the function of Phe 49—prevention that the dinuclear copper site reaches its lower energy state.

The dinuclear copper sites in the *Limulus* II and *Panulirus* structures and the interactions with Phe 49 are also illustrated in Kinemage 6.

The influence of calcium ions

Calcium ions play both a structural and a regulatory role in *Limulus* hemocyanin. They stabilize the multihexameric native structure as well as the homohexamers of *Limulus* II (Brenowitz et al., 1984). Calcium also decreases the oxygen affinity and increases the cooperativity of the multi-hexameric complex (Brouwer et al., 1983). Due to the presence of 10 mM EDTA in the crystallization medium, no calcium ion was expected to be bound in our structure. We did, however, observe a metal-binding site, which could very well be a calcium-binding site in vivo, although in our crystals it is most likely occupied by a sodium ion.

The binding site is not close to any intersubunit interface or to the interface between domains 1 and 2. Instead it is located at the outside of the hexamer and involves residues of a long loop (residues 505–541) that extends from the third domain and interacts with the first domain via helix 3.3 (see Fig. 5A,D; Kinemage 3). If the calcium site has a regulatory role, then the regulatory effect is most likely mediated by contacts between helix 3.3 and domain 1. If, in contrast, the calcium has a structural role, it could be involved in stabilizing interhexameric contacts or in stabilizing the local structure. The latter option is interesting because the calcium ion is close to the flexible loops 527–530 and 569–572 (see Fig. 5D), which may be stabilized by calcium binding.

The Bohr effect

In contrast to most other oxygen transport proteins, including most hemocyanins, native *Limulus* hemocyanin displays a reverse Bohr effect. This results in an increased oxygen affinity at lower pH. In addition, protons stabilize the quaternary structure. *Limulus* II monomers also show a negative Bohr effect in the pH range between 7 and 9 (Sullivan et al., 1974; Brenowitz et al., 1984). This monomeric response indicates the presence of an intrasubunit effect. In addition, protons stimulate hexamer formation of *Limulus* II hemocyanin in the same pH range if 10 mM $CaCl_2$ or 1 M NaCl is present (Brenowitz et al., 1984). This hexamerization leads to a positive Bohr effect, which overrides the negative Bohr effect observed for the free monomers.

The pH range in which the Bohr effect is observed suggests the possible involvement of histidine residues. Of the 45 histidines in the *Limulus* II sequence only three are involved in intersubunit contacts, viz. His 239, His 281, and His 342 (see Table 1). From these three histidines only the contact between His 239 and Glu 343 would become stronger upon lowering the pH. There are however several other interesting histidines. Of these, His 50 is particularly intriguing, because it is located in the domain 1–domain 2 interface, very close to the chloride-binding site and adjacent to Phe 49, which we propose is crucial for the mechanism of oxygen affinity regulation. An understanding of the mechanism(s) that lead to the Bohr effect will, however, require more structural and functional studies.

Cooperativity

In the paragraphs above, the regulation of oxygen affinity has been explained by the presence of two conformational states of the hemocyanin subunit: a closed and an open state. Several lines of evidence, presented above, suggest that the "closed" *Limulus* II structure represents the low oxygen affinity state and that the "open" *Panulirus* structure resembles the high oxygen affinity state. In order to obtain cooperativity, there must, however, be a mechanism by which the changes on the tertiary structural level influence the overall conformational state of the hexamer. This mechanism is almost certainly based on the domain 1 rotation, because this seems to be the only likely mechanism by which the observed effect of chloride ions on the conformational transition of the hexamer can take place. (In theory it is possible that a second chloride-binding site is involved, but because we have not observed such a site we will not discuss this further.)

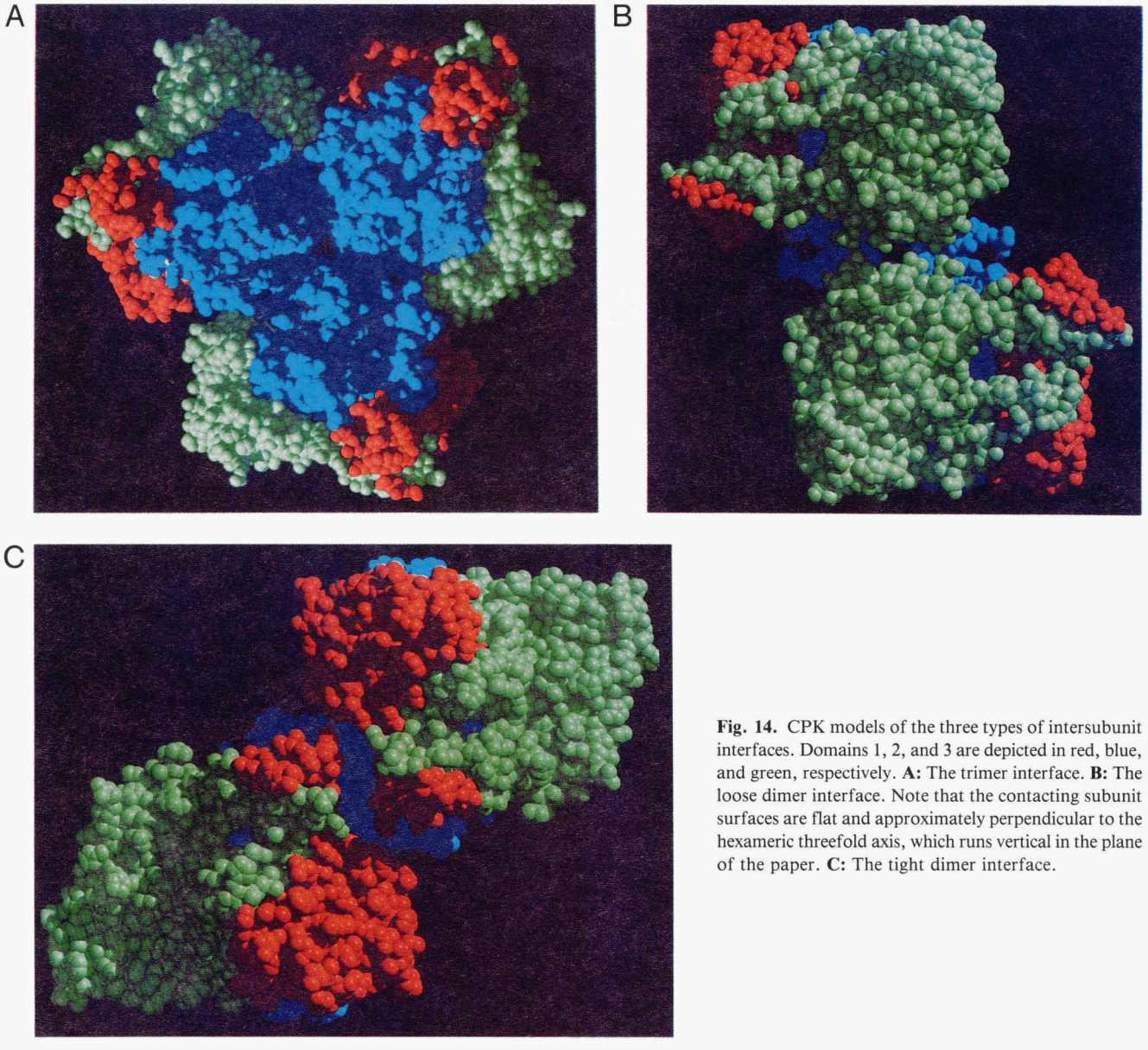

Fig. 14. CPK models of the three types of intersubunit interfaces. Domains 1, 2, and 3 are depicted in red, blue, and green, respectively. **A:** The trimer interface. **B:** The loose dimer interface. Note that the contacting subunit surfaces are flat and approximately perpendicular to the hexameric threefold axis, which runs vertical in the plane of the paper. **C:** The tight dimer interface.

In the hemocyanin hexamer three types of intersubunit interfaces can be distinguished: (1) a trimer interface between subunits in one trimer, (2) a loose dimer interface between two subunits in different trimers that interact via domains 2 and 3, and (3) a tight dimer interface between two subunits in different trimers that interact via domains 1 and 2. These three intersubunit interfaces are shown in Figure 14. Based on the superposition results given in Table 3, the trimers were found to behave as rigid bodies that can rotate with respect to each other about the threefold axis. The rigidity of the trimer agrees with the observation that domain 1 does not make significant interactions in the trimer interface in both the *Limulus* II and *Panulirus* structures (see Table 1). Movement of domain 1 will thus not directly influence the trimer interface. In this view it is more consistent to think of the hemocyanin hexamers as dimers of trimers than as trimers of dimers, as had been previously proposed (Gaykema et al., 1986). In the loose dimer, domain 1 is also not involved in any contacts. The loose dimer interface forms approximately a plane perpendicular to the threefold axis (see Fig. 14B). This allows the two trimers to rotate with respect to each other about the threefold axis, without severe steric collisions. The tight dimer interface, in contrast, does involve residues from domain 1 and is therefore the prime candidate to transmit the cooperative signal. In addition, the intersubunit interface is not perpendicular to the threefold axis (see Fig. 14C). Therefore, a rotation of the trimers with respect to each other as well as a rotation of domain 1 may be expected to lead to gross alterations in the tight

dimer interface unless both motions are linked in a way that maintains an energetically allowed subunit interface. Accordingly, we propose that the trimer and domain 1 rotations occur in a concerted fashion, linking the tertiary and quaternary transitions in the hexamer.

Conclusion

Our results on the structure of deoxygenated *Limulus* II hemocyanin, the published structure of *Panulirus* hemocyanin, and the knowledge of their functional properties have allowed us to make detailed proposals at the atomic level for some of the allosteric properties of arthropodan hemocyanins. This model invokes changes at both tertiary and quaternary levels of the protein structure. In this model the low and high oxygen affinity states are represented by closed and open conformational states. The reduced oxygen affinity of the closed conformation arises from unfavorable interactions of Phe 49 with the oxygenated dinuclear copper site. The high oxygen affinity of the open conformation is due to the movement of Phe 49 away from the dinuclear copper site. The movement of Phe 49 is associated with a 7.5° rotation of domain 1. We propose that the rotation of domain 1 can only occur in a concerted fashion, accompanied by a rotation of the trimers in the hexamer. This links the tertiary and quaternary changes in the hexamer and forms the basis for cooperativity. The rotation of domain 1 also predicts the preferential binding of chloride ions to the low affinity state correctly.

Phe 49, the key residue in the allosteric mechanism, is completely conserved in all arthropodan hemocyanins. This suggests that the proposed mechanism is general for all arthropodan hemocyanins. In contrast, the type of allosteric effectors and their role in oxygen binding varies considerably, even between closely related hemocyanins. This evolutionary flexibility in the use of regulatory compounds may arise from the fact that allosteric effectors can be targeted to a large area, including both the domain 1–domain 2 interface and the intersubunit interfaces.

Glu 309 has many special properties and blocks access of oxygen to the dinuclear copper site. This suggests a role in the control of oxygen entrance to the dinuclear copper site. In the *Panulirus* structure, the corresponding glutamate (Glu 329) has the same side-chain conformation. Accordingly, the Glu 309 side-chain conformation does not seem to be directly coupled to the overall conformational state of the subunit.

At this moment the proposed mechanism of allosteric regulation is speculative, since we cannot prove that the open *Panulirus* structure resembles the high oxygen affinity state. New crystal forms have been obtained very recently however, and we therefore hope to obtain a more complete knowledge of the structural changes accompanying oxygen binding in the near future.

Materials and methods

Crystallization

Purified *Limulus* II protein was prepared as described by Brenowitz et al. (1981). Crystals were obtained by the hanging drop technique as described by Magnus et al. (1991), with some small modifications leading to the following procedure.

The protein was dialyzed against 0.05 M glycine, 0.01 M EDTA, 0.05 M Tris/HCl, at pH 8.9. The protein concentration after dialysis was approximately 40 mg/mL. Reservoir solution was prepared by adding 40 μL 0.5 M glycine, 0.1 M EDTA, 0.5 M Tris/HCl at pH 8.9, 90 μL 1 M bis-tris at pH 6.0, 80 μL saturated NaCl, 80 μL PEG 6000, solution prepared by adding 50 mL water to 25 g of PEG 6000 and 340 μL water. Drops were made by adding 7.5 μL of reservoir solution to 7.5 μL protein solution on a siliconized coverslip. Following this procedure the final pH of the hanging drop is just below 7. Large blue (oxygenated) crystals with dimensions up to 1.0 × 0.5 × 0.5 mm^3 were normally obtained within a few days.

Deoxygenation of hemocyanin in the crystalline state

Crystals were deoxygenated by mounting them in an X-ray capillary, together with a column of buffer containing 250 mM Na-dithionite. The crystal and the dithionite solution were separated by air. The dithionite was dissolved in a modified mother liquor containing 2.5 mL 0.5 M glycine, 0.1 M EDTA, 0.5 M Tris/HCl at pH 8.9, 1 mL 1 M bis-tris at pH 6.0, 0.8 mL PEG 6000 solution prepared as described above, and 2.0 mL water. Deoxygenation was monitored by the crystal color, which normally changed from blue to colorless within a few hours. Colorless crystals were assumed to be fully deoxygenated.

Unit cell parameters

The crystals have space group R32, with unit cell parameters of $a = b = c = 117.0$ Å and $\alpha = 60.0°$ in the rhombohedral setting (or $a = b = 117.0$ Å, $c = 286.6$ Å in the hexagonal setting). This agrees well with the values reported by Magnus et al. (1991). The crystals contain one subunit of 73 kDa per asymmetric unit and have a volume/Dalton ratio of 2.59 $Å^3$/Da, which corresponds to a solvent content of 52%.

In space group R32 a special situation occurs when $\alpha = 60°$. In this case the unit cell axes a, b, and c have the same length as the axes running along the unit cell diagonals (a–b, a–c, and b–c). In our case α refined to 60.02°, which complicated crystal orientation determination, as will be described below.

Data collection

Two data sets of deoxygenated *Limulus* II crystals were collected. The first data set, to 2.5 Å resolution, was collected on an Enraf-Nonius FAST area detector, employing CuK_α radiation from an Elliot GX-21 rotating anode. This data set was used to find the correct molecular replacement solution and to do the initial rigid body refinement. The second data set was collected on the imaging plate of the X-31 beamline in the Hamburg Synchrotron. Both a high- and a low-resolution data set could be collected from a single crystal, using a wavelength of 1.009 Å. For both the high and the low resolution data set 76° of data were collected, starting from the same spindle axis position. The threefold axis was slightly misaligned with respect to the spindle axis, such that cusp data collection was not necessary.

Data processing

The FAST data was processed with the MADNES software (Pflugrath & Messerschmidt, 1986). Unfortunately, the autoindexing procedure in MADNES was not able to find the correct crystal orientation. This was caused by the fact that the autoindexing procedure uses the length of the cell axes and not the symmetry. In space group R32 with $\alpha = 60°$, the actual unit cell axes and the diagonals have identical lengths. Accordingly, the autoindexing procedure could not distinguish between the correct axes and the diagonals. This problem was overcome by initially processing the data in space group P1. A special program was written to determine the correct setting based on symmetry analysis.

The Synchrotron data were processed with the MOSCO package (Machin et al., 1983), adapted for the use of the imaging plate. In this case the autoindexing program REFIX gave the correct crystal setting as the best-fitting solution. Data reduction was carried out by programs from the Groningen BIOMOL software package for both the FAST and the Synchrotron data. Statistics for both data sets are given in Table 4.

Molecular replacement model

The *Limulus* II starting structure was based on the *Panulirus* structure named SU4 in Volbeda and Hol (1989a), using the sequence alignment as given by Linzen et al. (1985). Because the sequence identity between *Limulus* II and *Panulirus* hemocyanin is rather low (33%) a conservative strategy was chosen to minimize the chance of introducing errors into the model. In generating the *Limulus* II starting structure, the following residues were completely left out of the *Panulirus* structure:

1. Residues that were not included in the original *Panulirus* structure due to uninterpretable density (15 in all).
2. Residues that should be deleted according to the amino acid sequence alignment.
3. The residues immediately bordering insertions and deletions.
4. The first 42 N-terminal residues. These were left out because in this region a 21-residue deletion occurs and it was not clear where the deletion had to take place.

Side chains of the *Panulirus* model were retained if *Limulus* II contained an identical residue in the sequence alignment. A modified *Panulirus* side chain was obtained by adding or deleting relevant atoms if one of the following substitutions needed to be made: Thr → Ser, Tyr → Phe, Phe → Tyr, Asn → Asp, Asp → Asn, Gln → Glu, Glu → Gln, Ile → Val, all → Ala, all → Gly. All other residues were included as alanines. B-values were kept as in the *Panulirus* structure. The model did not contain the two copper ions nor any waters.

In this way 3,626 atoms were included in the *Limulus* II starting structure, which represents approximately 70% of the expected scattering mass.

Molecular replacement

In the general molecular replacement problem six parameters have to be determined. However, in our case the search model was anticipated to be a hemocyanin hexamer with 32-point group symmetry. Based on the volume/Dalton ratio of 2.59 $Å^3$/Da, it was likely that only one subunit per asymmetric unit was present. This can only be achieved when the hexameric 32-point group symmetry coincides with a crystallographic 32-point group symmetry. In space group R32 this leaves only two independent solutions to the molecular replacement problem (the two

Table 4. *Data collection statistics*

Data set	Maximum resolution (Å)	Completeness overall (%)	Completeness in last shell (%)	Rsym[a] (%)	Number of reflections
Fast	2.5	81.2	31.0	5.0	21,607
Synchrotron	2.18	97.0	92.0	4.6	38,255

[a] $\text{Rsym} = 100 * \Sigma\, |I - \langle I \rangle| / \Sigma\, \langle I \rangle$.

solutions are related by a 180° rotation about the threefold axis). *R*-factors calculated for both solutions were 53% and 56% for data between 10.0 and 3.5 Å. In Figure 15 the *R*-factor is shown as function of resolution for both molecular replacement solutions. For the solution with the lower *R*-factor, the *R*-factor clearly increases with increasing resolution, whereas for the other solution the *R*-factor seems to be rather independent of the resolution. This suggests that the solution with the lower *R*-factor is indeed correct.

Map calculation and model building

All electron density maps were calculated with the XTAL package (Hall & Stewart, 1987) using weight factors as described by Read (1986). Model building was carried out with the program FRODO (Jones, 1978) on a PS390 graphics system. Protein atoms added during model building were given an initial *B*-value of 20 Å^2. Waters were added automatically with a local program. Putative water positions (peaks in the $F_o - F_c$ map $>3\sigma$) were included if at least one hydrogen bond was made with a distance between 2.5 and 3.4 Å and no other nonbonded interactions with a distance shorter than 2.9 Å occurred. Waters were given an initial *B*-factor of 25 Å^2. After each refinement cycle all waters with *B*-factors above 50 Å^2 were removed from the model.

Refinement

Refinement was started with TNT (Tronrud et al., 1987) using the molecular replacement solution that gave the lower *R*-value. A rigid body refinement of the whole subunit was first carried out to allow for deviations in the quaternary structures of *Limulus* II and *Panulirus* hemocyanin. In the resulting $2F_o - F_c$ map, density showed up for many atoms not yet included in the model, indicating the correctness of the chosen solution. Many side chains and a number of additional residues were added to the model at this stage. At the same time, parts of the model for which there was no clear density were removed.

After the first model-building step, the Hamburg Synchrotron data became available and were used in all subsequent refinement runs. Refinement was continued with a further rigid body refinement that allowed the three hemocyanin domains to move independently. Subsequently, positional refinement was started, slowly extending the resolution of the data from 3.2 to 2.7 Å. This refinement converged at an *R*-value of 38.5%. An unconstrained *B*-value refinement was used to reduce the weight of the bad parts of the structure before a new $2F_o - F_c$ map was calculated for a second model-building session. The map had significantly improved and a large number of atoms could again be added to the model. At this stage there was also much more density for the coppers, which were still not included in the model.

At increased resolution, TNT refinement (running on a VAX3200) had become quite slow. By switching to the program X-PLOR (Brünger et al., 1987) (running on a Convex C120) an enormous gain in speed could be obtained. In X-PLOR four slow-cooling MD runs were carried out. In each MD run the structure was instantaneously heated to 3,000 K followed by slow cooling at a rate of 1 K/fs. This is identical to the protocol suggested in the X-PLOR manual release 2.1 (Brünger, 1990). All reflections in the used resolution range (2.18–8.0 or 2.18–10.0 Å) were included in the refinement.

At 2.2 Å resolution the MD protocol also became very time-consuming. In addition, the number of significant structural corrections made during the MD runs was quite limited. Therefore refinement was continued with only energy minimization (EM) runs, followed by manual rebuilding. During rebuilding, attention was restricted mostly to the problematic regions indicated by unfavorable model geometry and/or high *B*-values. After each refinement cycle a constrained *B*-value refinement was carried out. Restraints used were: 2.5 Å^2 for bonded main-chain atoms and 4 Å^2 for main-chain atoms related by a bond angle. For the side-chain atoms these values were 5 Å^2 and 8 Å^2, respectively.

During refinement, special care was taken not to introduce bias into the geometry of the dinuclear copper site. For this reason no restraints on the copper–ligand and copper–copper geometry were imposed. Actually the coppers were included only after the last MD run. By this time the density for both coppers had become the strongest feature in the $F_o - F_c$ map, at about 14 times the standard deviation of the map. The coppers, introduced at the density maxima of the map by the pekpik option of the XTAL program (Hall & Stewart, 1987), were added

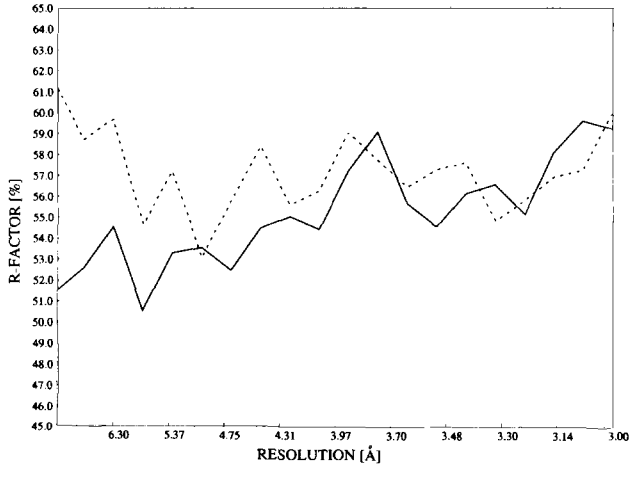

Fig. 15. *R*-factor as a function of resolution for the two allowed solutions of the molecular replacement problem. Continuous line = the correct solution; dashed line = the incorrect solution. Note that only the *R*-factor of the correct solution increases with increasing resolution.

to the model as free atoms without charge and with a 0.15-Å van der Waals radius. In this way the copper positions are effectively determined by the density alone.

Together with the copper ions, 64 water molecules were added to the model by the automated procedure described above. Individual inspection of the placed waters showed that this procedure was reliable, and subsequently added waters were no longer individually inspected. However, in later cycles the automated procedure also included waters in the second and even third water shell, which was judged unlikely to be correct. Therefore all waters were removed and replaced again in subsequent refinement cycles. The automatic procedure was used again, but now only for first shell waters. Strong $F_o - F_c$ density peaks not selected by this automatic procedure were inspected individually on a graphics device.

Acknowledgments

This work was supported by National Institutes of Health center grant ESO 1908, in part by NSF grant DMB-9004561 (K.A.M.), and by the Dutch Foundation for Chemical Research (SON) with financial aid from the Netherlands Organisation for Scientific Research (NWO).

References

Bak, H.J. & Beintema, J.J. (1987). *Panulirus interruptus* hemocyanin: The elucidation of the complete amino acid sequence of subunit a. *Eur. J. Biochem. 169*, 333–348.

Bonaventura, C., Sullivan, B., Bonaventura, J., & Bourne, S. (1974). CO binding by hemocyanins of *Limulus polyphemus*, *Busycon carica*, and *Callinectes sapidus*. *Biochemistry 13*, 4784–4789.

Brenowitz, M., Bonaventura, C., & Bonaventura, J. (1984). Self-association and oxygen-binding characteristics of the isolated subunits of *Limulus polyphemus* hemocyanin. *Arch. Biochem. Biophys. 230*, 238–249.

Brenowitz, M., Bonaventura, C., Bonaventura, J., & Gianazza, E. (1981). Subunit composition of a high molecular weight oligomer: *Limulus polyphemus* hemocyanin. *Arch. Biochem. Biophys. 210*, 748–761.

Brouwer, M., Bonaventura, C., & Bonaventura, J. (1977). Oxygen binding by *Limulus polyphemus* hemocyanin: Allosteric modulation by chloride ions. *Biochemistry 16*, 3897–3902.

Brouwer, M., Bonaventura, C., & Bonaventura, J. (1982). Chloride and pH dependence of cooperative interactions in *Limulus polyphemus* hemocyanin. In *Physiology and Biology of Horseshoe Crabs: Studies on Normal and Environmentally Stressed Animals* (Bonaventura, J., Bonaventura, C., & Tesh, S., Eds.), pp. 231–256. Alan R. Liss, Inc., New York.

Brouwer, M., Bonaventura, C., & Bonaventura, J. (1983). Metal ion interactions with *Limulus polyphemus* and *Callinectes sapidus* hemocyanins: Stoichiometry and structural and functional consequences of calcium (II), cadmium (II), zinc (II), and mercury (II) binding. *Biochemistry 22*, 4713–4723.

Brouwer, M. & Serigstad, B. (1989). Allosteric control in *Limulus polyphemus* hemocyanin: Functional relevance of interactions between hexamers. *Biochemistry 28*, 8819–8827.

Brown, J.M., Powers, L., Kincaid, B., Larrabee, J.A., & Spiro, T.G. (1980). Structural studies of the hemocyanin active site. 1. Extended X-ray absorption fine structure (EXAFS) analysis. *J. Am. Chem. Soc. 102*, 4210–4216.

Brünger, A.T. (1990). *XPLOR v2.1 Manual*. Yale University, New Haven, Connecticut.

Brünger, A.T., Kuriyan, J., & Karplus, M. (1987). Crystallographic *R* factor refinement by molecular dynamics. *Science 235*, 458–460.

Co, M.S. & Hodgson, K.O. (1981). Copper site of deoxyhemocyanin. Structural evidence from X-ray absorption spectroscopy. *J. Am. Chem. Soc. 103*, 3200–3201.

de Haas, F., Bijlholt, M.M.C., & Van Bruggen, E.F.J. (1991). An electron microscopic study of two-hexameric hemocyanins from the crab *Cancer pagurus* and the tarantula *Eurypelma californicum*: Determination of their quaternary structure using image processing and simulation models based on X-ray diffraction data. *J. Struct. Biol. 107*, 86–94.

Dooley, D.M., Scott, R.A., Ellinghaus, J., Solomon, E.I., & Gray, H.B. (1978). Magnetic susceptibility studies of laccase and oxyhemocyanin. *Proc. Natl. Acad. Sci. USA 75*, 3019–3022.

Eickman, N.C., Himmelwright, R.S., & Solomon, E.I. (1979). Geometric and electronic structure of oxyhemocyanin: Spectral and chemical correlations to met apo, half met, met, and dimer active sites. *Proc. Natl. Acad. Sci. USA 76*, 2094–2098.

Eyerle, F. & Schartau, W. (1985). Hemocyanins in spiders, XX[11]: Sulfhydryl groups and disulfide bridges in subunit d of *Eurypelma californicum* hemocyanin. *Biol. Chem. Hoppe-Seyler 366*, 403–409.

Freedman, T.B., Loehr, J.S., & Loehr, T.M. (1976). A resonance Raman study of the copper protein, hemocyanin. New evidence for the structure of the oxygen-binding site. *J. Am. Chem. Soc. 98*, 2809–2815.

Fujii, T., Sakurai, H., Izumi, S., & Tomino, S. (1989). Structure of the gene for the arylphorin-type storage protein SP 2 of *Bombyx mori*. *J. Biol. Chem. 264*, 11020–11025.

Gaykema, W.P.J., Hol, W.G.J., Vereijken, J.M., Soeter, N.M., Bak, H.J., & Beintema, J.J. (1984). 3.2 Å structure of the copper-containing, oxygen-carrying protein *Panulirus interruptus* haemocyanin. *Nature (Lond.) 309*, 23–29.

Gaykema, W.P.J., Volbeda, A., & Hol, W.G.J. (1986). Structure determination of *Panulirus interruptus* haemocyanin at 3.2 Å resolution. *J. Mol. Biol. 187*, 255–275.

Hall, S.R. & Stewart, J.M., Eds. (1987). *XTAL2.2 User's Manual*. Universities of Western Australia and Maryland.

Hazes, B. & Hol, W.G.J. (1992). Comparison of the hemocyanin β-barrel with other Greek key β-barrels: Possible importance of the "β-zipper" in protein structure and folding. *Proteins Struct. Funct. Genet. 12*, 278–298.

Herzberg, O. & Moult, J. (1991). Analysis of the steric strain in the polypeptide backbone of protein molecules. *Proteins Struct. Funct. Genet. 11*, 223–229.

Jackman, M.P., Hajnal, A., & Lerch, K. (1991). Albino mutants of *Streptomyces glaucescens* tyrosinase. *Biochem. J. 274*, 707–713.

Johansen, K. & Petersen, J.A. (1975). In *Eco-physiology of Estuarine Organisms* (Vernberg, F.J., Ed.), pp. 129–145. University of South Carolina Press, Columbia.

Jones, G., Brown, N., Manczak, M., Hiremath, S., & Kafatos, F.C. (1990). Molecular cloning, regulation, and complete sequence of a hemocyanin-related, juvenile hormone-suppressible protein from insect hemolymph. *J. Biol. Chem. 265*, 8596–8602.

Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr. 11*, 268–272.

Kraulis, P.J. (1991). MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr. 24*, 946–950.

Lamy, J., Lamy, J., Sizaret, P.-Y., Billiald, P., Jollés, P., Jollés, J., Feldmann, R.J., & Bonaventura, J. (1983). Quaternary structure of *Limulus polyphemus* hemocyanin. *Biochemistry 22*, 5573–5583.

Lang, W.H. & van Holde, K.E. (1991). Cloning and sequencing of *Octopus dofleini* hemocyanin cDNA: Derived sequences of functional units Ode and Odf. *Proc. Natl. Acad. Sci. USA 88*, 244–248.

Lesk, A.M. & Hardmann, K.D. (1985). Computer-generated pictures of proteins. *Methods Enzymol. 115*, 381–390.

Linzen, B., Soeter, N.M., Riggs, A.F., Schneider, H.-J., Schartau, W., Moore, M.D., Yokota, E., Behrens, P.Q., Nakashima, H., Takagi, T., Nemoto, T., Vereijken, J.M., Bak, H.J., Beintema, J.J., Volbeda, A., Gaykema, W.P.J., & Hol, W.G.J. (1985). The structure of arthropod hemocyanins. *Science 229*, 519–524.

Machin, P.A., Wonacott, A.J., & Moss, D. (1983). *Daresbury Lab. News. 10*, 3–9.

Maddaluno, J. & Giessner-Prettre, C. (1991). Nonempirical calculations on $(Cu^+)_2$-O_2: A possible model for oxyhemocyanin and oxytyrosinase active sites. *Inorg. Chem. 30*, 3439–3445.

Magnus, K.A., Lattman, E.E., Volbeda, A., & Hol, W.G.J. (1991). Hexamers of subunit II from *Limulus* hemocyanin (a 48-mer) have the same quaternary structure as whole *Panulirus* hemocyanin molecules. *Proteins Struct. Funct. Genet. 9*, 240–247.

McKee, V., Zvagulis, M., Dagdigian, J.V., Patch, M.G., & Reed, C.A. (1984). Hemocyanin models: Synthesis, structure, and magnetic properties of a binucleating copper (II) system. *J. Am. Chem. Soc. 106*, 4765–4772.

Messerschmidt, A., Rossi, A., Ladenstein, R., Huber, R., Bolognesi, M., Gatti, G., Marchesini, A., Petruzzelli, R., & Finazzi-Agró, A. (1989). X-ray crystal structure of the blue oxidase ascorbate oxidase from zucchini. *J. Mol. Biol. 206*, 513–529.

Morris, A.L., MacArthur, M.W., Hutchinson, E.G., & Thornton, J.M. (1992). Stereochemical quality of protein structure coordinates. *Proteins Struct. Funct. Genet. 12*, 345–364.

Mousseron, S. (1986). Thesis, Paris. (Cited in Linzen et al., 1985.)

Nakashima, H., Behrens, P.Q., Moore, M.D., Yokota, E., & Riggs, A.F. (1986). Structure of hemocyanin II from the horseshoe crab, *Limulus polyphemus*. *J. Biol. Chem. 261*, 10526–10533.

Naumann, U. & Scheller, K. (1991). Complete cDNA and gene sequence of the developmentally regulated arylphorin of *Calliphora vicina* and its homology to insect hemolymph proteins and arthropod hemocyanins. *Biochem. Biophys. Res. Commun. 177*, 963–972.

Nemoto, T. & Takagi, T. (1983). Report at the 56th Annual Meeting of the Japanese Biochemical Society, 29 September to 2 October. (Cited in Naumann & Scheller, 1991.)

Neuteboom, B., Dokter, W., van Gijsen, J., Rensink, H., de Vries, J., & Beintema, J.J. (1989). Partial amino acid sequence of a hemocyanin subunit from *Palinurus vulgaris*. *Comp. Biochem. Physiol. 94B*, 593–597.

Neuteboom, B., Jekel, P.A., & Beintema, J.J. (1992). Primary structure of hemocyanin subunit c from *Panulirus interruptus*. *Eur. J. Biochem. 206*, 243–249.

Pflugrath, J.W. & Messerschmidt, A. (1986) *MADNES User's Guide*. Max-Planck Institute für Biochemie, Martinsried, Germany.

Ramakrishnan, C. & Ramachandran, G.N. (1965). Stereochemical criteria for polypeptide and protein chain conformation. *Biophys. J. 5*, 909–933.

Rao, S.T. & Rossmann, M.G. (1973). Comparison of super-secondary structures in proteins. *J. Mol. Biol. 76*, 241–256.

Read, R.J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A42*, 140–149.

Ross, P.K. & Solomon, E.I. (1991). An electronic structural comparison of copper–peroxide complexes of relevance to hemocyanin and tyrosinase active sites. *J. Am. Chem. Soc. 11*, 3246–3259.

Sakurai, H., Fujii, T., Izumi, S., & Tomino, S. (1988). Complete nucleotide sequence of gene for sex-specific storage protein of *Bombyx mori*. *Nucleic Acids Res. 16*, 7717–7718.

Schneider, H.-J., Voll, W., Lehmann, L., Grißhammer, R., Goettgens, A., & Linzen, B. (1986). Partial amino acid sequence of crayfish (*Astacus leptodactylus*) hemocyanin. In *Invertebrate Oxygen Carriers* (Linzen, B., Ed.), pp. 173–176. Springer-Verlag, Berlin.

Soeter, N.M., Jekel, P.A., Beintema, J.J., Volbeda, A., & Hol, W.G.J. (1987). Primary and tertiary structures of the first domain of *Panulirus interruptus* hemocyanin and comparison of arthropod hemocyanins. *Eur. J. Biochem. 169*, 323–332.

Sonner, P., Voit, R., & Schartau, W. (1990). Primary structure of subunit b of *Eurypelma californicum* hemocyanin. In *Invertebrate Dioxygen Carriers* (Préaux, G. & Lontie, R., Eds.), pp. 77–80. Leuven University Press, Leuven, Belgium.

Sorrell, T.N. & Jameson, D.L. (1982). An explanation for the observed stoichiometry of carbon monoxide binding to hemocyanin. *J. Am. Chem. Soc. 104*, 2053–2054.

Sorrell, T.N. & Jameson, D.L. (1983). Synthesis, structure, and reactivity of monomeric two-coordinate copper (I) complexes. *J. Am. Chem. Soc. 105*, 6013–6018.

Strynadka, N.C.J. & James, M.N.G. (1991). Towards an understanding of the effects of calcium on protein structure and function. *Curr. Opin. Struct. Biol. 1*, 905–914.

Sullivan, B., Bonaventura, J., & Bonaventura, C. (1974). Functional differences in the multiple hemocyanins of the horseshoe crab, *Limulus polyphemus* L. *Proc. Natl. Acad. Sci. USA 71*, 2558–2562.

Thamann, T.J., Loehr, J.S., & Loehr, T.M. (1977). Resonance Raman study of oxyhemocyanin with unsymmetrically labeled oxygen. *J. Am. Chem. Soc. 99*, 4187–4189.

Tronrud, D.E., Ten Eyck, L.F., & Matthews, B.W. (1987). An efficient general-purpose least-squares refinement program for macromolecular structures. *Acta Crystallogr. A43*, 489–501.

van Bruggen, E.F.J, Schutter, W.G., Vam Breemen, J.F.L., Bijlholt, M.M.C., & Wichertjes, T. (1982). Arthropodan and molluscan haemocyanins. In *Electron Microscopy of Proteins* (Harris, M., Ed.), pp. 11–28. Academic Press, New York.

van Holde, K.E. & Miller, K.I. (1982). Haemocyanins. *Q. Rev. Biophys. 15*, 1–129.

Voit, R. & Feldmaier-Fuchs, G. (1990). Arthropod hemocyanins: Molecular cloning and sequencing of cDNAs encoding the tarantula hemocyanin subunits a and e. *J. Biol. Chem. 265*, 19447–19452.

Volbeda, A. & Hol, W.G.J. (1989a). Crystal structure of hexameric haemocyanin from *Panulirus interruptus* refined at 3.2 Å resolution. *J. Mol. Biol. 209*, 249–279.

Volbeda, A. & Hol, W.G.J. (1989b). Pseudo 2-fold symmetry in the copper-binding domain of arthropodan haemocyanins. Possible implications for the evolution of oxygen transport proteins. *J. Mol. Biol. 206*, 531–546.

Voll, W. & Voit, R. (1990). Characterization of the gene encoding the hemocyanin subunit *e* from the tarantula *Eurypelma californicum*. *Proc. Natl. Acad. Sci. USA 87*, 5312–5316.

Willott, E., Wang, X.-Y., & Wells, M.A. (1989). cDNA and gene sequence of *Manduca sexta*, an aromatic amino acid-rich larval serum protein. *J. Biol. Chem. 264*, 19052–19059.

Woolery, G.L., Powers, L., Winkler, M., Solomon, E.I., & Spiro, T.G. (1984). EXAFS studies of binuclear copper site of oxy-, deoxy-, metaquo-, metfluoro-, and metazidohemocyanin from arthropods and molluscs. *J. Am. Chem. Soc. 106*, 86–92.

Yen Fager, L. & Alben, J.O. (1972). Structure of the carbon monoxide binding site of hemocyanins studied by Fourier transform infrared spectroscopy. *Biochemistry 11*, 4786–4792.