9 August 1985, Volume 229, Number 4713 SCIENCE

# The Structure of Arthropod Hemocyanins

B. Linzen, N. M. Soeter, A. F. Riggs, H.-J. Schneider
W. Schartau, M. D. Moore, E. Yokota, P. Q. Behrens
H. Nakashima, T. Takagi, T. Nemoto, J. M. Vereijken, H. J. Bak
J. J. Beintema, A. Volbeda, W. P. J. Gaykema, W. G. J. Hol

*Summary.* Hemocyanins are large multi-subunit copper proteins that transport oxygen in many arthropods and molluscs. Comparison of the amino acid sequence data for seven different subunits of arthropod hemocyanins from crustaceans and chelicerates shows many highly conserved residues and extensive regions of near identity. This correspondence can be matched closely with the three domain structure established by x-ray crystallography for spiny lobster hemocyanin. The degree of identity is particularly striking in the second domain of the subunit that contains the six histidines which ligate the two oxygen-binding copper atoms. The polypeptide architecture of spiny lobster hemocyanin appears to be the same in all arthropods. This structure must therefore be at least as old as the estimated time of divergence of crustaceans and chelicerates, about 540 to 600 million years ago.

Three classes of oxygen-carrying proteins occur in animals: hemoglobins, hemerythrins, and hemocyanins. The site of oxygen binding is entirely different in each: a single Fe(II) contained in a heme group in hemoglobins, two Fe(II) atoms bound by amino acid side chains in hemerythrins, and two Cu(I) atoms bound directly to side chains in hemocyanins. Hemocyanins are freely dissolved in the hemolymph of many molluscs and arthropods, but the quaternary structure and the ratio of copper to protein is very different in the two phyla. Molluscan hemocyanins are large cylindrical molecules with 10 to 20 subunits and a mass of up to 9 million daltons. Subunits have eight domains, each of which has a pair of copper atoms and a mass of 50,000 daltons. Arthropod hemocyanins vary in size from 500,000 to about 3.5 million daltons and are hexamers or oligohexamers built from individual subunits of about 75,000 daltons (*1*). Although this difference in architecture has led to the suggestion that they may have originated independently, the oxygen binding sites appear to be quite similar spectroscopically. Binuclear copper sites similar to those of both arthropod and molluscan hemocyanins are found in tyrosinase, ceruloplasmin, laccase, and ascorbate oxidase (*2*).

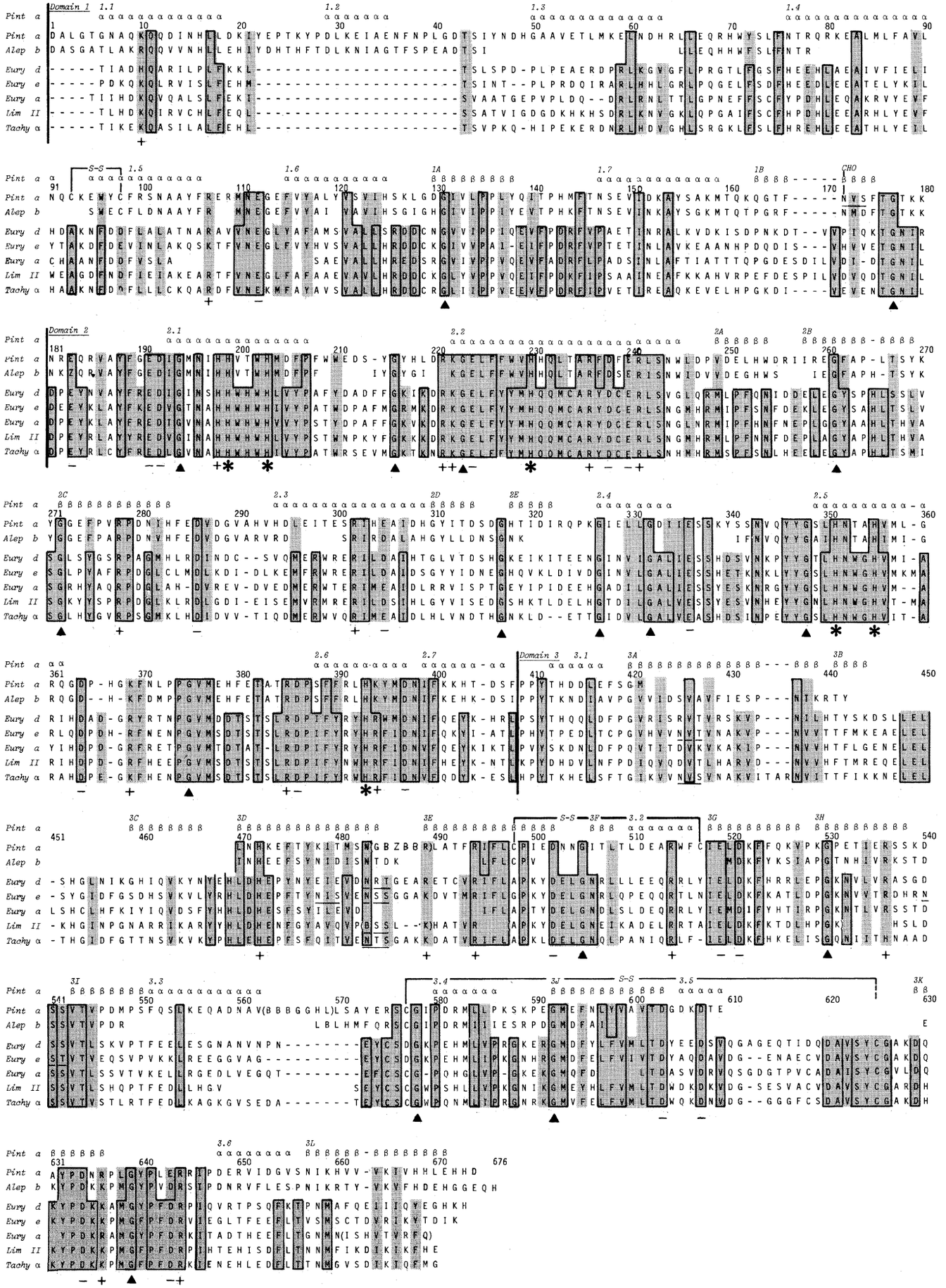
Although hemoglobins and hemerythrins (*3*, *4*) have well-established primary, tertiary, and quaternary structures, the structure of hemocyanins has remained elusive until very recently because of the great length of the polypeptide chains, the heterogeneity of the subunits, and the complex quaternary structures. Much structural information has been collected recently for the chelicerate and crustacean hemocyanins. The quaternary structures of two large chelicerate hemocyanins were solved by immunoelectron microscopy in 1981 (*5*). This was followed by a wealth of primary structure information: the first two complete amino acid sequences of chains d and e of hemocyanin from *Eurypelma californicum* (tarantula) were reported in 1983 (*6*); a third chain, the α subunit from *Tachypleus tridentatus* (Japanese horseshoe crab) has also been completed (*7*). Four other sequences, including two of crustacean hemocyanins, are far advanced.

The three-dimensional structure of hemocyanin of the spiny lobster, *Panulirus interruptus*, has been determined very recently at a resolution of 3.2 Å (*8*). We can now correlate the primary structures with the three-dimensional data on a broader basis, and provide an integrated description of this class of proteins for the first time.

## Amino Acid Sequence Comparisons

The amino acid sequences of the seven hemocyanin subunits from two crustaceans and three chelicerates have been aligned in Fig. 1. Many segments are identical or nearly so in the different sequences. Examples are residues 190 to 207 and 220 to 243. Regions of close correspondence are widespread. The polypeptide chain lengths are roughly the same: the chelicerate hemocyanins have about 625 residues and the crustacean hemocyanins about 660. The considerable differences in chain length estimated from sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis (*9*) have not been confirmed by the sequence and x-ray data. These differences must result from variations in the binding of SDS. So far, 117 positions (17 percent) are identical for all seven chains and 212 positions (33 percent) for the chelicerate chains. These figures increase to 195 (29 percent) or 298 (46 percent), respectively, if we include po-

B. Linzen, H.-J. Schneider, and W. Schartau are from Zoologisches Institut, Universität München, Luisenstrasse 14, 8000 München 2, Federal Republic of Germany; N. M. Soeter, J. M. Vereijken, H. J. Bak, J. J. Beintema, A. Volbeda, W. P. J. Gaykema, and W. G. J. Hol are from Biomolecular Study Centre (BIOS), University of Groningen, Nijenborgh 16, 9747 AG Groningen, The Netherlands; A. F. Riggs, M. D. Moore, E. Yokota, P. Q. Behrens, and H. Nakashima are from the Department of Zoology, University of Texas, Austin 78712, and T. Takagi and T. Nemoto are from Biological Institute, Faculty of Science, Tohoku University, Sendai, Japan 980.


Pint a
Alep b
Eury d
Eury e
Eury a
Lim II
Tachy α
Domain 1
Domain 2
Domain 3
S-S
CHO

sitions with isofunctional residues. These are shaded gray in Fig. 1.

Some remarkable features of this comparison are as follows. Identities are often clustered in contiguous regions, for example, positions 176 to 207, 216 to 243, 346 to 356, 374 to 400, 573 to 609, and 620 to 646. The most strongly conserved regions are those near the histidine residues that have been identified as copper ligands (*8*).

The glycine residues as elements of chain bending are extremely conserved. More than half of all the glycines shown in Fig. 1 occur at the same positions in 5, 6, or 7 subunits; only about 10 percent have no counterparts at the same or neighboring positions of another subunit. Many of the completely conserved glycines occur at or close to the beginning of a β strand or an α helix.

The similarity of the chains becomes more striking if the hydrophobic or hydrophilic nature of the amino acid side chains is considered. Hydrophilicity profiles for three hemocyanins resemble each other greatly (Fig. 2). For example, exchanging an acidic for a basic residue would not alter the hydrophilic nature of a position. This suggests that the microenvironment of every residue along the chain is a rather stable characteristic which is resistant to random change.

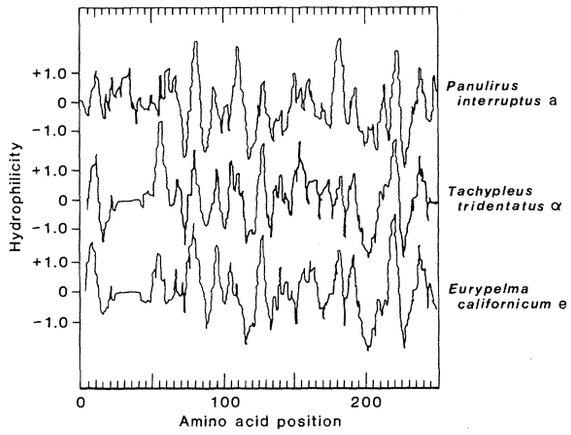



Fig. 2. Hydrophilicity profiles of positions 1 to 250 for three arthropod hemocyanin subunits. From top to bottom: *Panulirus interruptus* a, *Tachypleus tridentatus* α, and *Eurypelma californicum* e. Hydrophilicity values were calculated according to Hopp and Woods (*18*) and deletions were given the value 0.0. Means were calculated for groups of six residues, and the group shifted in steps of one residue. The degree of similarity is high in spite of large evolutionary distance between chains (divergence some 400 to 600 million years ago).

These observations taken together leave no doubt that the general morphology of the subunits of arthropod hemocyanins has been highly conserved. The three-dimensional structure established for the *Panulirus* hemocyanin subunit, its organization into three domains, and its arrangement of α helices and β strands including the β barrel of domain 3 must have been conserved for at least 540 to 600 million years, the estimated time of divergence of chelicerate and crustacean hemocyanins (*14*).

## Three-Dimensional Structure

The morphology of the hexamer (molecular weight, 450,000) and of the individual subunit of *Panulirus interruptus* hemocyanin is shown in Fig. 3. This monohexameric structure is the smallest functional hemocyanin known. A similar hexamer can be considered as the building block of all multihexameric hemocyanins. *Eurypelma* hemocyanin is a tetramer of hexamers (24 chains) and *Limulus* hemocyanin comprises 48 chains (2 × 24). The hexamer consists of two trimers facing each other in a staggered arrangement (Fig. 3A). Each subunit may be viewed as having a convex "back surface" facing the other two subunits in a trimer, while the two surface regions perpendicular to the threefold axis of the hexamer are either completely exposed to the solvent ("top-surface") or make contact with the other trimer ("bottom surface"). The contacts between the six subunits are complex, but can be briefly summarized as follows. The contacts between the subunits of one trimer are virtually all hydrophilic. Each subunit of one trimer has contact mainly with only one subunit of the other trimer, related by a twofold axis running close and roughly parallel to helix 1.7. These contacts are hydrophilic as well as hydrophobic and are closer than the contacts within a trimer. Therefore, the hexamer may be described as a trimer of dimers, rather than as a dimer of trimers. The dimer may have been conserved for a very long time among the arthropod hemocyanins because residues $Gly^{177}$, $Gly^{261}$, $Pro^{279}$, and $Phe^{370}$ at the interface are highly conserved (Fig. 1). In contrast, the interactions between the members of a trimer are more variable.

The subunits have three domains, two of which are mainly α-helical while the third has a β barrel and two large loops which make contact with both the other domains (Fig. 3B). The active site is located in the center of the second domain.

## Structure of the Subunit Domains

*Domain 1.* Domain 1 (Fig. 1, residues 1 to 180 and Fig. 3C) is the most variable of the three domains. This variability holds especially for the $NH_2$-terminal region. A number of conserved residues are also present. One of the latter, $Phe^{75}$, the residue of domain 1 closest to the oxygen binding site, probably functions to stabilize the hydrophobic core of domain 2. A larger number of conserved residues begins at position 110 close to the $NH_2$-terminus of helix 1.6. This buried helix is strongly hydrophobic and is in contact with three other helices (1.1, 1.4, and 1.5) within the domain. It may be important that helix 1.6 also contacts helix 2.1 of domain 2 which provides two copper ligands and hence may pick up any movements of the latter upon oxygenation.

Domain 1 contains a quite interesting appendix consisting of β strand 1B and α helix 1.7. This appendix makes contact

Fig. 1 (opposite page). Alignment of arthropod hemocyanin amino acid sequences and secondary structure of *Panulirus interruptus* hemocyanin. Complete and partial sequences are shown for two crustacean species, *Panulirus interruptus*, a spiny lobster, subunit a (*8*, *15*) and the crayfish, *Astacus leptodactylus*, subunit b (*6*), and of the following Chelicerata: *Eurypelma californicum*, a tarantula, subunit d, e, and a (*6*), and two horseshoe crabs, *Limulus polyphemus*, subunit II (*16*), and *Tachypleus tridentatus*, subunit α (*7*). The one-letter code for amino acid residues is used (*17*). Regions with identical or nearly identical residues have been aligned so as to minimize the number of gaps. Dashes represent gaps introduced for alignment, blank spaces are parts not yet determined, and sequences in brackets are presumptive. For the purpose of this article a special numbering scheme has been used. The secondary structure elements (α helices 1.1–3.6 and β strands 1A–3L), S–S bridges, carbohydrate attachment site (CHO), copper ligands (*), and domain borders of *Panulirus interruptus* hemocyanin are indicated. Positions with identical residues in all or in the chelicerate sequences are boxed (also if B and Z had been placed instead of D, N, E, and Q, respectively, or if I and L could not be distinguished in the *Astacus* b, or in the *Eurypelma* d and a sequences). Boxes are open if no sequence data of Crustacea are available. Positions with isofunctional residues are shaded (T and S; D and E; K, R, and H; Y, F, and W; I, V, L, and M). The symbols beneath the sequences indicate positively charged (+) or negatively charged (−) residues or conserved glycines (▲) in all sequences. Positions with the structural requirements for carbohydrate attachment (N X T/S) are underlined.

with the first large extension of domain 3: 1B forms a β sheet with β strands 3B and 3C. Being at the "bottom" of the subunit, helix 1.7 faces the "bottom surface" of the other trimer in the hexamer, that is, it participates in the extensive contacts between the trimers.

The significance of the presence of a carbohydrate side chain in domain 1 is unknown. Carbohydrate has been found only in *Panulirus* hemocyanin where it is attached to $Asn^{172}$. This site is different even in the hemocyanin of the closely related crayfish *Astacus*. No carbohydrate has been found in any of the other hemocyanins sequenced, although several Asn-X-Ser or Thr sequences occur (underlined in Fig. 1). Such sequences are often associated with carbohydrate binding.

*Domain 2*. This domain (Fig. 1, residues 181 to 408, and Fig. 3D) contains the binuclear copper site, which binds oxygen (see below). The crustacean hemocyanins have a cluster of basic residues around position 180. This site is easily accessible for proteolytic cleavage in *Panulirus* hemocyanin (*10*). Another cluster of charged residues occurs in the chelicerate hemocyanins at positions 217 to 224 at the beginning of helix 2.2. These include $Arg^{221}$ and $Glu^{224}$ which form a well-defined salt bridge in *Panulirus* which is completely conserved. Four conserved charged residues occur close to each other in the COOH-terminal half of helix 2.2. Two of these form conserved salt bridges: $Asp^{238}$ with $Lys^{394}$ and $Arg^{241}$ with $Glu^{183}$. The former one may be particularly important because residue 394 is next to one of the histidine ligands of the oxygen binding site. Conservation of salt bridges has been observed in highly conserved proteins (*11*).

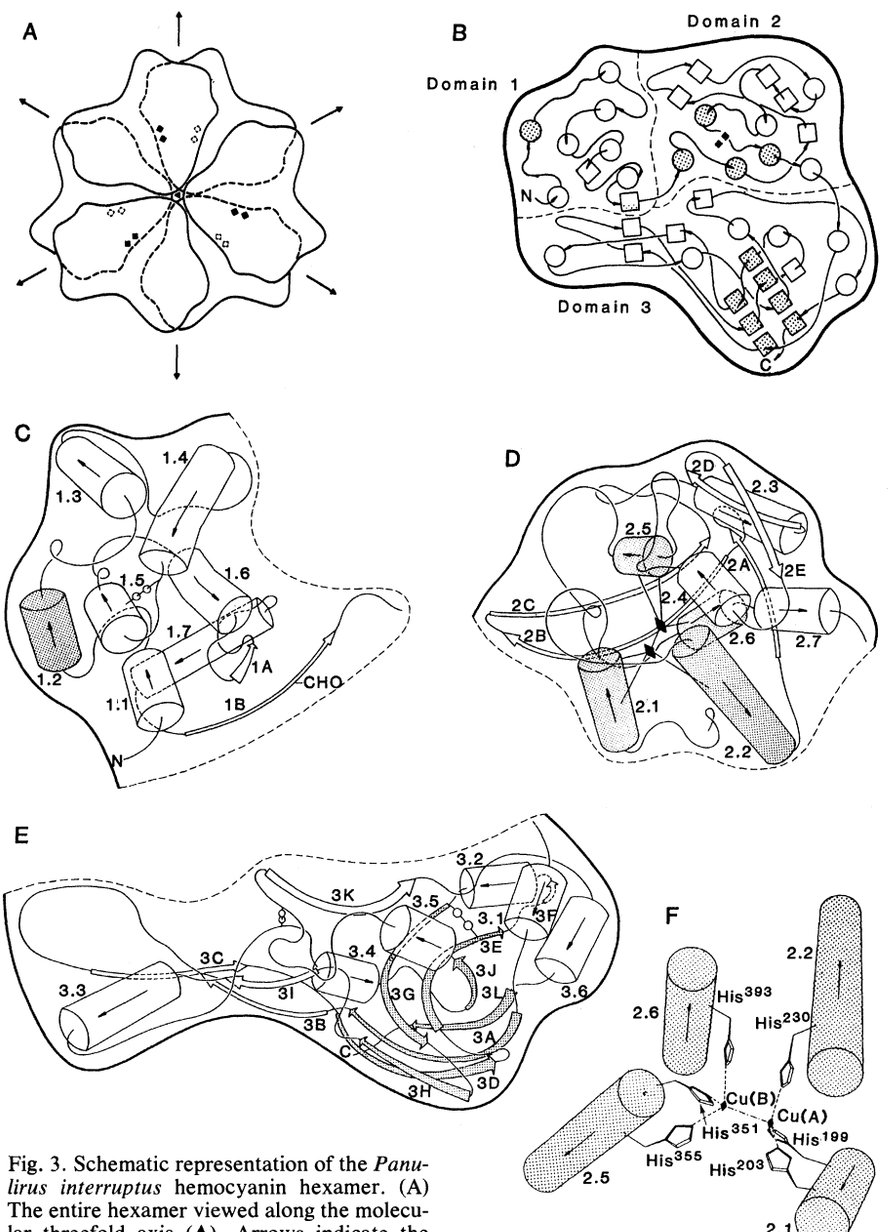



Fig. 3. Schematic representation of the *Panulirus interruptus* hemocyanin hexamer. (A) The entire hexamer viewed along the molecular threefold axis (▲). Arrows indicate the three molecular twofold axes, running perpendicular to the threefold axis. The parts of each molecular twofold axis across which the most intimate subunit contacts occur are indicated with solid lines and the other parts with dashed lines. Diamonds indicate the copper ions in the upper (◆) and the lower (◇) trimer. (B) One subunit with α helices (○), β strands (□), and the course of the polypeptide chain. The three domains are separated by dashed lines. The α helix that is missing in the Chelicerata and part of the β strand containing the carbohydrate moiety in domain 1, the four α helices providing Cu ligands in domain 2, and the β barrel strands in domain 3 are shaded. (C) Domain 1 with the two-character identifiers for each secondary structure element. The first character is the domain number. The second character is a number for α helices and a letter for β strands. Helices are represented by cylinders, β strands by arrows, and the disulfide bridges by –o–o–. (D) Domain 2. Diamonds indicate copper ions. (E) Domain 3. (F) The four helices surrounding the binuclear copper site and providing the six histidine copper ligands. Some contacts between domains are not possible to depict in the two-dimensional representations shown in parts B, C, D, and E.

## Binuclear Copper Site in Domain 2

The active site is located in the center of the second domain (Fig. 3, B and D). The two copper atoms A and B are each ligated by three histidines (Fig. 3F): the first one by $His^{199}$, $His^{203}$, and $His^{230}$ and the second by $His^{351}$, $His^{355}$, and $His^{393}$ (the histidines are indicated by * in Fig. 1). The six histidines are provided by four α helices, 2.1, 2.2, 2.5, and 2.6, which are all highly conserved and in the interior of the molecule.

Helix 2.1 at copper site A contains a remarkable histidine-tryptophan cluster (Fig. 1, residues 198 to 203). A tryptophan is also present between the two histidines of helix 2.5 at copper site B in chelicerate hemocyanins (Fig. 1, residue 353). Most of the tryptophan residues in the sequence are near the copper ligands (four of five in *Eurypelma* d, three of five in *Eurypelma* e, three of four in *Eurypelma* a, four of eight in *Limulus* II, and four of eight in *Tachypleus* α). This proximity explains the exceptionally strong quenching of tryptophan fluorescence that is observed in *Eurypelma* hemocyanin (*12*). The relative number of tryptophan residues near the active site is lower in crustacean hemocyanins. The tryptophans may serve to anchor the active site helices in the surrounding structure. A large number of aromatic residues in addition to these tryptophans can be found in the neighborhood of the copper binding site in *Panulirus* hemocy-

anin; many of these are conserved. Although several tyrosines are near the copper ligands in the sequence, none comes close to the coppers in the three-dimensional structure.

*Domain 3.* Domain 3 (Fig. 1, residues 409 to 676, and Fig. 3E) contains a seven-stranded β barrel. The barrel shows a close topological correspondence to the antiparallel β barrels observed in Cu-Zn, superoxide dismutase, immunoglobulin (*8*), and actinoxanthine (*13*). The β strands in the barrel belong to the most conserved parts of domain 3 (Fig. 1), which indicates that they are a universal feature of arthropod hemocyanins.

Two long loops extend from this barrel. The first makes contact with domain 1 as already discussed. The second one comprises residues 540 to 590 which includes the segments β 3I, α 3.3, and α 3.4. The region in this loop between α 3.3 and α 3.4 is highly variable in length and in nature of residues. *Panulirus interruptus* hemocyanin a contains three S-S bridges and no free SH-groups and chelicerate hemocyanins contain both S-S bridges and SH-groups. Comparison of the sequences shows that only the third disulfide bridge of *Panulirus* hemocyanin may be conserved in the other sequences although not at exactly the same positions: 575 to 625 or 577 to 625 (Fig. 1). This bridge links together two highly conserved regions occurring at positions 573 to 603 and 620 to 646 near the COOH-terminus. Both segments are involved in contacts with domain 2 (especially α 2.1 and α 2.2) as well as domain 1 (especially β 1A). Residues 592 to 603 belong to the β barrel (3J).

## Comparison of Crustacean and Chelicerate Hemocyanins

What distinguishes crustacean from chelicerate hemocyanins? One striking difference is the occurrence of a 21-residue loop from $Tyr^{22}$ to $Asp^{42}$ which contains helix 1.2. This loop is completely absent from chelicerate hemocyanins. Helix 1.2 in *Panulirus* hemocyanin bulges out from the surface of domain 1, supported only by two straps which should give it greater flexibility than that exhibited by other elements of the secondary structure. This might explain why it is the most poorly defined helix in the electron density map. The helix makes no contacts with any other domain or subunit, and its excision can probably be accommodated without losing the essential functions. A second, shorter loop occurs in a highly variable segment just before the conserved disulfide bridge linking $Cys^{577}$ to $Cys^{625}$. This loop is part of the second arm in *Panulirus* hemocyanin which stretches out from domain 3 and embraces domain 1; it also faces the outside. The shorter arm in chelicerate hemocyanins could have the effect of tilting helix 3.3 upward. Alternatively helix 3.3 could be partially unwound and be one turn shorter. Finally, the two crustacean hemocyanins show an extension of five amino acids at the $NH_2$-terminus; this extension is absent from the chelicerate hemocyanins.

A second major distinction between crustacean and chelicerate hemocyanins appears in the structure of the active site. The chelicerate sequence, -His-His*-Trp-His-Trp-His*-, for the first two ligands of copper A is different in crustacean hemocyanins. The first Trp (position 200) is replaced by Val and the third His (position 201) by Thr. In the copper B site in helix 2.5 (Fig. 1) $Trp^{353}$ is replaced by Thr.

Table 1. Comparison of the percent differences between amino acid sequences of domains and complete subunits of hemocyanins. The number of differences between amino acid sequences for each domain were counted for those hemocyanin chains with complete or nearly completed sequences for that domain. The alignments presented in Fig. 1 were used. The table gives the percent differences between sequences derived according to Dayhoff *et al.* (*19*). Positions where one sequence has an amino acid and the other a deletion were taken into account. Positions where either sequence is still unknown or where both contain deletions are ignored. A residue that may be either aspartic acid or asparagine, or glutamic acid or glutamine, were not considered different from either possible alternative.

**Domain 1**

| | Pint | Eury | | | Lim II |
|---|---|---|---|---|---|
| | a | d | e | a | |
| Eury d | 81 | | | | |
| Eury e | 78 | 56 | | | |
| Eury a | 84 | 61 | 60 | | |
| Lim II | 80 | 58 | 54 | 50 | |
| Tachy α | 79 | 46 | 46 | 60 | 52 |
| NPC* | 166–180 | 141–154 | | | |

**Domain 2**

| | Pint | Eury | | | Lim II |
|---|---|---|---|---|---|
| | a | d | e | a | |
| Eury d | 59 | | | | |
| Eury e | 59 | 35 | | | |
| Eury a | 59 | 37 | 33 | | |
| Lim II | 63 | 37 | 32 | 31 | |
| Tachy α | 58 | 38 | 37 | 37 | 36 |
| NPC | 226–227 | 224–227 | | | |

**Domain 3**

| | Eury | | Lim II |
|---|---|---|---|
| | d | e | |
| Eury e | 49 | | |
| Lim II | 51 | 51 | |
| Tachy α | 53 | 46 | 45 |
| NPC | 238–257 | | |

**Complete subunit**

| | Eury | | Lim II |
|---|---|---|---|
| | d | e | |
| Eury e | 45 | | |
| Lim II | 47 | 45 | |
| Tachy α | 46 | 43 | 43 |
| NPC | 618–632 | | |

*NPC, the numbers of positions compared.

## Multihexameric Hemocyanins

*Panulirus* hemocyanin crystals contain two similar subunits, a and b (*8*). The large chelicerate hemocyanins have seven to eight different kinds of subunits which occupy specific sites in the hexamers so that multiple lock-and-key structures must be postulated for the points of contact. All chelicerate chains have additional residues at positions 167 to 171. These residues are near the intersubunit contact area. The longest additions in this region, which are also almost identical, are in *Eurypelma* a and *Limulus* II. It is striking that these two subunits occupy identical positions in the native oligomer and are immunologically related (*14*).

## Molecular Evolution

The differences between amino acid sequences in the separate domains and in the complete hemocyanin chains are summarized in Table 1. These data show that the three domains have been conserved to quite different extents. The oxygen-binding domain 2 contains the largest number of identical residues; the other two domains vary more. Domains 2 and 3, of all chelicerate hemocyanin chains, vary to approximately the same extent (7 to 8 percent). However, greater variation exists in domain 1 (15 percent). The comparison of the *Eurypelma* d and e chains with the *Tachypleus* α chain shows the least difference. The *Eurypelma* a and *Limulus* II comparison shows only a slightly larger difference.

During hemoglobin evolution the gene duplication giving rise to the α and β chains preceded the divergence of teleosts and tetrapods, and it is not difficult to recognize a hemoglobin chain as ei-

Table 2. Evolutionary rates of separate domains and complete chains of hemocyanins. The percent differences (± standard deviation) from Table 1 are converted to "accepted point mutations per 100 residues" (PAM's) according to Dayoff (*20*, p. 375) and rates were calculated by assuming a common ancestry of about 400 million years ago for the Chelicerata (*21*) and of about 600 million years ago for the divergence of the Crustacea and the Chelicerata (*22*). The hemoglobin and myoglobin data are from Dayhoff (*20*, p. 3).

| Item | Difference (%) | PAM's | Ancestor (million years ago) | PAM's per $10^8$ years |
|---|---|---|---|---|
| Hemocyanin | | | | |
| Domain 1 | | | | |
| Pint a—Chelicerata | 80 ± 2 | 234 | 600 | 19.5 |
| Chelicerate chains | 54 ± 6 | 91 | 400 | 11.4 |
| Domain 2 | | | | |
| Pint a—Chelicerata | 60 ± 2 | 108 | 600 | 9.0 |
| Chelicerate chains | 35 ± 2 | 47 | 400 | 5.9 |
| Domain 3 | | | | |
| Chelicerate chains | 49 ± 3 | 77 | 400 | 9.6 |
| Complete subunit | | | | |
| Chelicerate chains | 45 ± 2 | 67 | 400 | 8.4 |
| Hemoglobin (α and β chains, most mammalian orders) | | | | 12.0 |
| Myoglobin (most mammalian orders) | | | | 8.9 |

ther α or β from its sequence. Such sister proteins evolving in parallel during the evolution of species are called paralogous proteins. The situation is different for the hemocyanin chains of the Chelicerata. Without further analysis of the sequence data we must assume that gene duplications leading to separate chains in one species and the divergence between the Arachnida (*Eurypelma*) and Xiphosura (*Limulus* and *Tachypleus*) occurred at close to the same time. For instance, it is evident that the *Euryplema* chains a, d, and e are paralogous gene products with a common ancestry going back to the divergence of all chelicerate hemocyanin chains. As already mentioned, the chain of *Eurypelma* a and that of *Limulus* II are immunologically related and occupy identical positions in the hexamer. This relationship is also shown by the difference matrices of both domains 1 and 2 (Table 1).

Some evolutionary rates are presented in Table 2. These do not differ much from those observed for hemoglobin and myoglobin. The data suggest that, although the hemocyanin molecules are much more complicated structures than those of the globins, both types of oxygen carriers were subject to similar selection pressure.

## Conclusion

It may be appreciated from the foregoing that these hemocyanins represent a class of extremely complicated oxygen carriers. The central domain with the active site is two-thirds larger than a hemoglobin subunit, and is flanked by two additional domains of similar size. We suggest that these flanking domains pick up the movement of the central domain during oxygenation and convey it to the other subunits in the hexamer. Domain 3, the largest, is also probably essential for the integrity of the structure of the subunit by virtue of its long arms which stretch out past domain 2 to reach domain 1. The three domains are very different. No evidence exists to suppose that the domains are (or contain) repeating sequences.

Hemocyanins are certainly less efficient than hemoglobins in terms of the size of a functional unit. Each hemocyanin subunit comprises some 10,000 atoms to grasp and release just one molecule of oxygen—and up to 48 such subunits are combined in the native aggregate. Nature certainly had the alternative of using hemoglobin which was in the inventory of the articulate ancestors and which appears in many Crustacea. However, globin genes in some organisms may have been absent or may have lacked the capability to produce multisubunit structures exhibiting allosteric control, a requirement of efficient $O_2$ transport. Any gene that evolved to fit this function might be used.

### References and Notes


1. K. E. van Holde and K. I. Miller, *Q. Rev. Biophys.* **15**, 1 (1982); H. D. Ellerton, N. F. Ellerton, H. A. Robinson, *Progr. Biophys. Mol. Biol.* **41**, 143 (1983).
2. E. I. Solomon, in *Copper Proteins*, T. G. Spiro, Ed. (Wiley, New York, 1981), p. 43.
3. R. E. Dickerson and J. Geis, *Hemoglobin: Structure, Function, Evolution and Pathology* (Benjamin, Menlo Park, Calif., 1983).
4. G. L. Klippenstein, J. W. Holleman, I. M. Klotz, *Biochemistry* **7**, 3868 (1968); R. E. Stenkamp, L. C. Sieker, L. H. Jensen, J. E. McQueen, *ibid.* **17**, 2499 (1978); W. A. Hendrickson and J. L. Smith, in *Invertebrate Oxygen-Binding Proteins*, J. Lamy and J. Lamy, Eds. (Dekker, New York, 1981), p. 343; I. M. Klotz and D. M. Kurtz, Jr., *Acc. Chem. Res.* **17**, 16 (1984).
5. J. Lamy, M. M. C. Bijlholt, P.-Y. Sizaret, J. Lamy, E. F. J. van Bruggen, *Biochemistry* **20**, 1849 (1981); J. Markl, B. Kempter, B. Linzen, M. M. C. Bijlholt, E. F. J. van Bruggen, *Hoppe-Seyler's Z. Physiol. Chem.* **362**, 1631 (1981).
6. H.-J. Schneider, R. Drexel, G. Feldmaier, B. Linzen, F. Lottspeich, A. Henschen, *Hoppe-Seyler's Z. Physiol. Chem.* **364**, 1357 (1983); W. Schartau, F. Eyerle, P. Reisinger, H. Geisert, H. Storz, B. Linzen, *ibid.*, p. 1383. The unpublished data on *Astacus* hemocyanin was obtained by H.-J. Schneider *et al.* and that on the *Eurypelma* a chain by Schartau *et al.*; the manuscripts are in preparation.
7. T. Nemoto and T. Takagi, *Report at the 56th Annual Meeting of the Japanese Biochemical Society*, 29 September to 2 October 1983.
8. W. P. J. Gaykema *et al.*, *Nature (London)* **309**, 23 (1984).
9. J. Markl, A. Hofer, G. Bauer, A. Markl, B. Kempter, M. Brenzinger, B. Linzen, *J. Comp. Physiol.* **133**, 167 (1979); J.-P. van Eerd and A. Folkerts, in *Invertebrate Oxygen-Binding Proteins: Structure, Active Site and Function*, J. Lamy and J. Lamy, Eds. (Dekker, New York, 1981), pp. 139–149.
10. J. M. Vereijken, E. H. Schwander, N. M. Soeter, J. J. Beintema, *Eur. J. Biochem.* **123**, 283 (1982).
11. D. J. Barlow and J. M. Thornton, *J. Mol. Biol.* **168**, 867 (1983).
12. R. Loewe, *J. Comp. Physiol.* **128**, 161 (1978).
13. V. Z. Pletnev, A. P. Kuzin, S. D. Trakhanov, P. V. Kostetsky, *Biopolymers* **21**, 287 (1982).
14. B. Kempter, J. Markl, M. Brenowitz, C. Bonaventura, J. Bonaventura, *Hoppe-Seyler's Z. Physiol. Chem.* **366**, 77 (1985).
15. J. M. Vereijken, J. de Vlieg, J. J. Beintema, *Biochim. Biophys. Acta* **788**, 298 (1984); unpublished data on *Panulirus* hemocyanin by N.M. Soeter, H. J. Bak, B. Neuteboom, T. A. Jekel, J. J. Beintema, in preparation.
16. E. Yokota and A. F. Riggs, *J. Biol. Chem.* **259**, 4739 (1984). Additional data on *Limulus* hemocyanin were obtained by M. D. Moore, P. Q. Behrens, H. Nakashima, and A. Riggs; the manuscripts are in preparation.
17. The single letter abbreviations for the amino acid residues are: alanine, A; arginine, R; asparagine, N; aspartic acid, D; cysteine, C; glutamic acid, E; glutamine, Q; glycine, G; histidine, H; isoleucine, I; leucine, L; lysine, K; methionine, M; phenylalanine, F; proline, P; serine, S; threonine, T; tryptophan, W; tyrosine, Y; valine, V.
18. T. P. Hopp and K. R. Woods, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3824 (1981).
19. M. O. Dayhoff, Ed., *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Washington, D.C., 1972), vol. 5, p. D6.
20. ———, *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Washington, D.C., 1978), vol. 5, suppl. 3, p. 11.
21. W. D. I. Rolfe, in *The Terrestrial Environment and the Origin of Land Vertebrates*, A. L. Panchen, Ed. (Systematics Association, special volume No. 15; Academic Press, New York, 1980), pp. 117–157.
22. J. Bergstrom, in *Arthropod Phylogeny*, A. P. Gupta, Ed. (Van Nostrand Reinhold, New York, 1979), pp. 3–56.


23. The Munich laboratory was supported by the Deutsche Forschungsgemeinschaft (Li 107, Scha 317, and Schn 226). We thank R. Paul for computer work and G. Feldmaier-Fuchs and H. Storz for technical assistance. The University of Texas group was supported by NIH grants GM 21314 and GM 28410 and by a grant from the Robert A. Welch Foundation, F-213. We thank F. Waddill, M. Ervin, and K. Haschke for assistance. The Groningen laboratory was supported by the Dutch Foundation for Chemical Research (SON) with financial aid from the Dutch Organization for the Advancement of Pure Research (ZWO). We thank P. A. Jekel, K. H. Kalk, and P. Wietzes for technical assistance and E. F. J. van Bruggen and J. Drenth for continuous interest. The Tohoku group thanks K. Konishi for encouragement. This paper is the result of a workshop held in Groningen in the period 22 to 25 February 1984. Most authors of this paper attended this workshop, which was supported by grants from the Groningen University Fund and the Fund for Stimulating Research of the Faculty of Mathematics and Natural Sciences of the University of Groningen.