Minireview

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 19, Issue of May 11, pp. 15563–15566, 2001

Hemocyanins and Invertebrate Evolution*

Published, JBC Papers in Press, March 15, 2001,
DOI 10.1074/jbc.R100010200

Kensal E. van Holde‡§, Karen I. Miller‡§¶, and Heinz Decker‖**

From the ‡Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331-7305 and ‖Institute for Molecular Biophysics, University of Mainz, D55128 Mainz, Germany

The circulatory transport of oxygen is essential for efficient aerobic metabolism in most animals. A variety of proteins has evolved to facilitate this process. Most familiar are the hemoglobins, with representatives in almost every phylum (1). However, hemoglobins are not the unique oxygen carriers; there exist two less well known protein classes: (a) the hemerythrins, non-heme iron proteins found in a few invertebrates (2) and (*b*) hemocyanins, binuclear type 3 copper proteins utilized by many arthropods and molluscs (3). Because they are found freely dissolved in the blood, hemocyanins are easily purified and were the first proteins to be physically characterized as defined, multisubunit structures (4). However, their great size (some are over 10^7 Da) and subunit complexity hindered further study for many years. Only in the last decade have we gained a detailed understanding of how these giant molecules are constructed and how their structures might have evolved from simpler molecules. This, in turn, has provided intriguing clues concerning the evolution of invertebrates and their respiratory functions. Connecting the molecular biology of hemocyanins to invertebrate evolution is the object of this review.

Hemocyanin Structure: One Protein or Two?

The copper-containing oxygen transport proteins of both molluscs and arthropods were originally given the same name: hemocyanin. This was justified by similarities in the mode of oxygen binding. In both phyla, the oxygen-binding site involves a pair of copper atoms, which are in the Cu(I) state in the deoxy form but become Cu(II) upon oxygenation, binding the oxygen as O_2^{2-}. This change accounts for the blue color developed upon oxygenation. Indeed, recent structural analysis has shown that the O_2^{2-}-binding sites of molluscan and arthropod hemocyanins are very similar, both in the coordination of copper via histidine ligands and the way in which oxygen is bound (5, 6) (Fig. 1). In contrast to these similarities molluscan and arthropod hemocyanins are profoundly different in molecular structure at all levels. This can be seen immediately in the differences in quaternary structure schematically depicted in Fig. 2. Because of these differences, it has now become customary to consider the molluscan and arthropod hemocyanins as different proteins (3, 7, 8). In the following section we briefly summarize what is now known concerning the structures of these proteins and show that the nature of their relationship may be more subtle than previously thought.

Arthropod hemocyanins are built as multiples of hexamers, each hexamer made of monomers of about 75 kDa. An example of a four-hexamer structure is shown in Fig. 2*A*. Sequence analysis shows that a given arthropod hemocyanin may contain several variants of the common monomer sequence (see for example Refs. 9 and 10), each variant occupying a specific position in the whole molecule. The combination of chain variants determines the level to which hexamer association can occur (11). The sequences are sufficiently similar that all arthropod hemocyanin subunits probably have a tertiary structure similar to the example shown in Fig. 2*B*. Each subunit is organized into three domains; the second, highly helical domain carries the active site copper pair. Each copper is ligated by three histidine residues, as shown in Fig. 1, and lies within a 4α-helix bundle reminiscent of hemerythrin or perhaps even the globin fold (12).

Molluscan hemocyanins are built on an entirely different plan, as Fig. 2*C* shows. The polypeptide chains are very large, about 350–450 kDa each, and each consists of 7 or 8 globular "functional units" connected by linker peptide strands (13, 14). In the blood of cephalopod molluscs, like squids or octopi, the circulating hemocyanin exists as decamers of these large subunits, forming hollow cylindrical arrays with 5- or 10-fold axial symmetry. In some other molluscs (chiefly gastropods) dimers or even higher oligomers of these decamers can be found (see Fig. 2*C*). Such molecules are truly immense; the structure shown in Fig. 2*C* has a molecular mass of about 9×10^6 Da and contains 160 oxygen-binding sites! Sequence analysis (13, 14) has revealed that the functional units within a molluscan hemocyanin monomer are quite similar, with 40–50% identity. The tertiary structure of one such functional unit has been determined (6) and is depicted in Fig. 2*D*. Note that the arthropod subunit and the molluscan functional unit have a quite different tertiary structure. The molluscan unit is smaller (about 50 kDa) than the arthropod subunit and consists of only two domains, an N-terminal highly helical domain carrying the O_2 site and a C-terminal domain that is largely β-sheet. Clearly, at most levels, molluscan and arthropod hemocyanins appear to be quite different proteins.

Nevertheless, a closer comparison of molluscan functional units with arthropod hemocyanin subunits raises intriguing questions. Although there appears to be very little overall similarity in amino acid sequences, we do detect in the copper-binding regions what appear to be meaningful local similarities. In each case there are two well separated copper-binding regions: that nearer the N terminus is called the "A" site and that nearer the C terminus the "B" site. Fig. 3 depicts sequences in the A and B regions of representative arthropod and molluscan hemocyanins and several related binuclear copper proteins. What sequence similarity exists between these proteins is concentrated in these regions. Moreover, x-ray diffraction studies have shown that octopus hemocyanin (6), two arthropod hemocyanins (5, 12), and catechol oxidase (15) exhibit remarkable similarity in the spatial arrangement of the six histidine ligands that hold the copper atoms. Fig. 4 shows that a very particular arrangement of reactive groups forming the active site is necessary to bind dioxygen as is done by type 3 copper proteins. This conservation lies in the three-dimensional structure rather than in the linear sequences. In fact, the sequences determining the similar structures are quite different in the two phyla. This is particularly clear in the A site, where the molluscan and arthropod hemocyanins have a somewhat different sequence arrangement of the three copper-binding histidines. It is often said that the A and B sites in arthropod hemocyanins are very similar, but aside from the location of the binding histidines, this is not so at least insofar as sequence is concerned (Fig. 3). The overall arrangement of the copper-binding histidines in the A site differs in molluscs mainly in the change in sequence position of the second of these three residues (Fig. 3). The only wholly conserved positions in this whole group of proteins are histidines 1 and 3, and the phenylalanine lying four residues upstream from histidine 3. Considering all of these observations, we must conclude that it is an oversimplification to state that molluscan and arthropod hemocyanins are unrelated. However, the relationship must be distant.

* This minireview will be reprinted in the 2001 Minireview Compendium, which will be available in December, 2001.

§ Supported by a research grant from the National Science Foundation.

¶ To whom correspondence should be addressed. Tel.: 541-737-2264; Fax: 541-737-0481; E-mail: millerk@ucs.orst.edu.

** Supported by the Deutsche Forschungsgemeinschaft.

From What Did Hemocyanins Evolve?

A clue to the origins of the hemocyanins can be seen in Fig. 3 and from examination of the much larger body of data of which this represents a small sample. Both classes of hemocyanins appear to be related to proteins exhibiting phenoloxidase activity (16, 17). The arthropod hemocyanins exhibit some sequence similarity to arthropod phenol oxidases (see Refs. 18 and 19, for example), whereas the molluscan hemocyanins resemble more closely that group of enzymes known as tyrosinases (20) and catechol oxidases (15). These differential sequence similarities are especially clear in the region of the active site, as shown in Fig. 3. Furthermore, structural comparison of sweet potato catechol oxidase and the hemocyanin-binding domain from octopus reveals a specific similarity with respect to an unusual Cys-His thioether bridge, which holds one of the copper-binding histidines in the proper orientation. The fact that both kinds of hemocyanins have weak phenol oxidase activity further supports the idea of close affinity (16, 17, 21, 22). Both tyrosinases and phenol oxidases are widely distributed and ancient. If hemocyanins did evolve from phenoloxidases and tyrosinases the substrate binding capacity of the enzyme must have been inhibited without hindering O_2 binding. This seems to have been accomplished by the addition of residues blocking the active site to large substrates (17, 23, 24). Indeed, just such a difference has been noted by comparing the crystal structure of sweet potato catechol oxidase with that of molluscan hemocyanin (15).

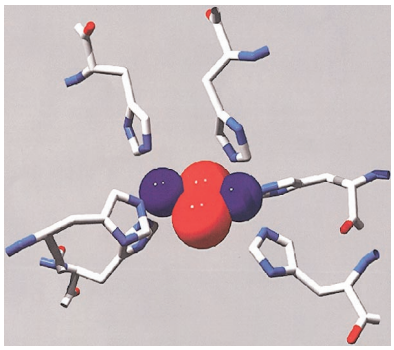

FIG. 1. **The oxygen-binding site common to type 3 copper proteins.** The specific structure shown here is taken from an x-ray diffraction study of *Limulus polyphemus* hemocyanin (5). Virtually identical structures have been observed in several other type 3 copper proteins. Coppers are *blue* and oxygens are *red*.

The Emergence of Hemocyanin Function

If the precursor to hemocyanins was a cytoplasmic enzyme, some mechanism had to be provided to permit transport into the circulatory system. Leader sequences are found on molluscan hemocyanins (25). Arthropod hemocyanins are exported into the hemolymph either by leader sequences or by cell lysis (8).

However, providing for export is not sufficient; a low molecular weight monomeric protein dissolved in the hemolymph is not efficient for oxygen transport. If its concentration in the hemolymph is high enough to transport a significant amount of oxygen, it will yield an unbearably high osmotic pressure. This problem can be avoided if the subunits aggregate and/or polymerize to yield giant molecules. For example, octopus hemocyanin exists in the hemolymph at a concentration of about 100 mg/ml, yet because of its great molecular mass gives the same osmotic pressure contribution as about 2 mg/ml vertebrate hemoglobin would if it were free in the blood. The evolutionary solution found for this problem differed in the arthropods and molluscs. The former developed subunits that non-covalently associated; the latter appear to have used gene duplication (25) to generate long, multiunit chains, which then further polymerize. These strategies, which generated large macromolecular structures with many binding sites, not only decreased the osmotic pressure but also provided a new versatility in facing the problem of loading oxygen with high affinity at the respiratory interface and unloading the oxygen where it is needed. This is accomplished by cooperative binding, a behavior confined to molecules with multiple binding sites. Some hemocyanins possess the highest cooperativity found in nature, with Hill coefficients of more than 9 (26).

Thus, current evidence indicates how molluscan and arthropod hemocyanins could have arisen independently from distinct but similar enzymes with phenol oxidase activities. Faint sequence

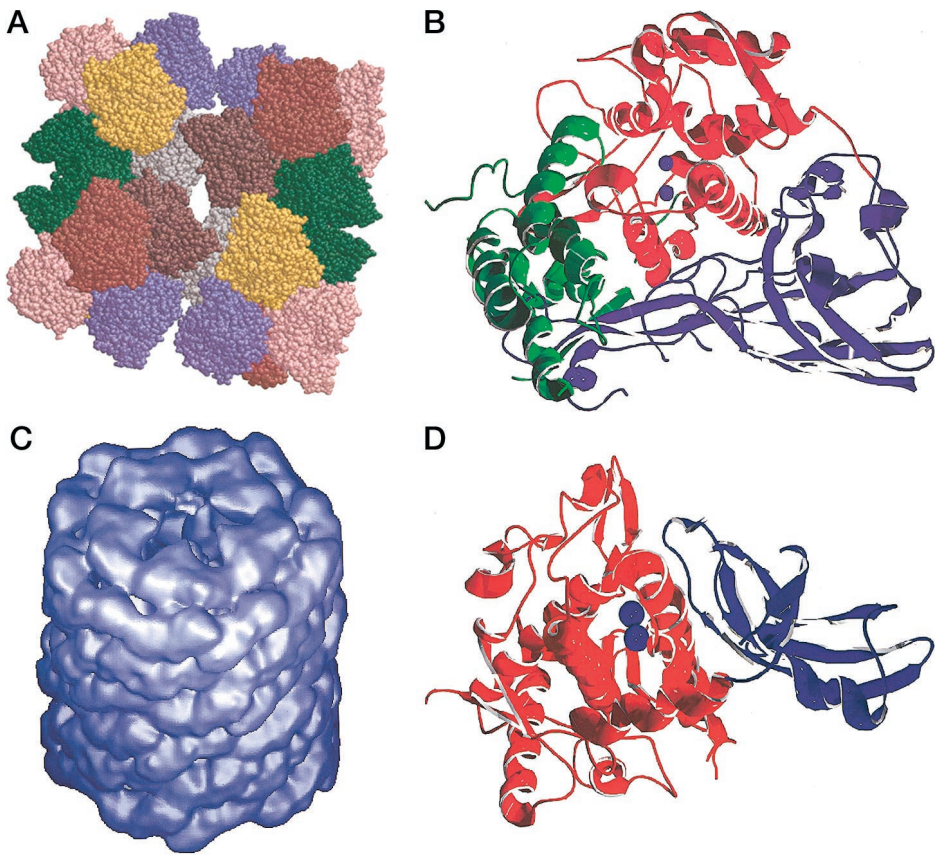

FIG. 2. **Representative hemocyanin structures.** *A,* quaternary structure of an arthropod hemocyanin, the 4-hexamer (24 subunits) hemocyanin from the spider *Eurypelma californicum* (10). The seven different subunit types are colored differently. *B,* tertiary structure of an arthropod hemocyanin subunit from *L. polyphemus* (35). The three domains I (*green*), II (*red*), and III (*purple*) are colored differently. *C,* quaternary structure of a representative molluscan hemocyanin, that of the abalone, *Haliotis tuberculata*. This consists of 20 polypeptide chains, each containing 8 functional units and therefore 8 binding sites. The functional units are divided (6/2) between the cylindrical wall and an internal collar (36) (courtesy of Dr. U. Meissner). *D,* tertiary structure of a molluscan hemocyanin functional unit, the C-terminal unit from the *Octopus dofleini* hemocyanin (6). The two domains are colored differently. Quaternary structures are from high resolution electron microscopy, and subunit structures are from x-ray diffraction studies.

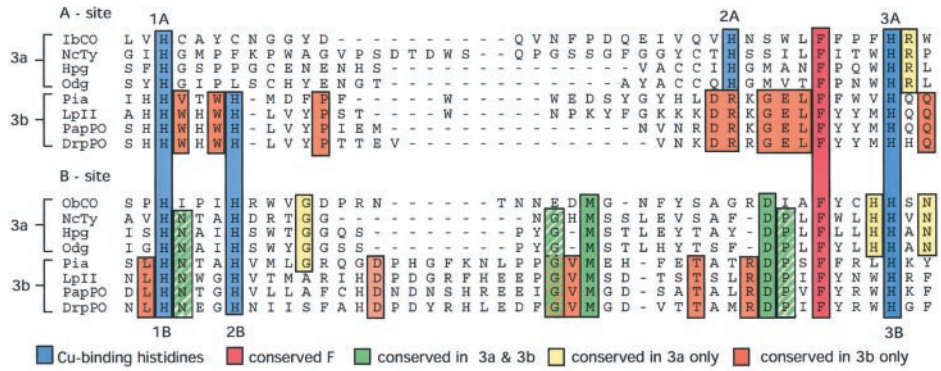

FIG. 3. **Comparison of sequences in the A and B sites of a number of type 3 copper proteins.** The copper-ligating histidines are denoted by *1*, *2*, and *3*. We define two classes of type 3 proteins; class 3a includes the molluscan hemocyanins, tyrosinases, and catechol oxidase, and class 3b comprises arthropod phenol oxidases and hemocyanins. Data are taken from wider listings (see for example, Refs. 9 and 19). Species code: *IbCO,* sweet potato catechol oxidase; *NcTy, Neurospora crassa* tyrosinase; *Hpg, Helix pomatia* hemocyanin, functional unit g; *Odg, Octopus dofleini* hemocyanin, functional unit g; *Pia, Panulirus interruptus* hemocyanin, subunit a; *LpII, Limulus polyphemus* hemocyanin, subunit II; *PapPO, Pacifasticus leniusculus* phenol oxidase; *DrpPO, Drosophila melanogaster* phenol oxidase.

similarities between these enzymes (mainly near the O_2-binding site) suggest in turn an even more remote common ancestor.

A Common Ancestor for all Type 3 Copper Proteins?

It has often been suggested that the difference in the A site between molluscan and arthropod hemocyanins indicates fundamentally different ancestries for these two classes of proteins (3, 7, 8). In particular, it was proposed that the arthropod copper-binding region arose from a simple duplication of a primordial B site, whereas the molluscan copper-binding domain resulted from a fusion between two genes, one carrying an A site type of structure and the other a B site (3, 7, 16). However, closer examination of Fig. 3 casts doubt as to whether so complex an explanation is necessary. Alignment as shown here indicates that the A and B sites are very similar in the placement of histidines 1B and 1A, and 3B and 3A. Also, the conserved phenylalanine residue is in exactly the same registry in both sites. Of the truly critical residues, only histidine 2 has shifted significantly between what we call type 3a and type 3b copper proteins. Perhaps the simplest explanation for the present sequence differences is that all type 3 copper proteins have evolved from a binuclear predecessor, which itself arose from duplication of a B-like copper site. The branches of these proteins, segregating into proto-molluscan and proto-arthropodan lines then underwent independent evolution to yield the quite different sequences found today (see Fig. 5). The greater dissimilarity between the molluscan hemocyanins sequence and tyrosinases as compared with that between arthropod hemocyanins and insect phenol oxidases suggests that the molluscan hemocyanin-tyrosinase split is more ancient (see below). Conservation of active site geometry, even through evolutionary changes that greatly modify the surrounding protein framework, is by no means unknown. A well recognized example is found in subtilisin and the mammalian serine proteases (27).

When Did Hemocyanins Originate?

Because arthropod and molluscan hemocyanins seem to have evolved independently from somewhat different (but related) proteins, we should consider their origins as separate events, occurring sometime after the split between the Lophotrochozoa (which includes the molluscs) and the Ecdysozoa (of which the arthropoda are now recognized as members (see Refs. 28 and 29)). In either case, the evolution of a fully functional hemocyanin must have been quite rapid, as judged by the starlike branching patterns for the diversification of either arthropod subunits (even within one species (10)) or molluscan functional units (13, 14). That is, once circulating hemocyanins appeared they rapidly evolved into something like the present structures. Attempts to define the points in evolutionary time at which the molluscan and arthropod hemocyanins first appear as functional entities encounter (and may contribute to) an ongoing debate concerning the time of divergence of the major phyla. Estimates for the onset of this divergence range over

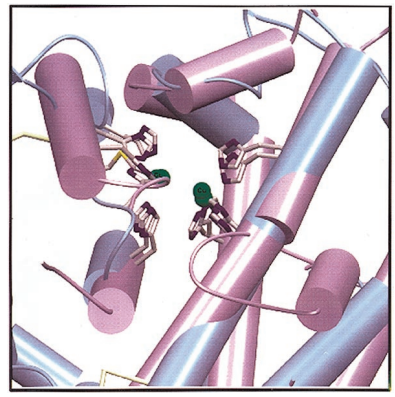

FIG. 4. **A superposition of the binding site regions in the arthropod (*Limulus polyphemus* (*purple*)) subunit and the molluscan (*Octopus dofleini* (*blue*) functional unit.** The A site is at the *left*. Note the close juxtaposition of structure on the B site and of both coppers and all ligating histidines (taken from Ref. 6 with permission).

half a billion years, from the beginning of the Cambrian, 550 MYA[1] (29) to 800–1000 MYA (30, 31). The early Cambrian demonstrates the first profusion of shelly fossils, but fossils of softer bodied metazoans are well known back into the Vendian (550–650 MYA). It seems likely that some of these creatures, although difficult to classify, represented primitive arthropods and molluscs (29, 32). In this context, what do the hemocyanin sequence data tell us? By using as calibration dates events such as the divergence of cephalopods from gastropods or chelicerate arthropods from other arthropods, it is possible to use comparisons of hemocyanin subunit (or functional unit) sequences to estimate the time at which the multiple units arose. For arthropods, the data point to a rapid divergence of the chelicerate subunits (and presumably the first fully functional arthropod hemocyanin) at about 550–600 MYA (10, 33). This is a date with which the more conservative models for metazoan evolution would be comfortable. The molluscan hemocyanin data, on the other hand, yield an apparently surprising result, that the diversification of molluscan functional units occurred between about 700 and 800 MYA (14). If correct, this is indeed strong evidence for the "phyla-early" models.

On the basis of these results, we can propose a very tentative scenario for the evolution of the hemocyanins. Very early in the evolution of life, the environment was anaerobic (34). Under these circumstances, the oxygen production produced by photosynthesis must have been toxic to many creatures. A variety of enzymes were

[1] The abbreviation used is: MYA, million years ago.

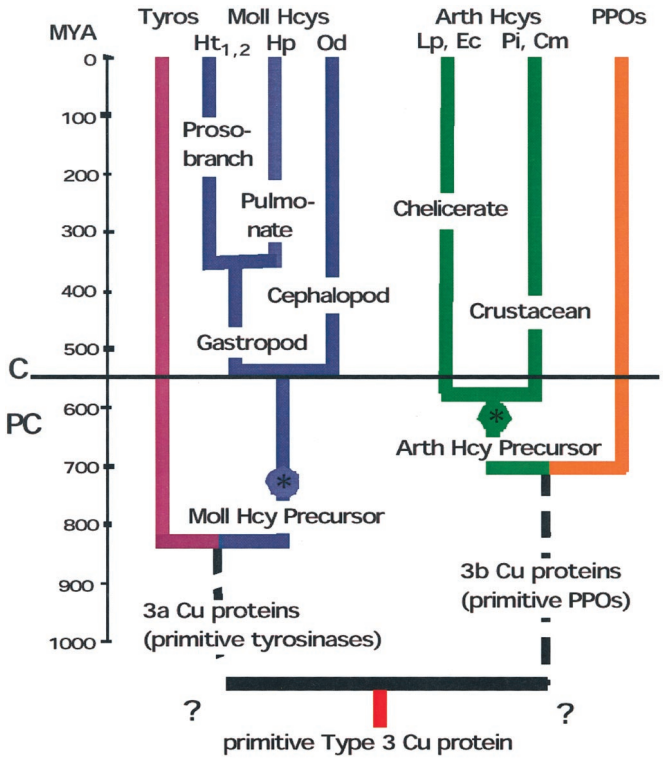

FIG. 5. **A speculative map of the evolution of type 3 copper proteins.** The *circled asterisks* indicate the presumed origins of the two types of hemocyanins. The horizontal line (*C/PC*) denotes the Cambrian/Precambrian boundary. *Tyros,* tyrosinases; *PPO,* prophenol oxidases. Symbols designating hemocyanins: $Ht_{1,2}$, *Haliotis tuberculata* types 1 and 2; *Hp, Helix pomatia; Od, Octopus dofleini; Lp, Limulus polyphemus; Ec, Eurypelma californicum; Pi, Panulirus interruptus; Cm, Cancer magister*.

evolved to neutralize oxygen by carrying out oxidation reactions. For this two metal ions were used, iron and copper. In the case of copper, a type 3 copper center evolved, in which two coppers reversibly bond oxygen as a peroxide (5, 6). The resulting tyrosinases and other phenol oxidases must represent an extremely ancient class of binuclear copper proteins, predating the emergence of higher metazoan phyla. It is possible that these all originated from a single protein whose function was to protect primitive organisms from the new toxin, oxygen. By the time the major metazoan phyla began to emerge (700–800 MYA?) oxygen levels were close to those at the present (34), and aerobic metabolism had been well established. An increase in animal size, as well as the development of impermeable integuments, made simple diffusion inadequate for oxygen supply. Therefore, a circulating oxygen transport protein became essential to utilize the advantages of aerobic metabolism. We postulate that such transport proteins developed in several independent ways, hemoglobins from myoglobins, hemerythrins from myohemerythrins, and the two kinds of hemocyanins from two different classes of phenol oxidases. The development of such diverse transport systems, from different roots, strongly suggests that it occurred *after* the divergence of the major phyla. This in turn argues that the phylogenetic divergence occurred before 750 MYA. That hemocyanin evolution occurred earlier in the molluscan line is indicated by both the earlier dating of molluscan functional unit divergence and by the apparent greater evolutionary distance between molluscan hemocyanins and their tyrosinase cousins (Fig. 3). A reasonable scenario is shown in Fig. 5. At the present, such speculations must of course be regarded with caution. Much more sequence data on these and other invertebrate proteins are urgently needed.

REFERENCES

1. Hardison, R. C. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 5675–5619
2. Kurtz, D. M. (1992) *Adv. Comp. Environ. Physiol.* **13,** 151–171
3. van Holde, K. E., and Miller, K. I. (1995) *Adv. Protein Chem.* **47,** 1–81
4. Svedberg, T., and Chirnoaga, E. (1928) *J. Am. Chem. Soc.* **50,** 1399–1411
5. Magnus, K. A., Hazes, B., Ton-That, H., Bonaventura, C., Bonaventura, J., and Hol, W. G. (1994) *Proteins* **19,** 302–309
6. Cuff, M. E., Miller, K. I., van Holde, K. E., and Hendrickson, W. A. (1998) *J. Mol. Biol.* **278,** 855–870
7. van Holde, K. E., Miller, K. I., and Lang, W. H. (1992) *Adv. Comp. Environ. Physiol.* **13,** 251–300
8. Markl, J., and Decker, H. (1992) *Adv. Comp. Environ. Physiol.* **13,** 325–376
9. Durstewitz, G., and Terwilliger, N. (1997) *Mol. Biol. Evol.* **14,** 266–276
10. Voit, R., Feldmaier-Fuchs, G., Schweikardt, T., Decker, H., and Burmester, T. (2000) *J. Biol. Chem.* **275,** 39339–39344
11. Markl, J. (1986) *Biol. Bull.* **171,** 90–115
12. Volbeda, A., and Hol, W. G. J. (1989) *J. Mol. Biol.* **206,** 531–546
13. Miller, K. I., Cuff, M. E., Lang, W. E., Varga-Weisz, P., Field, K. G., and van Holde, K. E. (1998) *J. Mol. Biol.* **278,** 827–842
14. Lieb, B., Altenheim, B., and Markl, J. (2000) *J. Biol. Chem.* **275,** 5675–5681
15. Klabunde, T., Eicken, C., Sacchettini, J. C., and Krebs, B. (1998) *Nat. Struct. Biol.* **5,** 1084–1090
16. Decker, H., and Terwilliger, N. (2000) *J. Exp. Biol.* **203,** 1777–1782
17. Decker, H., and Tuczek, F. (2000) *Trends Biochem. Sci.* **25,** 392–397
18. Aspan, A., Huang, F. S., Cerenius, L., and Söderhäll, K. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 939–943
19. Burmester, T., and Scheller, K. (1996) *J. Mol. Evol.* **42,** 713–728
20. van Gelder, C. W. G., Flurkey, W. H., and Wichers, H. J. (1997) *Phytochemistry* **45,** 1309–1323
21. Salvato, B., Jori, G., Piazzese, A., Ghiretti, F., Beltramini, M., and Lerch, K. (1983) *Life Sci. Rep. Suppl.* **1,** 313–317
22. Salvato, B., Santamaria, M., Beltramini, M., Alzuet, G., and Casella, L. (1998) *Biochemistry* **37,** 14065–14077
23. Decker, H., Dillinger, R., and Tuczek, F. (2000) *Angew. Chem. Int. Ed. Engl.* **39,** 1591–1595
24. Decker, H., and Rimke, T. (1998) *J. Biol. Chem.* **273,** 25889–25892
25. Lieb, B., Altenheim, B., Markl, J., Vincent, A., van Olden, E., van Holde, K., and Miller, K. (2001) *Proc. Natl. Acad. Sci. U. S. A.*, in press
26. Loewe, R. (1978) *J. Comp. Physiol. B* **128,** 161–168
27. Creighton, T. (1993) *Proteins*, 2nd Ed., W. H. Freeman, New York
28. Adoutte, A., Balavoine, G., Lartillot, N., Lespinet, O., Prud'homme, B., and de Rosa, R. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 4453–4458
29. Conway Morris, S. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 4426–4429
30. Gu, X. (1998) *J. Mol. Evol.* **47,** 369–371
31. Wray, G. A., Levinton, J. S., and Shapiro, L. H. (1996) *Science* **274,** 568–573
32. Fedonkin, M. A., and Waggoner, B. M. (1997) *Nature* **388,** 868–871
33. Burmester, T. (2001) *Mol. Biol. Evol.* **18,** 184–195
34. Canfield, D. E., and Teske, A. (1996) *Nature* **382,** 127–132
35. Hazes, B., Magnus, K. A., Bonaventura, C., Bonaventura, J., Dauter, Z., Kalk, K. H., and Hol, W. G. (1993) *Protein Sci.* **2,** 597–619
36. Meissner, U., Dube, P., Harris, J. R., Stark, H., and Markl, J. (2000) *J. Mol. Biol.* **298,** 21–34