Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 244–248, January 1991
Biochemistry

Cloning and sequencing of *Octopus dofleini* hemocyanin cDNA: Derived sequences of functional units Ode and Odf

(proteins/evolution/copper binding/oxygen binding)

WALTER H. LANG* AND KENSAL E. VAN HOLDE†

Department of Biochemistry and Biophysics, Oregon State University, Corvallis, OR 97331-6503

Contributed by Kensal E. van Holde, October 4, 1990

ABSTRACT **A number of additional cDNA clones coding for portions of the very large polypeptide chain of *Octopus dofleini* hemocyanin were isolated and sequenced. These data reveal two very similar coding sequences, which we have denoted "A-type" and "G-type." We have obtained complete A-type sequences coding for functional units Ode and Odf; consequently, a total of three such unit sequences are now known from a single subunit of one molluscan hemocyanin. This presents the opportunity to make sequence comparisons within one hemocyanin subunit. Domains within one subunit show on the average 42% identity in amino acid residues; corresponding functional units from hemocyanins of different species show degrees of identity of 53–75%. Therefore, molluscan hemocyanins already existed before the individual molluscan classes diverged in the early Cambrian. Sequence comparisons of molluscan hemocyanins with arthropodan hemocyanins and tyrosinases allow us to identify the ligands of the "Copper B" site with high probability. Possible ligands for the "Copper A" site are proposed, based on sequence comparisons between molluscan hemocyanins and tyrosinases. Besides two histidine side chains, a methionine side chain might be involved in binding of Copper A, a result not in conflict with spectroscopic studies.**

Hemocyanins serve as oxygen-transport molecules in a variety of molluscs and arthropods (1, 2). In both molluscan and arthropodan hemocyanins, the oxygen binding site is formed by a binuclear copper center. X-ray diffraction studies show that each copper atom is complexed by three histidine ligands in arthropod hemocyanin (3). Spectroscopic evidence suggests, but does not prove, a similar binding in molluscan hemocyanins (4). The two copper atoms are antiferromagnetically coupled, and it has been suggested, but not yet proved, that an additional endogenous bridging ligand is involved.

Despite their superficial similarities in appearance and function, molluscan and arthropodan hemocyanins are fundamentally different with respect to their molecular architecture (1, 2). They both occur as very large subunit aggregates, but arthropodan hemocyanins are found as multiples of a basic hexameric assembly of heterogeneous subunits (molecular mass ≈ 75 kDa), with each subunit containing one oxygen-binding site. Molluscan hemocyanins are composed of 10 or 20 very large subunits (molecular mass ≈ 350–450 kDa), which are arranged to form cylinders. Each subunit is composed of seven or eight oxygen-binding functional units (molecular mass per unit = 45–55 kDa). The limited sequence information available to date on molluscan hemocyanins has revealed almost no sequence similarity between these and arthropodan hemocyanins except in a small region that corresponds to the "Copper B" site in *Panulirus interruptus* hemocyanin (5–7).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Our laboratory is investigating the structure and function of the hemocyanin from the giant Pacific octopus (*Octopus dofleini*). The native 51S molecule is composed of 10 apparently identical 350-kDa subunits (8). Electron microscopy of dissociated hemocyanin reveals the presence of seven globular functional units per subunit. Upon mild proteolytic digestion, the 360-kDa subunit can be cleaved into single functional units and multiunit fragments (9). Originally these functional units were arbitrarily numbered in their order of identification. Their order within the whole subunit was then determined and found to be 7 - 4 - 3 - 6 - (5 - 2) - 1, with some uncertainty regarding the order of domains 2 and 5 (9). This ambiguity has now been resolved with the tentative order above being confirmed (ref. 10 and this work). To be consistent with other hemocyanin nomenclature, we propose to relabel the functional units by letters a–g, with the prefix Od to signify *O. dofleini*:

7 - 4 - 3 - 6 - 5 - 2 - 1

Oda-Odb-Odc-Odd-Ode-Odf-Odg

It should be noted that most molluscan hemocyanins are composed of eight functional units, in which case the C-terminal unit is labeled with h.

The N-terminal sequences of units Oda, Ode, Odf, and Odg have been determined by peptide sequencing (9), and the complete sequence of Odg was reported earlier from our laboratory (6). A model for the structure of the "Copper A" site in molluscan hemocyanins was also presented in the same paper. Here we present the sequences of Ode and Odf of this hemocyanin.‡ Sequences of the *Octopus* units will be compared with one another and with functional units of other molluscan hemocyanins. The extended data set allows us to present a revised model for the structure of the Copper A site.

MATERIALS AND METHODS

Oligonucleotides. The oligonucleotide WL3 (5'- CAA CTG AGG GGA ATT CCA TG -3') was synthesized complementary to base pairs 43–62 of the pHC1 cDNA insert (6). The oligonucleotide WL4 (5'- ATT CCA TGG TCG CTT TC -3') was synthesized complementary to a sequence portion close to the 5' end of the coding strand of cDNA clone pHC2 (underlined in Fig. 2). Both oligonucleotides were synthesized on an Applied Biosystems DNA synthesizer. They were gel-purified before use and served as probes for screening libraries and as primers.

cDNA Cloning and Sequencing. The preparation of the specifically primed cDNA library and the screening process has been described (6). Hybridization temperatures used were 55°C for probe WL3 and 47°C for probe WL4. Restric-

*Present address: Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, WA 98104.
†To whom reprint requests should be addressed.
‡The sequence reported in this paper has been deposited in the GenBank data base (accession no. M57288).

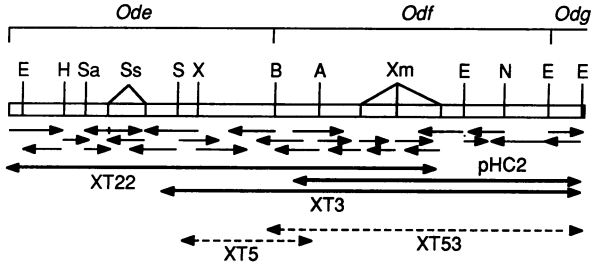

FIG. 1. Restriction map of cDNA clones. Solid heavy arrows indicate A-type sequence; broken arrows indicate G-type sequence. The light arrows show the direction of sequencing of subcloned A-type DNA fragments used to deduce the sequence in Fig. 2. A, *Acc* I; B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; N, *Nco* I; S, *Sac* I; Sa, *Sau* 3AI; Ss, *Ssp* I; X, *Xba* I; Xm, *Xmn* III.

tion fragments of the isolated cDNA clones were subcloned into pUC 19, and plasmid DNA was purified in small scale and sequenced as described (6). Details of all of these procedures have been given (7).

RESULTS

After transformation with our specifically primed cDNA library, we obtained about 3000 clones. In a first round, ≈10% of these were screened with probe WL3. Ninety-eight clones were obtained and checked for insert size. The clone with the longest insert (1350 bp) was termed pHC2. A restriction map of it was assembled, and it was sequenced as indicated in Fig. 1. Clone pHC2 had a single open reading frame, and comparison of the translated protein sequence with that of Odg showed that pHC2 coded for almost the entire functional unit Odf and the N-terminal 42 amino acids of Odg.

From analysis of the first-strand products by alkaline agarose gel electrophoresis, it could be expected to isolate clones with an insert size of up to 2 kilobases (kb). Therefore, the entire library was screened with probe WL4. We would expect our probe to hybridize only to clones of at least the size of pHC2, assuming that second-strand synthesis was full-length. We isolated 65 more cDNA clones, which were rescreened, and their inserts were analyzed for size. Clones XT3 and XT22 had the longest inserts, 1.8 and 1.9 kb, respectively. A restriction map from both clones was made (Fig. 1), and the regions not overlapping with pHC2 were subcloned and sequenced. The sequencing strategy is shown in Fig. 1. The combined nucleotide sequence of pHC2 and XT22 together with the corresponding translated protein sequence of Ode and Odf is shown in Fig. 2. The N-terminal sequence obtained for Odf is almost identical to that obtained by protein sequencing of functional unit 2, which confirms the tentative order of units proposed earlier (9).

During sequence analysis of other clones from the same library, we have detected the presence of a slightly different second hemocyanin cDNA sequence. We denote the sequence described above and this second sequence as "A-

```
1     GAGAAGGCAATGAA TATCTGGTACGAAAG AATGTCGAAAGACTT TCTCTGTCTGAAATG AATTCTTTGATACAC GCTTTCAGAAGAATG CAGAAGGACAAGTCA
1       E  G  N  E   Y  L  V  R  K   N  V  E  R  L   S  L  S  E  M   N  S  L  I  H   A  F  R  R  M   Q  K  D  K  S

105   TCTGACGGTTTTGAG GCAATCGCTTCATTC CATGCTCTTCCTCCT CTCTGTCCAAGCCCA ACTGCCAAACATAGG CACGCTTGTTGTCTT CATGGTATGGCTACG
35    S  D  G  F  E   A  I  A  S  F   H  A  L  P  P   L  C  P  R  P   T  A  K  H  R   H  A  C  C  L   H  G  M  A  T

210   TTCCCTCACTGGCAC AGGCTCTACGTTGTT CAGTTTGAACAAGCT TTACATAGACATGGA GCTACGGTTGGCGTT CCTTACTGGGATTGG ACCCGTCCTATTTCA
70    F  P  H  W  H   R  L  Y  V  V   Q  F  E  Q  A   L  H  R  H  G   A  T  V  G  V   P  Y  W  D  W   T  R  P  I  S

315   AAGATCCCTGATTTC ATAGCGTCGGAAAAG TATTCTGATCCTTTC ACTAAAATAGAGGTT TATAACCCATTTAAT CATGGTCATATTTCT TTCATTAGTGAGGAC
105   K  I  P  D  F   I  A  S  E  K   Y  S  D  P  F   T  K  I  E  V   Y  N  P  F  N   H  G  H  I  S   L  I  S  E  D

420   ACTACGACTAAACGA GAGGTCAGCGAATAT TTGTTTGAACATCCT GTACTGGGAAAACAG ACGTGGCTCTTTGAT AACATCGCTTTAGCT TTGGAACAGACCGAT
140   T  T  T  K  R   E  V  S  E  Y   L  F  E  H  P   A  L  G  K  Q   T  W  L  F  D   N  I  A  L  A   L  E  Q  T  D

525   TATTGCGATTTCGAA ATACAATTAGAGATT GTTCACAATGCCATT CACTCGTGGATTGGC GGGAAAGAAGAGCAT TCCTTGAACCATTTA CATTATGCAGCCTAC
175   Y  C  D  F  E   I  Q  L  E  I   A  H  N  A  I   H  S  W  I  G   G  K  E  E  H   S  L  N  H  L   H  Y  A  A  Y

630   GACCCAATATTCTAT CTACATCATTCTAAT GTCGATCGTTTGTGG GTTATTTGGCAAGAA TTGCAGAAATTGAGA GGTCTCAATGCTTAT GAATCCCATTGTGCT
210   D  P  I  F  Y   L  H  H  S  N   V  D  R  L  W   V  I  W  Q  E   L  Q  K  L  R   G  L  N  A  Y   E  S  H  C  A

735   CTTGAACTTATGAAA GTTCCGTTGAAACCG TTCTCTTTCGGAGCT CCTTACAATTTGAAT GATCTAACTACCAAA TTGTCTAAACCTGAA GATATGTTTAGATAC
245   L  E  L  M  K   V  P  L  K  P   F  S  F  G  A   P  Y  N  L  N   D  L  T  T  K   L  S  K  P  E   D  M  F  R  Y

840   AAAGACAACTTCCAT TACGAATACGACATT CTAGATATTAACAGT ATGTCTATTAATCAG ATAGAGTCGTCGTAC ATCAGACACCAAAAA GACCACGATCGTGTT
280   K  D  N  F  H   Y  E  Y  D  I   L  D  I  N  S   M  S  I  N  Q   I  E  S  S  Y   I  R  H  Q  K   D  H  D  R  V

940   TTTGCCGGCTTTTTG TTGAGTGGATTTGGT TCATCAGCTTATGCA ACCTTTGAAATCTGT ATTGAAGGAGGAGAA TGTCATGAAGGAAGT CACTTTGCTGTGTTG
315   F  A  G  F  L   L  S  G  F  G   S  S  A  Y  A   T  F  E  I  C   I  E  G  G  E   C  H  E  G  S   H  F  A  V  L

1050  GGAGGCAGCACAGAA ATGCCATGGGCCTTC GACCGTCTCTATAAG ATAGAAATTACTGAT GTACTTTCTGATATG CACTTAGCGTTCGAT TCAGCTTTCACTATT
350   G  G  S  T  E   M  P  W  A  F   D  R  L  Y  K   I  E  I  T  D   V  L  S  D  M   H  L  A  F  D   S  A  F  T  I

1155  AAAACGAAAATAGTT GCTCAAAATGGAACT GAACTGCCAGCTAGC ATTCTACCAGAAGCA ACTGTAATAAGGATC CCACCTTCCAAGCAA GACGCCGATATTGAC
385   K  T  K  I  V   A  Q  N  G  T   E  L  P  A  S   I  L  P  E  A   T  V  I  R  I*  P  P  S  K  Q   D  A  D  I  D

1260  ATCCCACTAAATCAT ATCCGACGAAATGTA GAGTCTTTGGACGAA AGAGATATTCAGAAT CTTATGGCAGCTCTT ACTCGGGTTAAGAAA GATGAAAGCGACCAT
420   I  P  L  N  H   I  R  R  N  V   E  S  L  D  E   R  D  I  Q  N   L  M  A  A  L   T  R  V  K  K   D  E  S  D  H

1365  GGATTCCAGACTATT GCTAGTTATCATGGT TCAACGCTGTGTCCG AGTCCAGAGGAGCCC AAATATGCCTGTTGT CTACATGGAATGCCC GTCTTCCCACATTGG
455   G  F  Q  T  I   A  S  Y  H  G   S  T  L  C  P   S  P  E  E  P   K  Y  A  C  C   Y  H  G  M  P   V  F  P  H  W

1470  CACCGTGTTTACTTA TTACATTTTGAAGAT TCTATGCGCCGGCAT GGCTCCAGTGTTGCC ACTCCTTATTGGGAT TGGACACAACCTGGT ACGAAACTGCCTAGA
490   H  R  V  Y  L   L  H  F  E  D   S  M  R  R  H   G  S  S  V  A   T  P  Y  W  D   W  T  Q  P  G   T  K  L  P  R

1575  CTTTTAGCAGATTCT GACTACTATGATGCT TGGACTGATAATGTG ATCGAGAATCCATTC CTGAGGGGCTACATT ACATCTGAAGACACC TACACAGTCAGGGAC
525   L  L  A  D  S   D  Y  Y  D  A   W  T  D  N  V   I  E  N  P  F   L  R  G  Y  I   T  S  E  D  T   Y  T  V  R  D

1680  GTAAAGCCAGAGCTA TTTGAAATCGGTGGA GGAGAGGGATCTACT CTTTACCAACAAGTA CTACTGATGCTTGAA CAAGAAGACTACTGT GACTTCGAAGTTCAG
560   V  K  P  E  L   F  E  I  G  G   G  E  G  S  T   L  Y  Q  Q  V   L  L  M  L  E   Q  E  D  Y  C   D  F  E  V  Q

1785  TTTGAAGTAGTTCAT AACTCTATTCACTAC CTTGTTGGTGGTCAC CAGAAATATGCCATG TCCAGTTTGGTCTAT AGTTCCTTTGATCCT ATCTTCTACGTTCAT
595   F  E  V  V  H   N  S  I  H  Y   L  V  G  G  H   Q  K  Y  A  M   S  S  L  V  Y   S  S  F  D  P   I  F  Y  V  H

1890  CATTCAATGGTTGAT CGTCTCTGGGCTATT TGGCAAGCTCTCCAG GAACACAGACATTTG CCGTTTGATAAGGCT TACTGTGCCCTGGAA CAACTGTCATTCCCA
630   H  S  M  V  D   R  L  W  A  I   W  Q  A  L  Q   E  H  R  H  L   P  F  D  K  A   Y  C  A  L  E   Q  L  S  F  P

1995  ATGAAGCCTTTCGTT TGGGAGTCCAACCCA AATCTACATACACGA GCTGCATCAACACCA CAACATCTATTTGAC TACAATAAACTCGGT TACAAATATGATGAC
665   M  K  P  F  V   W  E  S  N  P   N  L  H  T  R   A  A  S  T  P   Q  H  L  F  D   Y  N  K  L  G   Y  K  Y  D  D

2100  CTCGAATTCCATGGA ATGAATATAGATCAA TTGGAAAATGCTATT CATAAAACGCAGAAC AAAGACAGAGTTTTT GCTTCCTTCTTACTC TTTGGTATTAAAACC
700   L  E  F  H  G   M  N  I  D  Q   L  E  N  A  I   H  K  T  Q  N   K  D  R  V  F   A  S  F  L  L   F  G  I  K  T

2205  TCAGCTGATGTCCAT TTGAAACTTTGCAAA GATGAAACTTGTGAA GATGCTGGTGTAGTC TTTGTACTTGGTGGT GACAATGAAATGCCA TGGCCCTTCGATAGA
735   S  A  D  V  H   L  K  L  C  K   D  E  T  C  E   D  A  G  V  V   F  V  L  G  G   D  N  E  M  P   W  P  F  D  R

2310  ACGTACAAGATGGAT ATTACCAATGTTCTA CATAAAATGCACATT CCTTTGGAAGATCTG TATGTTCATGGAAGC ACTATTCACCTTGAA GTAAAGATTGAATTC
770   T  Y  K  M  D   I  T  N  V  L   H  K  M  H  I   P  L  E  D  L   Y  V  H  G  S   T  I  H  L  E   V  K  I  E  S

2415  GTAGATGGAAAAGTA TTAGATTCTAGCTCA TTGCCAGTTCCTTCA ATGATATATGTCCCA GCTAAAGAATTCACA AAAGAGATAGAAAAG
805   V  D  G  K  V   L  D  S  S  S   L  P  V  P  S   M  I  Y  V  P   A  K  E  F  T   K  E  I  E  K
```

FIG. 2. Sequence of functional units Ode and Odf from *O. dofleini* hemocyanin at the DNA and protein level. The underlined DNA sequence is complementary to probe WL4.

type" and "G-type," respectively, after the first nucleotide difference observed. As Fig. 1 shows, clones XT5 and XT53 are G-type; since they cover a "narrower" range of sequence, we shall not describe them further here; details have been given (7). An analysis of 27 clones revealed that 13 coded for the G-type and 14 for the A-type.

DISCUSSION

Presence of G-Type and A-Type Sequences. The earlier experimental evidence suggested the presence of only one kind of polypeptide chain in *O. dofleini* hemocyanin. However, cDNA cloning and sequencing show that two slightly different polypeptide chains are encoded. In 1750 bases available from both G-type and A-type sequences, 57 nucleotide substitutions are found, which cause 34 amino acid changes, most of which are clustered in the putative linker regions between functional units. The presence of two sequences resolves apparent contradictions between DNA and protein sequencing results. The sequence for the N terminus of Odg obtained by protein sequencing was Thr-Val-Gly-Asp-Ala-Ile-Ile-Arg-Lys (9). However, by DNA sequencing, Glu-Ala-Val-Arg-Gly-Thr-Ile-Ile-Arg-Lys was found initially (6). Now it is clear that the N-terminal peptide sequence is that encoded by the G-type sequence. The presence of two cDNAs clearly tells us that there were two hemocyanin genes present in the animal from which the mRNA used for construction of the cDNA libraries was obtained. The hemocyanin used for sequencing came from a different animal. Since we do not know the number of hemocyanin genes in the genome of *O. dofleini*, we cannot tell whether these two genes are the result of a very recent gene duplication or whether they represent an allele.

Interspecies and Intraspecies Comparison of Functional Unit Sequences. Four complete sequences are now known for functional units of molluscan hemocyanins. In addition to Ode, Odf, and Odg, functional unit d of *Helix pomatia* β_c-hemocyanin (designated Hpd) is entirely sequenced (9). Almost complete sequences are available for functional unit g of *H. pomatia* β_c-hemocyanin (designated Hpg) (11), and

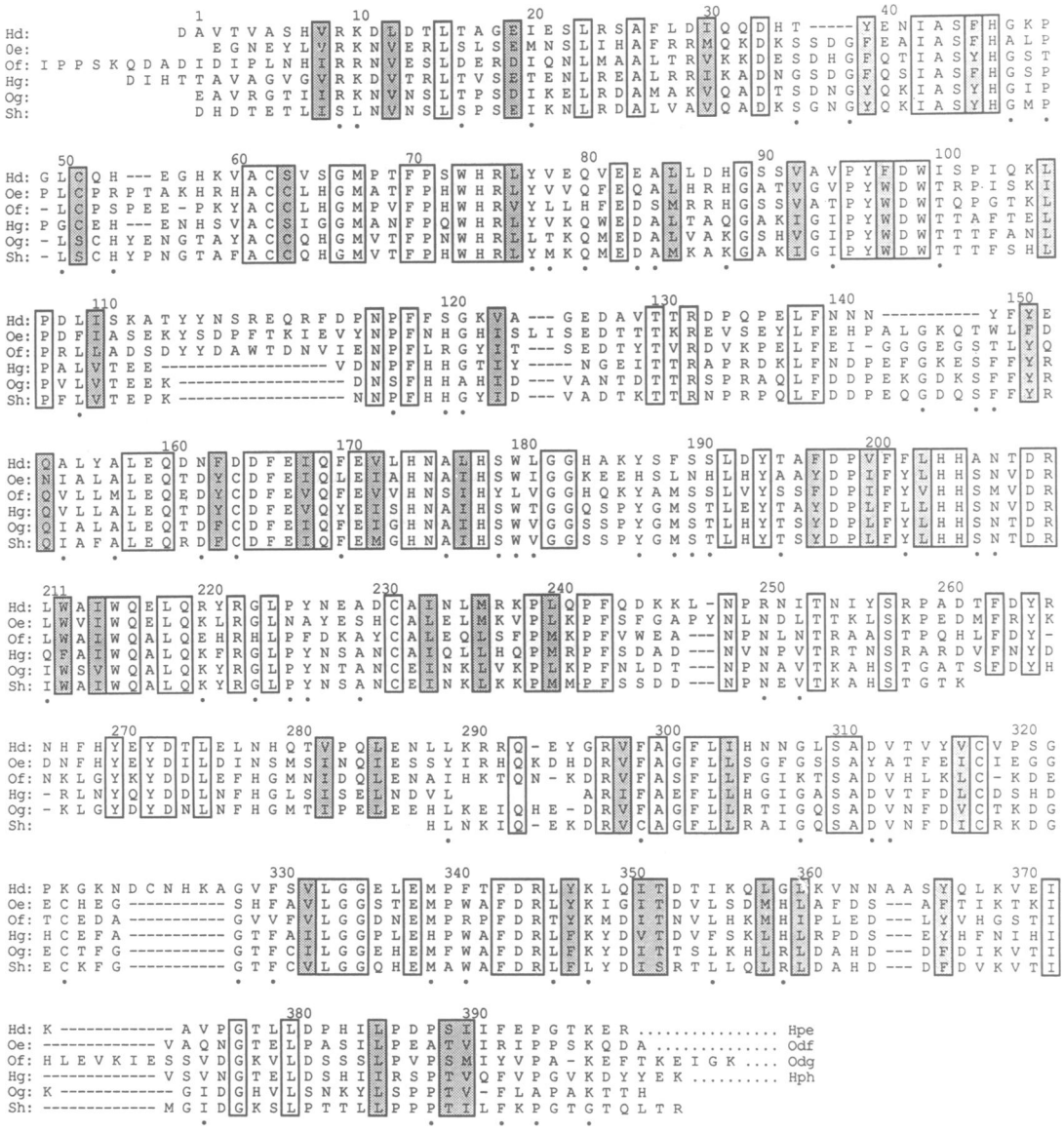

FIG. 3. Alignment of complete or almost complete sequences for functional units of molluscan hemocyanins. The sequences are from domains Ode (Oe), Odf (Of), Odg (Og), Soh (Sh), Hpd (Hd), and Hpg (Hg) as defined in text. Unshaded boxes show identical residues, and shaded boxes show similar residues shared by all six sequences. Identical or similar residues shared by five of the six sequences are marked with bullets. The following groups of amino acid residues are considered to be similar: E and D; N and Q; S and T; S and C; M, L, I, and V; F, Y, and W; H, K, and R. Numbering refers to the sequence of Odg.

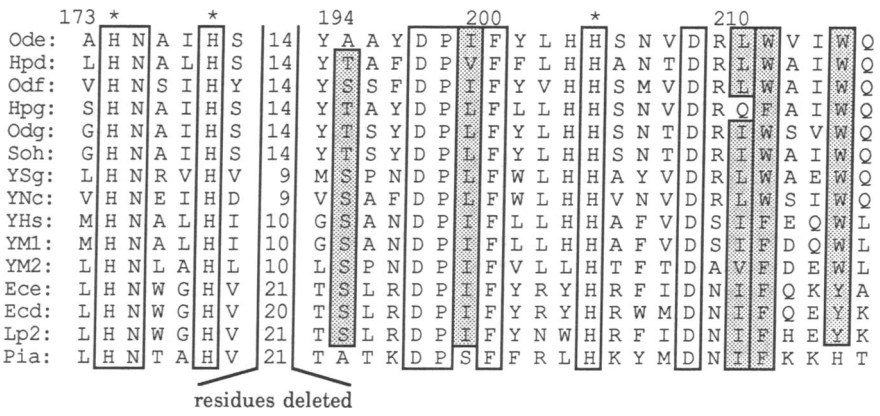

FIG. 4. Alignment of sequences around the putative Copper B site in molluscan and arthropodan hemocyanins and tyrosinases. Histidine residues potentially serving as copper ligands are marked with an asterisk. Code: Ysg, *Streptomyces glaucescens* tyrosinase; YNc, *Neurospora crassa* tyrosinase; YM1, mouse tyrosinase 1; YM2, mouse tyrosinase 2; YHs, human tyrosinase; Ece, *Eurypelma californicum* hemocyanin chain e; Ecd, *Eurypelma californicum* hemocyanin chain d; Pia, *P. interruptus* hemocyanin chain a; Lp2, *Limulus polyphemus* hemocyanin component II. These sequences were obtained from the Protein Identification Resource, National Biomedical Research Foundation (Release 24.0) sequence data bank. Protein domains of molluscan hemocyanins are as in the text. Numbers are as in Fig. 3.

functional unit h of *Sepia officinalis* hemocyanin (designated Soh) (12). The alignment of these six sequences is shown in Fig. 3. Overall they are remarkably conserved and show a high degree of similarity. Ninety positions are occupied by identical residues and 44 by isofunctional residues. In addition partial sequences are known for functional unit f (Sof) and g (Sog) of *S. officinalis* hemocyanin (12, 13).

Comparison between pairs of functional units reveals a significant result. When we compare noncorresponding units between species or different units within one species, we find an average identity of only 42% ± 3.7%. But functional units occupying corresponding positions in two different hemocyanins are much more similar: Soh vs. Odg = 75%; Soh vs. Hpg = 53%; and Odg vs. Hpg = 53%. The higher identity score between the two cephalopods, as compared with cephalopod vs. gastropod, is in agreement with paleontological evidence, which suggests that the squid and the octopi diverged about 200–300 million years ago, in contrast to the divergence of cephalopods from gastropods (about 500–600 million years ago) (14). Because the degree of identity in corresponding units is higher than in noncorresponding ones, we conclude that molluscan hemocyanin polypeptide chains already existed in their multiunit form before the molluscan orders diverged from one another.

Structure of the Active Site. From the determination of the crystal structure of *Panulirus interruptus* hemocyanin (3), the structure of the copper-binding sites in arthropodan hemocyanins is well known. The two copper-binding sites (A and B) are very similar, each copper atom having three histidine ligands: two are separated by three amino acid residues with the side chains projecting from the same side of an α-helix, and the third is furnished by a second α-helix running antiparallel to the first one.

The only similarities between molluscan and arthropodan hemocyanins are found in a small region that corresponds to the "Copper B" binding site in *P. interruptus* hemocyanin (5, 6). A very similar sequence is also conserved in tyrosinases. Fig. 4 shows an alignment of sequences around the B site in tyrosinases and molluscan and arthropodan hemocyanins. The positions of the Copper B ligands of *Panulirus* hemocyanin are marked with asterisks. Therefore, it is very likely that histidines-174,-178, and -205 (numbered as in Fig. 3) also serve as ligands for copper in the Copper B site in tyrosinases and molluscan hemocyanins. Site-directed mutagenesis of His-174 and His-205 (numbered as in Fig. 4) in *Streptomyces glaucescens* tyrosinase led to inactivation of the enzyme (15, 16). The His-174 mutant has been shown to retain only one mole of copper per mole of protein (15). Photoinactivation of *N. crassa* tyrosinase leads to destruction of His-205 and the loss of one copper atom (17). The sum of the evidence strongly suggests that tyrosinases and both kinds of hemocyanin share one type of copper-binding site.

This leaves the question as to what the ligands of Copper A may be in molluscan hemocyanins. The current spectroscopic evidence (reviewed in refs. 1 and 2) suggests that each copper atom is bound by at least two, possibly three, histidine side chains. It is reasonable to assume that the copper ligands would be conserved in all sequences of molluscan hemocyanins; therefore, the obvious thing to do is to look for conserved histidines. In an earlier paper from our laboratory, His-46, His-53, and His-74 have been suggested as candidates (6). However, in contrast to His-74 and His-46, His-53 is not

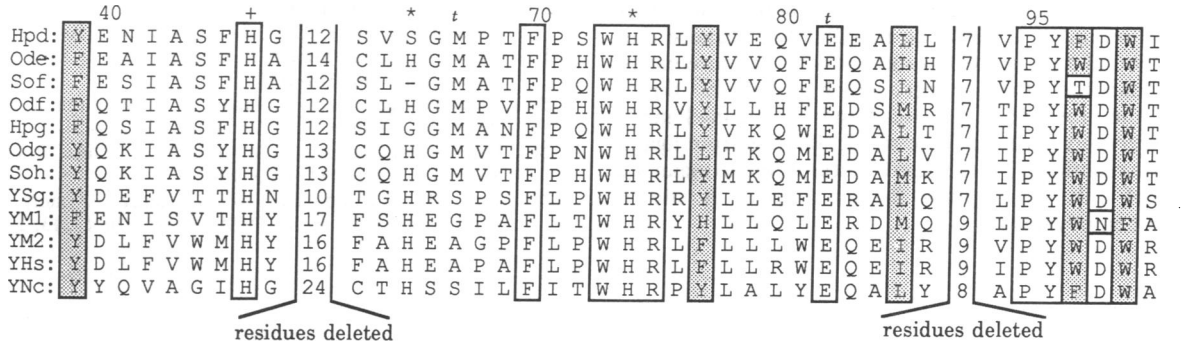

FIG. 5. Alignment of sequences around the putative copper A site in molluscan hemocyanins and tyrosinases. The code is the same as in Fig. 4. Numbers are as in Fig. 3.

conserved in all molluscan hemocyanin sequences (Fig. 3) and therefore is probably not involved in copper binding. Site-directed mutagenesis of His-74 in *S. glaucescens* tyrosinase leads to inactivation of the enzyme with loss of one copper per molecule of protein (15). Besides His-74, there is only one more conserved histidine residue in the vicinity where we would expect the A site: His-46. This residue is also conserved in all tyrosinases (Fig. 5). If we accept this histidine as a likely second ligand, what is the third? There is only one other conserved histidine residue: His-204, which is positioned next to the putative Copper B ligand His-205 and therefore would be very close to the active site. It is conserved in tyrosinases as well but apparently is not destroyed by photoinactivation in *N. crassa* tyrosinase (17), which argues against its role as a ligand for Copper A. In contrast, His-65, which is conserved in all tyrosinases but not all molluscan hemocyanins (Fig. 5), was also shown to be involved in copper binding in *S. glaucescens* tyrosinase by site-directed mutagenesis (16).

Because there are no other conserved histidine residues in molluscan hemocyanins, we have to consider now three possibilities:

(*i*) First, the third ligand for Copper A may be a histidine but may not be conserved between individual units of molluscan hemocyanins. It is difficult to guess which histidine then would be a good candidate, since there are several in the vicinity. However, it is hard to accept that in an otherwise highly conserved protein, the structure of the active site would not be conserved.

(*ii*) The third side chain involved in Copper A binding may be another kind of amino acid residue. Cysteine and methionine side chains are known to serve as copper ligands in other copper proteins [e.g., plastocyanin (18)]. Cysteine has been considered earlier as a possible copper ligand but has been subsequently ruled out (19). Extended x-ray absorption fine structure (EXAFS) studies have been argued to also exclude sulfur ligation of copper in hemocyanin (20), but EXAFS studies of plastocyanin are only able to detect cysteine sulfur as a copper ligand, not methionine (21). Therefore, methionine cannot yet be excluded as a possible copper ligand in molluscan hemocyanins. Indeed, there is one methionine (Met-67) that is conserved in all molluscan hemocyanin functional units but not in tyrosinases (Fig. 5). As a close neighbor of His-65, which has been implicated in copper binding in tyrosinases, it could be substituting as a Copper A ligand in molluscan hemocyanins. Aspartate and glutamate side chains (Glu-82 is conserved in all sequences in Fig. 5) are also known to serve as metal ligands in metalloproteins [e.g., hemerythrin (22)] but have not yet been observed as copper ligands in proteins.

(*iii*) The third possibility is that there are only two Copper A ligands, both of them histidines. There is some experimental evidence that this could be the case. It is found that, unlike in arthropodan hemocyanins, one of the two coppers in molluscan hemocyanins is removed much more easily than the second one (23). It could be thought that because one copper is bound only by two ligands, it is much more easily removed than the other copper, which is held by three. However, the presence of one weak and two strong ligands could conceivably give the same result. Results from ligand substitution reactions (24) can also be interpreted to mean the presence of only two ligands for the Copper A site in molluscan hemocyanins.

In summary it can be said that whereas the ligands of the Copper B site can now be identified with confidence in molluscan hemocyanins, the full structure of the active site in molluscan hemocyanin remains an unsolved question. X-ray studies (25) hopefully can provide an answer to this problem.

This research was supported by Grant No. DMB-8818697 from the National Science Foundation. K.E.v.H. acknowledges an American Cancer Society Research Professorship.

1. Van Holde, K. E. & Miller, K. I. (1982) *Q. Rev. Biophys.* **15**, 1–129.
2. Ellerton, H. D., Ellerton, N. F. & Robinson, H. A. (1983) *Prog. Biophys. Mol. Biol.* **41**, 143–248.
3. Gaykema, W. P. J., Hol, W. G. J., Vereijken, J. M., Soeter, N. M., Bak, H. J. & Beintema, J. J. (1984) *Nature (London)* **309**, 23–29.
4. Himmelwright, R. S., Eickman, N. C., LuBien, C. D. & Solomon, E. I. (1980) *J. Am Chem. Soc.* **102**, 5377–5388.
5. Drexel, R., Siegmund, S., Schneider, H. J., Linzen, B., Gielens, C., Preaux, G., Kellerman, J. & Lottspeich, F. (1987) *Biol. Chem. Hoppe-Seyler* **368**, 617–635.
6. Lang, W. H. (1988) *Biochemistry* **27**, 7276–7282.
7. Lang, W. H. (1990) Ph.D. thesis (Oregon State Univ., Corvallis, OR).
8. Miller, K. I. & van Holde, K. E. (1982) *Comp. Biochem. Physiol.* **73B**, 1013–1018.
9. Lamy, J., Leclerc, M., Sizaret, P.-Y., Lamy, J. N., Miller, K. I., McParland, R. & van Holde, K. E. (1987) *Biochemistry* **26**, 3509–3518.
10. Lang, W. H. & van Holde, K. E. (1990) in *Invertebrate Dioxygen Carriers*, eds. Preaux, G. & Lontie, R. (Univ. of Leuven Press, Louvain, Belguim), pp. 135–140.
11. Xin, X.-Q., Gielens, C., Witters, R. & Preaux, G. (1990) in *Invertebrate Dioxygen Carriers*, eds. Preaux, G. & Lontie, R. (Univ. of Leuven Press, Louvain, Belgium), pp. 113–117.
12. Declercq, L., Witters, R. & Preaux, G. (1990) in *Invertebrate Dioxygen Carriers*, eds. Preaux, G. & Lontie, R. (Univ. of Leuven Press, Louvain, Belgium), pp. 131–134.
13. Top, A., Gielens, C., Witters, R., van Beumen, J. & Preaux, G. (1990) in *Invertebrate Dioxygen Carriers*, eds. Preaux, G. & Lontie, R. (Univ. of Leuven Press, Louvain, Belgium), pp. 119–123.
14. Runnegar, B. & Pojeta, J., Jr. (1985) in *The Mollusca Vol. 10: Evolution*, eds. Trueman, E. R. & Clarke M. R. (Academic, Orlando, FL), pp. 1–57.
15. Huber, M. & Lerch, K. (1988) *Biochemistry* **27**, 5610–5615.
16. Jackman, M., Huber, M. & Lerch, K. (1989) *Highlights Mod. Biochem.* **1**, 281–290.
17. Dietler, C. & Lerch, K. (1982) in *Oxidases and Related Redox Systems*, eds. King, T. E., Morrison, M. & Mason, H. S. (Pergamon, Oxford), pp. 305–317.
18. Colman, P. M., Freeman, H. C., Guss, J. M., Murata, M., Norris, V. A., Ramshaw, J. A. M. & Venkatappa, M. P. (1978) *Nature (London)* **272**, 319–323.
19. Lontie, R. (1958) *Clin. Chim. Acta* **3**, 68–71.
20. Woolery, G. L., Powers, L., Winkler, M., Solomon, E. I. & Spiro, T. G. (1984) *J. Am Chem. Soc.* **106**, 86–92.
21. Scott, R. A., Hahn, J. E., Doniach, S., Freeman, H. C. & Hodgson, K. O. (1982) *J. Am. Chem. Soc.* **104**, 5364–5369.
22. Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. (1984) *J. Am. Chem. Soc.* **106**, 618–622.
23. Himmelwright, R. S., Eickman, N. C. & Solomon, E. I. (1978) *Biochem. Biophys. Res. Commun.* **81**, 243–247.
24. Preaux, G., Gielens, C., Witters, R. & Lontie, R. (1988) *Bull. Soc. Chim. Belg.* **97**, 1073–1044.
25. Cuff, M. E., Hendickson, W. A., Lamy, J., Lamy, J., Miller, K. I., van Holde, K. E. (1990) *J. Mol. Biol.* **213**, 11–15.