Structure of arthropod hemocyanin

Henk J. Bak, Ben Neuteboom, Peter A. Jekel, Nell M. Soeter, Johan M. Vereijken and Jaap J. Beintema

Biochemisch Laboratorium, Rijksuniversiteit Groningen, Nijenborgh 16; 9747 AG Groningen, The Netherlands

Received 23 June 1986

Hemocyanins are large multi-subunit copper proteins that transport oxygen in many arthropods and molluscs. The amino acid sequence of subunit a of *Panulirus interruptus* hemocyanin (657 residues) has been completed and fitted to the electron-density map (3.2 Å resolution). Comparison of amino acid sequence data for several different hemocyanin subunits of arthropod species indicated that the general features of the polypeptide architecture as found in spiny lobster hemocyanin occur in all arthropods. This structure must therefore be at least as old as the estimated time of divergence of crustaceans and chelicerates, 540–600 million years ago.

(Panulirus interruptus) *Hemocyanin* *Amino acid sequence* *Molecular evolution*

1. INTRODUCTION

Hemocyanins occur freely dissolved in the hemolymph of many molluscs and arthropods, but their respective molecular architectures differ completely. Molluscan hemocyanins are cylindrical oligomers with 10–20 subunits of about 400 kDa each. Every subunit contains seven or eight domains with one binuclear copper site per domain. Arthropod hemocyanins are composed of hexamers or multi-hexamers with subunits of about 75 kDa containing one binuclear copper site [1]. Similar binuclear copper sites have been found in tyrosinase, ceruloplasmin, laccase and ascorbate hydroxylase [2]. Therefore, structural information on the binuclear copper site in arthropod hemocyanin may lead to a better understanding of these sites in other proteins.

Due to the great length of the polypeptide chains, the heterogeneity of the subunits, and the complex quaternary structures, the structure of the hemocyanins has remained elusive for a long time. However, a lot of primary-structure information has been collected recently for two groups of arthropods, the crustaceans and chelicerates [3]. In addition, the three-dimensional structure of hemocyanin of the spiny lobster *Panulirus interruptus* has been determined at a resolution of 3.2 Å [4,5]. The amino acid sequence of subunit a of *Panulirus* hemocyanin (657 residues) has been completed and will be reported here together with the results of the amino acid sequence studies on subunits b and c, which were started later.

Dedicated to Professor Dr G. Braunitzer on the occasion of his 65th birthday

2. SPINY LOBSTER HEMOCYANIN

Hemocyanin of the spiny lobster *P. interruptus*, an arthropod belonging to the Crustacea, consists of heterogeneous hexamers containing at least three different subunits, named a, b and c, with molecular masses of about 75 kDa. After dissociation of the hexamers, the major subunits a and b have been found to account each for 40–45% of the monomers. The distribution of the subunits among the hexamers in vivo is not known. Reassociation experiments indicate that molecules with certain subunit compositions are formed preferentially [6].

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

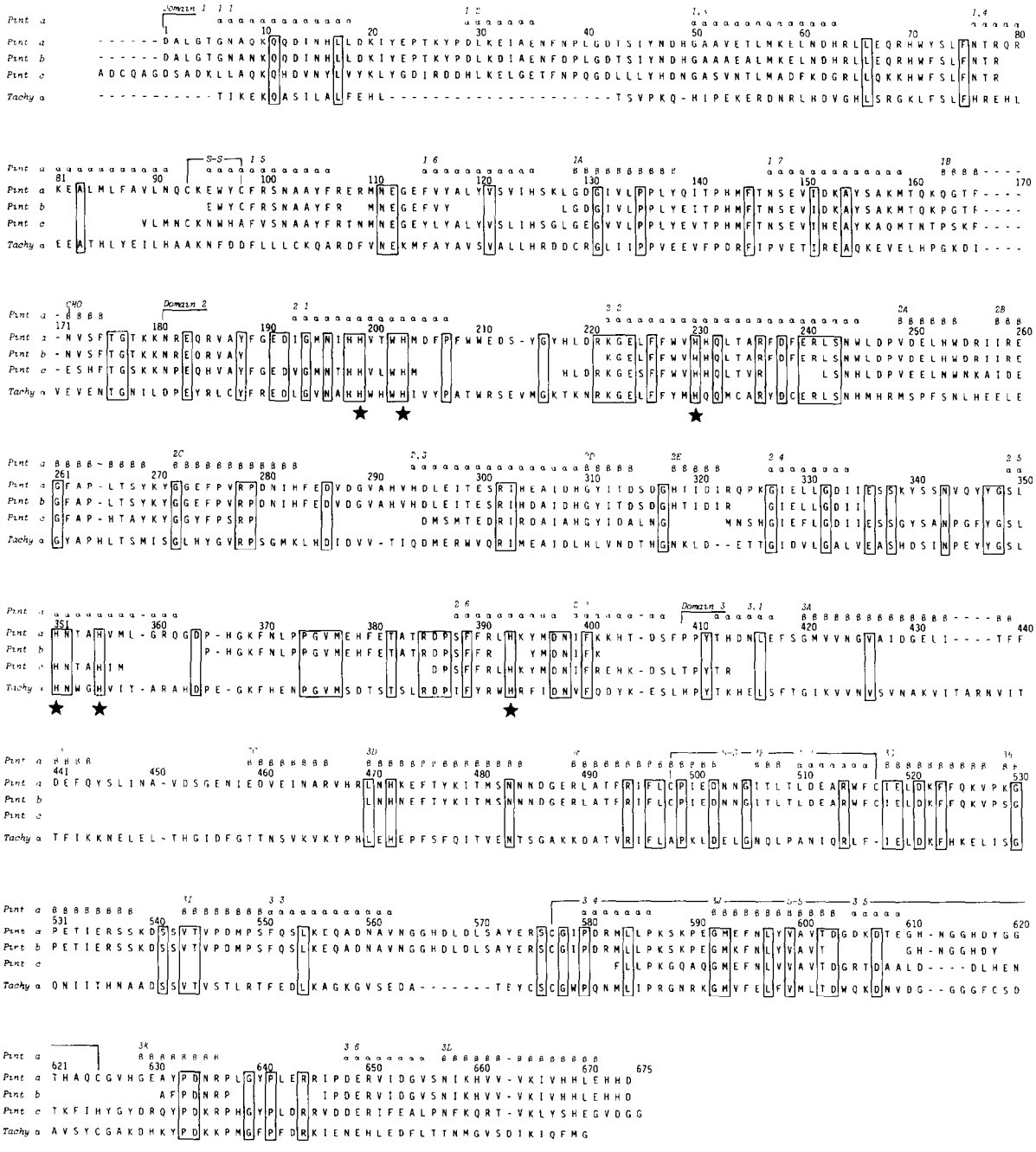

Fig.1. Alignment of arthropod hemocyanin amino acid sequences and secondary structure of *P. interruptus* hemocyanin. The numbering from fig.1 in [3] has been used. Complete and partial amino acid sequences are shown for a crustacean species, *P. interruptus* (*Pint*), subunits a, b and c, and for a chelicerate species, *Tachypleus tridentatus* (*Tachy*), subunit α [3]. The one-letter code for amino acid residues has been used. Residues identical for all nine subunits of arthropod hemocyanins sequenced so far [3] are boxed. Dashes represent gaps introduced for alignment, blank spaces are parts not yet determined. The secondary-structure elements (α and β), disulfide bridges (S-S), carbohydrate attachment site (CHO), copper ligands (★) and domain borders of *Panulirus* hemocyanin are indicated.

The three-dimensional structure of *Panulirus* hemocyanin has been determined at a resolution of 3.2 Å with crystals containing a mixture of subunits a and b [3–5]. Each polypeptide chain is folded into three distinct domains, the first and the second domain being mainly helical and globular whereas the third domain contains a β-barrel structure and two long loops.

The first domain contains the carbohydrate moiety present in chains a and b, which is linked to asparagine 172 (numbering according to fig.1). Chain c does not contain carbohydrate at this position. The second domain contains the active site, where each copper atom is ligated by three histidines, two of which are provided by a -His-X-X-X-His- sequence located in a helix. The third ligand of each copper ion comes from another helix.

The third and largest domain contains an antiparallel β-barrel consisting of seven strands. The two long loops extending from the center of this domain appear to function as arms holding the three domains together. A remarkable feature is the topology of the β-barrel. It is identical to that of the β-barrels found in immunoglobulins and in Cu,Zn-superoxide dismutase [4]. However, hardly any similarities have been observed in the amino acid sequences of these β-barrels. Therefore, the similarity is probably the result of convergent evolution.

3. AMINO ACID SEQUENCE COMPARISONS

Complete and partial sequences of two crustacean and five chelicerate chains could be aligned and correlated with the three-dimensional structure of *Panulirus* hemocyanin. The general features of the polypeptide architecture found in spiny lobster hemocyanin appeared to occur in all arthropods, although some differences between crustacean and chelicerate hemocyanins were found. This structure must therefore be at least as old as the estimated time of divergence of crustaceans and chelicerates, 540–600 million years ago. These findings were published recently [3], when only an incomplete amino acid sequence of chain a of *P. interruptus* hemocyanin was available.

Chain a of spiny lobster hemocyanin is the only crustacean subunit sequenced completely hitherto. Partial sequences of chains b and c of *Panulirus* and chain b of the crayfish *Astacus leptodactylus* [3] are available. The comparison will be extended now. It will offer the reader a more complete picture of nature's creation of 'the third way of oxygen transport'.

The amino acid sequences of the three subunits of *P. interruptus* hemocyanin, and as a representative of the chelicerate chains, chain α of *Tachypleus tridentatus* hemocyanin have been aligned in fig.1. The polypeptide chain lengths are roughly the same: the chelicerate subunits have about 625 residues and the crustacean hemocyanins about 660. So far, 110 positions (17%) are identical for all nine chains. This figure increases to 29% if we include isofunctional residues [3]. The identical residues are often clustered, for example in the stretches 176–207, 216–243, 347–356, 374–400, 576–607 and 633–644. The most strongly conserved regions are those near the histidine residues that have been identified as copper ligands. In table 1 of [3], the differences between the amino acid sequences of the separate domains and of the complete subunits were summarized. However, only an incomplete sequence of *Panulirus* chain a was available at that time.

Table 1 in this paper summarizes the percentage differences between the complete sequences of the subunits and that of chain a. The largest differences were found in domain 1 and domain 3, 80 and 70%, respectively, whereas the smallest dif-

Table 1

Comparison of the percentage differences between amino acids of complete subunits

	Pint a	Eury d	Eury c	Lim II
Eury d	70%			
Eury e	70%	45%		
Lim II	69%	47%	45%	
Tachy α	69%	46%	43%	43%
NPC	647–666	618–632		

The alignment of fig.1 (and fig.1 of [3]) was used. Percentages were calculated as before [3]. Data were taken directly from this paper and from table 1 of [3]. *Abbreviations:* Pint, *P. interruptus*; Eury, *Eurypelma californicum*; Lim, *Limulus polyphemus*; Tachy, *Tachypleus tridentatus*; NPC, the numbers of positions compared

Table 2

Comparison of the percentage differences between amino acids of partially sequenced crustacean subunits

	% Seq.	Pint a		Pint b		Pint c	
		NPC	% diff.	NPC	% diff.	NPC	% diff.
Pint b	72	474	3				
Pint c	64	428	41	323	42		
Asta b	72	478	31	377	31	341	38

The alignment of fig.1 (and fig.1 of [3]) was used. The percentages were calculated as before [3]. The percentage of residues sequenced, the number of positions compared (NPC) as well as the percentage differences between the subunits are given. *Abbreviations:* as in table 1, and Asta, *Astacus leptodactylus*

ference was found in domain 2, only 60% (not shown).

Only partial amino acid sequence information is available for chains b and c of *Panulirus* and chain b of *A. leptodactylus* [3]. Table 2 summarizes the percentage differences between these subunits in the positions which could be compared, as well as the percentage of residues sequenced to date. Subunits a and b are almost identical, differing only by a few percent, whereas subunit c is the most deviating of all four crustacean chains, having 40% different residues. Moreover, in contrast to subunit b of *Panulirus* and probably subunit b of *Astacus*, subunit c does not possess disulfide bridges at the same positions as chain a since Cys-98 and Cys-625 are both replaced by other residues. Since *Astacus* subunit b resembles *Panulirus* subunit a and b more closely than does subunit c, a gene duplication generating subunits a/b and c has taken place presumably before the species differentiation of *Astacus* and *Panulirus*.

ACKNOWLEDGEMENTS

We thank Dr R.N. Campagne for carefully reading this manuscript. This work was supported in part by The Netherlands Foundation for Chemical Research (SON) with financial aid from The Netherlands Organization for the Advancement of Pure Research (ZWO).

REFERENCES

[1] Van Holde, K.E. and Miller, K.I. (1982) Q. Rev. Biophys. 15, 1–129.

[2] Urbach, F.L. (1981) in: Metal Ions in Biological Systems (Sigel, H. ed.) vol.13, pp.73–115, Marcel Dekker, New York.

[3] Linzen, B., Soeter, N.M., Riggs, A.F., Schneider, H.-J., Schartau, W., Moore, M.D., Yokata, 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. and Hol, W.G.J. (1985) Science 229, 519–524.

[4] Gaykema, W.P.J., Hol, W.G.J., Vereijken, J.M., Soeter, N.M., Bak, H.J. and Beintema, J.J. (1984) Nature 309, 23–29.

[5] Gaykema, W.P.J., Volbeda, A. and Hol, W.G.J. (1986) J. Mol. Biol. 187, 255–275.

[6] Van den Berg, A.A., Gaastra, W. and Kuiper, H.A. (1977) in: Structure and Function of Hemocyanin (Bannister, J.V. ed.) pp.6–12, Springer, Berlin.