Fietzek, P. P., & Rexrodt, F. W. (1975) *Eur. J. Biochem. 59*, 113.
Fukae, M., & Mechanic, G. L. (1980) *J. Biol. Chem. 255*, 6511.
Fuller, F., & Boedtker, H. (1981) *Biochemistry 20*, 999.
Galloway, D. (1982) in *Collagen in Health and Disease* (Weiss, J. B., & Jayson, M. I. V., Eds.) p 528, Churchill Livingston, New York.
Gibson, D. R., Gracy, R. W., & Hartman, F. C. (1980) *J. Biol. Chem. 255*, 9369.
Golub, E., & Katz, E. P. (1977) *Biopolymers 16*, 1357.
Housley, T., Tanzer, M. L., Hensen, E., & Gallop, P. M. (1975) *Biochem. Biophys. Res. Commun. 67*, 824.
Hulmes, D. J. S., & Miller, A. (1971) *Nature (London) 293*, 239.
Jany, K. D., Keil, W., Meyer, H., & Kiltz, H. H. (1976) *Biochim. Biophys. Acta 453*, 62.
Jörnvall, H. (1974) *FEBS Lett. 38*, 329.
Katz, E. P. (1970) *Biopolymers 9*, 745.
Katz, E. P., & Li, S. T. (1973a) *J. Mol. Biol. 73*, 351.
Katz, E. P., & Li, S. T. (1973b) *J. Mol. Biol. 80*, 1.
Li, S. T., Sullman, S., & Katz, E. P. (1981) in *The Chemistry and Biology of Mineralized Tissues* (Veis, A., Ed.), pp 69–73, Elsevier/North-Holland, Amsterdam.
Mechanic, G. L. (1974) *Biochem. Biophys. Res. Commun. 56*, 923.
Oshima, G., Shimabukuro, H., & Nagasawa, K. (1979) *Biochim. Biophys. Acta 567*, 392.
Piez, K. A., & Trus, B. L. (1981) *Biosci. Rep. 1*, 801.
Rauterberg, J., Fietzek, P. P., Rexrodt, F. W., Becker, U., Stark, M., & Kuhn, K. (1972) *FEBS Lett. 21*, 7.
Robins, S. P., & Duncan, A. (1983) *Biochem. J. 215*, 175.
Ruggerri, A., Benazzo, F., & Reale, E. (1979) *J. Ultrastruct. Res. 68*, 101.
Stinson, R. H., & Sweeny, P. R. (1980) *Biochim. Biophys. Acta 621*, 158.
Tanzer, M. L., Housley, T., Berube, L., Fairweather, R. Franzblau, C., & Gallop, P. M. (1973) *J. Biol. Chem. 248*, 393.
Torchia, D. A. (1982) *Methods Enzymol. 82A*, 174.
Yamauchi, M., Noyes, C., Kuboki, Y., & Mechanic, G. L. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 7684.
Yamauchi, M., Katz, E. P., & Mechanic, G. L. (1986) *Biochemistry 25*, 4907.
Yamauchi, M., London, R. E., Guenat, C., Hashimoto, F., & Mechanic, G. L. (1987) *J. Biol. Chem.* (in press).
Yonath, A., & Traub, W. (1969) *J. Mol. Biol. 43*, 465.

Octopus dofleini Hemocyanin: Structure of the Seven-Domain Polypeptide Chain†

Josette Lamy,‡ Michèle Leclerc,‡ Pierre-Yves Sizaret,‡ Jean Lamy,‡ Karen I. Miller,§ Reginald McParland,§ and K. E. van Holde*,§

Laboratoire de Biochimie, Faculté de Pharmacie, 37032 Tours Cédex, France, and Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331

Received October 9, 1986; Revised Manuscript Received December 30, 1986

ABSTRACT: *Octopus dofleini* hemocyanin exists as a native molecule composed of 10 identical 350-kDa polypeptide chains. It is homogeneous by sedimentation methods, by N-terminal sequence homology, and by the presence of a single immunoprecipitation peak, corresponding to the subunit. Electron microscopy shows seven globular domains, and there are seven immunologically distinct products (*Od* 1 to *Od* 7) following limited proteolysis. By combining the results of proteolysis and immunological methods with amino acid sequencing, the order of the domains within the subunit has been established with one remaining uncertainty. Trypsin and serine protease from *Staphylococcus aureus* produce fragments corresponding to domain *Od* 1 and the six-domain fragment (2,3,4,5,6,7). The amino acid sequence and the digestion conditions for domain *Od* 1 are the same as those that Takagi [Takagi, T. (1986) in *Invertebrate Oxygen Carriers* (Linzen, B., Ed.) pp 259–262, Springer-Verlag, Berlin and Heidelberg] found for the C-terminal domain of *Paroctopus* hemocyanin. Thus, the sequence becomes N-(2,3,4,5,6,7)-1-C. Trypsin, subtilisin, and other enzymes produce fragments (1,2,5) and (3,4,7). This information permits ordering the domains additionally as follows: N-[(3,4,7),6]-(2,5)-1-C. α-Chymotrypsin produces fragments (1,2,5,6) and (2,5,6), allowing as the next step this sequence: N-(3,4,7)-6-(2,5)-1-C. Digestion with subtilisin produces the unique fragment (4,7), and a secondary digestion of fragment (3,4,7) by trypsin yields the fragment (3,4). The sequence now is N-(3-4-7)-6-(2,5)-1-C. Amino acid sequencing demonstrates that fragment (4,7) has the same N-terminal sequence as the whole subunit. Thus, the sequence now becomes N-7-4-3-6-(2,5)-1-C. The uncertainty of the order of the (2,5) fragment remains.

Molluscan hemocyanins are composed of 10 or 20 subunits of 350–450 kDa arranged to form hollow cylinders with molecular weights of approximately 3.5×10^6 to 9.0×10^6. Most of the information on the structure of molluscan hemocyanin has been obtained from studies of *Helix pomatia* and other gastropods (Ghiretti-Magaldi et al., 1981; van Bruggen et al., 1981a; van Holde & Miller, 1982; Lontie, 1983). These studies have revealed that gastropod hemocyanins are composed of a mixture of several different subunits, a fact that has made them difficult to study, because their association reactions are complex (Brouwer et al., 1979). One must sort out individual subunit types and determine their role in the

† This work was supported by RCP 080816 of Centre National de la Recherche Scientifique (CNRS) (J.L.) and Grant DMB 85 17310 from the National Science Foundation (K.E.v.H. and K.I.M.).
‡ Faculté de Pharmacie.
§ Oregon State University.

0006-2960/87/0426-3509$01.50/0

structure of the whole molecules, as well as examining the structure of the subunit itself. Despite these complexities, Gielens and co-workers have now demonstrated for *Helix* hemocyanin that the structure of the individual 450-kDa subunits of hemocyanin β consist of eight different oxygen binding domains, each about 55 kDa, arranged in a specific sequence along the polypeptide chain (Gielens et al., 1977, 1980, 1981; De Sadeleer et al., 1983).

Unlike gastropod hemocyanins, which can form 9×10^6 dalton whole molecules, the cephalopod hemocyanins are restricted to molecules roughly half this size, with molecular weights from 3.3×10^6 to 3.8×10^6, corresponding to sedimentation coefficients of 51 S to 59 S. Nevertheless, they are also quite variable and complex. *Nautilus* has a molecular weight of around 3.5×10^6 and probably contains more than one type of subunit (Bonaventura et al., 1981). *Sepia* has a molecular weight of around 3.8×10^6 and contains 10 identical subunits, each consisting of eight domains (Gielens et al., 1983; Wichertjes et al., 1986a,b). Similarly, *Loligo* has a molecular weight of 3.8×10^6 and has eight domains per subunit (van Holde & Miller, 1982; van Bruggen et al., 1981b). Therefore, it was of particular interest to discover whether *Octopus* hemocyanin had a similar structure.

Because the association behavior of *Octopus dofleini* hemocyanin is fully reversible, it seemed likely that it is a somewhat simpler system than has been the case for most other molluscan hemocyanins (van Holde & Miller, 1985). Electrophoretic and sedimentation equilibrium studies supported the existence of a single type of subunit. Furthermore, the low molecular weight of the subunit (350 000) suggested a smaller number of oxygen binding domains (Miller & van Holde, 1983; van Holde & Miller, 1982; Herskovits & Villanueva, 1986). These suspicions were confirmed independently by Lamy et al. (1986) for *O. dofleini* and Gielens et al. (1986) for *Octopus vulgaris* and presented simultaneously at Tutzing, West Germany, at the meeting Invertebrate Oxygen Binding Proteins in August 1985. They determined that indeed there was a single type of polypeptide chain containing seven immunologically distinct oxygen binding domains. The results of the two groups were in perfect agreement on this point.

This paper reports the details of the subunit structure of *O. dofleini* hemocyanin, including the sequence of domains within the subunit and immunological electron microscopic and amino acid sequencing evidence for their unique structure. This work was aided by the determination of a partial amino acid sequence for the C-terminal domain by Takagi (1986).

Materials and Methods

Hemocyanin. (*A*) *Taxonomy.* Since we were planning to produce amino acid sequence data, it was important that there be complete taxonomic information for this *Octopus*.

The giant octopus *Octopus dofleini* of the northern Pacific coast of the U.S. is the only very large octopus found in this region, so there is no possibility of confusion with any other species of *Octopus*. Although it is not often made, the taxonomists do make a distinction between three subspecies of *O. dofleini* (Pickford, 1964). The first, *Octopus dofleini dofleini*, is found in Japan; the second, *Octopus dofleini apollyon*, is found in Alaska; the third, *Octopus dofleini martini*, is found subtidally in Puget Sound and along the northern coast of the continental U.S. and British Columbia. Thus we have been working with *O. dofleini martini*.

(*B*) *Preparation of Samples.* Blood samples were obtained from *Octopus* and purified by gel filtration on Bio-Gel A-5m (Rio-Rad) as described previously (Miller & van Holde, 1982). The purification buffer was 0.1 *I* (ionic strength) Tris-HCl,[1] pH 7.65, 10 mM $CaCl_2$, and 50 mM $MgCl_2$. Dissociated 11S subunits were prepared by dialyzing against 0.1 *I* Tris-HCl, pH 8.0, and 10 mM EDTA.

Antisera. Rabbit antisera directed against native and dissociated hemocyanin were prepared according to a previously reported method (Lamy et al., 1979a). Domain-specific antisera were prepared as follows. First, a line immunoelectrophoresis of a partially purified hemocyanin domain was carried out vs. an antiserum specific for the whole polypeptide chain. Second, at the completion of the electrophoresis the agarose layer was covered with soft blotting paper and pressed according to Weeke (1973). Third, the precipitate line was cut from the gel and homogenized with saline and complete Freund's adjuvant (1 volume of saline/1 volume of Freund's adjuvant). Finally, an amount of immunoprecipitate equivalent to 2 nmol of pure domain was injected into the rabbit. The same dose was given for the booster injections. Antisera obtained by this method were used, as such, without purification.

Immunoelectrophoreses. Crossed immunoelectrophoreses, line immunoelectrophoreses, and crossed line immunoelectrophoreses were carried out according to the methods of Axelsen et al. (1973). Antigen deposits of 45 pmol and 2 nmol were used for analytical purposes (crossed immunoelectrophoresis) and for immunogen preparation (line immunoelectrophoresis), respectively.

Electron Microscopy. Specimens for electron microscopy were prepared by negative staining with 2% uranyl acetate. The specimens were observed in a Jeol 1200 EX electron microscope at an accelerating voltage of 80 kV. In the microscope the grid was oriented upside down with the carbon film facing the electron beam and the molecules facing the emulsion side of the photographic film. This procedure restores the correct orientation of the molecule on the print.

Polyacrylamide Gel Electrophoresis. Polyacrylamide gel electrophoreses were carried out in vertical gel slabs according to the method of Davis (1964) with a LKB 2001 vertical electrophoresis unit. For semipreparative purposes, an amount of hydrolysate corresponding to 35 nmol of polypeptide chain was submitted to electrophoresis in a 7% polyacrylamide gel. At the end of this step, two strips of gel were cut on both sides of the gel and stained with Coomassie Blue G-250. Then protein bands were cut from the gel, put in a 0.05 M Tris-HCl buffer, pH 8.9, containing 10 mM EDTA, and gently homogenized with a glass rod. After an overnight diffusion the mixture was filtered, and the gel was washed with the same buffer. Finally, the protein solution was concentrated by vacuum dialysis. SDS gels were performed according to the method of Laemmli (1970). The subunit and fragments of the subunit were separated on 7% gels.

Fast Protein Liquid Chromatography (FPLC). Purification of fragments and domains occurred on the fast protein liquid chromatography apparatus of Pharmacia, by ion-exchange chromatography and gel permeation. Ion-exchange chromatography was carried out on a Mono Q HR 5/5 column in a 0.05 M Tris-HCl buffer, pH 8.9, containing 10 mM EDTA. Elution was performed in the presence of a NaCl gradient between 0 and 0.5 M. Gel permeation occurred on a Superose 12 HR 10/30 column, in the same buffer containing 0.15 M NaCl.

Proteases. Eight different proteases were used to cleave *Octopus* hemocyanin subunits for further study. These pro-

[1] Abbreviations: Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; EDTA, ethylenediaminetetraacetic acid; SDS, sodium dodecyl sulfate.

Table I: Information on Proteases Used To Cleave *O. dofleini* Hemocyanin

enzyme	EC no.	source	buffer conditions
bovine pancreas α-chymotrypsin	3.4.21.1	Serva (ref 17160)	0.05 M borate, pH 8.0
bovine pancreas trypsin	3.4.21.4	Sigma (ref T 1005)	0.05 M ammonium carbonate, pH 8.0
coagulation factor IIa (thrombin)	3.4.21.5	Boehringer Mannheim (ref 60 2400)	0.05 M ammonium carbonate, pH 8.3
bovine plasma fibrinolysine (plasmin)	3.4.21.7	Boehringer Mannheim (ref 60 2370)	0.06 M phosphate, pH 7.4
porcine pancreas elastase	3.4.21.11	Boehringer Mannheim (ref 10 4566)	0.05 M borate, pH 8.0
subtilisin	3.4.21.14	Boehringer Mannheim (ref 16 5905)	0.025 M ammonium carbonate, pH 8.2
S. aureus strain V8 serine protease	3.4.21.19	Sigma (ref P 8400)	0.05 M ammonium carbonate, pH 8.0
papain (from *Papaya latex*)	3.4.22.2	Sigma (ref P 3125)	0.1 M phosphate, pH 6.4, 3.3 mM EDTA, 11 mM Cys-HCl

teases and their corresponding digestion buffers are summarized in Table I.

Amino Acid Sequencing. Approximately 1 nmol each of *Octopus* hemocyanin, hemocyanin domain, or domain dimer was used for sequencing. Samples were prepared by dialysis against 10 mM NaCN in 50 mM NH_4HCO_3, pH ~8.0, and then twice against 25 mM NH_4HCO_3 alone. Samples were then lyophilized, dissolved in a small amount of H_2O, and lyophilized again to remove trace amounts of NH_4HCO_3.

Peptide sequencing was performed on an Applied Biosystems Model 470A gas-phase protein sequencer. The standard ABI 02NVAC or 02CPTH programs were used for coupling and cleavage at a cartridge temperature of 45 °C. Anilinothiazolinone (ATZ) amino acid derivatives were converted to phenylthiohydantoin (PTH) derivatives with trifluoroacetic acid (TFA) at a conversion temperature of 55 °C.

PTH amino acid derivatives were analyzed on a Beckman μ-flow HPLC system equipped with an IBM cyano column. Chromatography conditions were similar to those described by Applied Biosystems.

Results

Homogeneity of the Hemocyanin and Its Subunits

In the electron microscope, *O. dofleini* hemocyanin shows a structure similar to those of other cephalopods (van Bruggen et al., 1981a). These results are shown in Figure 1. There are circular profiles with a diameter of 30 nm and rectangular profiles of 15 × 35 nm. The rectangles can be seen to be composed of three layers. The fact that they are somewhat wider than the circular profile is probably an artifact of the microscopy. There appear to be two different "faces" in the circular profile. Whether this is due to some difference in the uptake of stain or to real polarity of the molecule is not clear. Van Bruggen et al. (1981a) show similar pictures of *O. vulgaris*. There is no obvious polarity of the side view as there is for gastropod hemocyanins. There is no evidence for more than one type of hemocyanin on the basis of these photographs.

O. dofleini hemocyanin can be dissociated into its subunit polypeptide chains by dialysis against a buffer containing 10 mM EDTA at pH 8.0 or above. These subunits have been shown to be homogeneous in molecular weight by sedimentation equilibrium (Miller & van Holde, 1982). An additional criterion for homogeneity is shown by sedimentation velocity studies. We have calculated the integral distribution of sedimentation coefficients by the method of van Holde and Weischet (1978). This technique effectively removes the boundary spreading caused by diffusion during a sedimentation velocity experiment. The results show that the *Octopus* hemocyanin subunit is highly homogeneous (data not shown). Thus, by these available physical criteria, the *O. dofleini* hemocyanin subunits are homogeneous with respect to both molecular weight and hydrodynamic behavior. This does not, of course, rule out heterogeneity in polypeptide chain composition and/or sequence. To assess homogeneity at this level, more discriminating techniques are required.

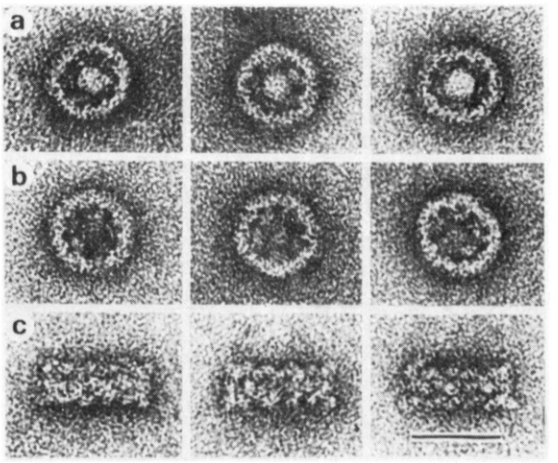


FIGURE 1: Electron micrographs of the whole 51S molecule of *O. dofleini* hemocyanin. The length of the bar is 30 nm. (a and b) Top views of the molecule; (c) side view.

The presence of a single N-terminal amino acid sequence is often given as evidence for subunit homogeneity. The first attempt at amino acid sequencing of the N-terminus of the entire subunit was made more difficult by the large size of the molecule. When enough protein was applied to the filter to provide a reasonable number of ends, the thickness of the protein layer could obscure some ends, leading to a high background level that intensified as the run continued and fresh ends were revealed. An attempt was made to improve the resolution by first purifying the hemocyanin as subunits by gel filtration on Sephadex G-150 in the presence of 10 mM EDTA. The resultant sample was reapplied to the sequencer, and the resolution was greatly improved. The results of this sequencing show a single N-terminus for the purified subunit. This suggests but does not yet prove conclusively that the subunits have the same amino acid sequence throughout.

The homogeneity of the polypeptide chains can be tested in one additional way. Figure 2 demonstrates homogeneity of the *Octopus* subunits. In this experiment, crossed immunoelectrophoresis of the subunit (Figure 2a) at pH 8.6 in the absence of calcium shows a single peak corresponding to the subunits. At pH 7.5 with 2.5 mM calcium lactate, there is a reduction in size of this peak and the corresponding appearance of a second more anodic peak (Figure 2b) representing the whole molecule. At pH 7.5 with 5 mM calcium lactate, the hemocyanin is nearly all in the form of whole molecules (Figure 2c). These results are supported by electron micrographs (Figure 2c–e) of small samples of material extracted from the gels themselves under the same experimental conditions. Thus by immunological criteria both the subunit and the whole molecule are homogeneous proteins.

Electron Microscopy of the Polypeptide Chains

Figure 3 shows a gallery of selected views of polypeptide chains resulting from the dissociation of the native molecule

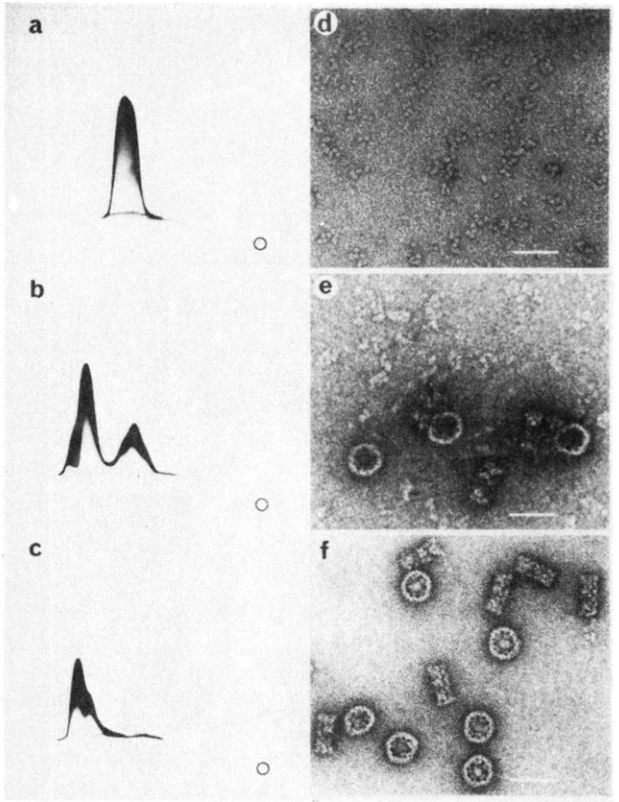


FIGURE 2: Demonstration of the homogeneity of the constituent polypeptide chains of *O. dofleini* hemocyanin by crossed immunoelectrophoresis and electron microscopy. (a–c) Crossed immunoelectrophoresis of whole *O. dofleini* hemocyanin. The first dimension occurred (a) at pH 8.6 in barbital buffer, (b) at pH 7.5 in 50 mM Tris-HCl buffer with 2.5 mM calcium lactate, and (c) at pH 7.5 in 50 mM Tris-HCl buffer with 5 mM calcium lactate; (d–f) electron micrographs of hemocyanin eluted from gels separated as in (a–c), respectively. The anode is on the left and the cathode on the right in each picture.

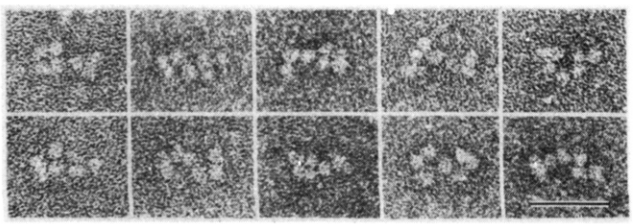


FIGURE 3: Selected electron microscopy views of multidomain polypeptide chains of *O. dofleini* hemocyanin. The length of the bar is 30 nm.

by dialysis against a 0.1 *I* Tris-HCl buffer, pH 8.0, containing 10 mM EDTA. Seven globular domains are visible on the micrographs. It is interesting that a number of these views show the molecule arranged into a block of four domains, plus a more loosely arranged group of three. There is no structural explanation for this observation at the present time.

Preparation of Six Domain Specific Antisera

Identification and Nomenclature of the Domains. When it appeared probable that the unique (major) polypeptide chain of *O. dofleini* hemocyanin was composed of seven ~50-kDa domains, the problem of their nomenclature arose. Since at the beginning there was absolutely no way to identify various domains within the polypeptide chain, the characterization was done according to their immunochemical properties, a method that proved to be successful with the subunits of *Androctonus australis* (Lamy et al., 1979a) and *Limulus polyphemus* hemocyanins (Lamy et al., 1979b).

Initially the whole subunit was treated by trypsin, a method used with success by Gielens et al. (1980, 1981) on *Helix* hemocyanin. The lysate, submitted to an immunoelectrophoresis vs. the antiserum specific for the whole hemocyanin, produced five major peaks. These five peaks were further shown to correspond to four one-domain fragments and one two-domain fragment by thin-layer gel filtration and SDS gel electrophoresis. Arbitrarily, the names *Od* 1, 2, 3, and 4 were assigned to the four one-domain fragments producing the four major peaks in immunoelectrophoresis (*Od* = *Octopus dofleini*). The two-domain major fragment was later shown to cross-react with domain *Od* 2 and to contain another domain termed *Od* 5. Therefore, it was concluded that the two-domain major fragment produced by trypsin was composed of domain *Od* 2 and domain *Od* 5. Domains *Od* 6 and 7 were at first missed because they are very labile in the presence of trypsin. Domain *Od* 6 was identified in subtilisin digests, and the existence of domain *Od* 7 was deduced from the crossed line immunoelectrophoresis pattern of a three-domain fragment (see Description of a Typical Experiment of Proteolytic Digestion).

Domain-Specific Antisera Permit Identification of Domains in Proteolytic Fragments. Six antisera specific for domains *Od* 1, 2, 3, 4, 5, and 6 were prepared. The antigens given to the rabbits were purified by ion-exchange fast protein liquid chromatography (FPLC) on Mono Q columns and by gel permeation on Superose 12. The last step of purification was the line immunoelectrophoresis described under Antisera. All the six antisera were domain-specific in immunoprecipitation. For example, Figure 4A shows the precipitation pattern against an anti-dissociated hemocyanin antiserum of a subtilisin digest of *O. dofleini* hemocyanin. The pattern is complex and difficult to interpret. However, the six domain specific antisera permit identification of various domains in the digest fragments. Thus, a comparison of panels A and B of Figure 4 shows that domain *Od* 6 produces the highest peak on the cathodic side of Figure 4A. Similarly, a comparison of panels A and F of Figure 4 shows that domain *Od* 2 produces the most anodic peak of Figure 4A. The anti-*Od* 5 antiserum precipitates three antigens with different relative mobilities in the first dimension (electrophoresis at pH 8.6). Of these three antigens, the most anodic one is also precipitated by the anti-*Od* 2 antiserum. Obviously it is an *Od*(2,5) fragment. The two anodic fused peaks of Figure 4C both correspond to one-domain fragments. The best explanation is that the subunit is attacked by subtilisin in different positions producing two antigenically identical domains *Od* 5 with different electrical charges, perhaps because of interdomain loops with different length. The anti-*Od* 4 antiserum precipitates a single fragment of the subtilisin digest. One first would conclude that this fragment is domain *Od* 4. In fact, it is a two-domain fragment composed of one domain *Od* 4 and one domain *Od* 7. This was established after purification of the fragment by the crossed line electrophoresis technique described under Description of a Typical Experiment of Proteolytic Digestion (Figure 7).

An anti-domain *Od* 7 antiserum could not be prepared by the above-described method for two reasons. First, domain *Od* 7 appears very labile by most proteases, and second, the interdomain link between domain *Od* 7 and domain *Od* 4 seems much more resistant to proteolysis than domain *Od* 7 itself. Therefore, there was not enough pure domain *Od* 7 to prepare the specific antiserum.

An interesting aspect of these antisera was their high specificity. Actually, as shown in Figure 4, each antiserum

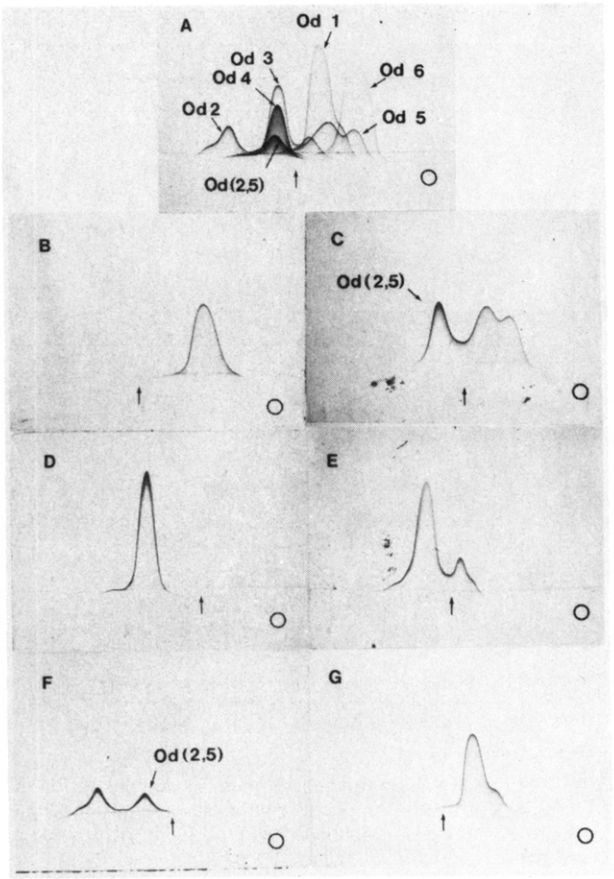


FIGURE 4: Demonstration of the specificity of antisera prepared against the purified domains of *O. dofleini* hemocyanin. (A) Precipitation pattern against an anti-dissociated hemocyanin antiserum of a subtilisin digest of *Od* hemocyanin; (B–G) precipitation pattern against antisera specific for purified domains of the same subtilisin digest of *Od* hemocyanin: (B) anti-*Od* 6, (C) anti-*Od* 5, (D) anti-*Od* 4, (E) anti-*Od* 3, (F) anti-*Od* 2, and (G) anti-*Od* 1. The anode is on the left and the cathode on the right in each picture. The arrows point to identical electrophoretic mobilities.

precipitates a single type of domain. No strong cross-reactivity occurs as with the subunits of arthropodan hemocyanins (Lamy et al., 1983). This suggests that despite their similarity in molecular weight and function the domains may have very different amino acid sequences.

Sequence of the Domains in the Polypeptide Chain

Once the domain names were assigned, the determination of their sequence in the polypeptide chain required the preparation of various fragments to determine their domain composition and to look for overlaps, in a way similar to the alignment of the fragments in an amino acid sequence determination. In the case of *O. dofleini* hemocyanin, the polypeptide chain was cleaved by proteolytic digestions, the overlaps were identified by immunochemistry, and the N-terminal and C-terminal orientations of the fragments were determined by N-terminal amino acid sequences.

Description of a Typical Experiment of Proteolytic Digestion. As an example this section describes the hydrolysis of *O. dofleini* hemocyanin subunit by *elastase*. As a first step, a convenient enzyme/substrate ratio is determined, and the proteolysis is studied as a function of time. In this preliminary experiment the test used to estimate the extent of the hydrolysis is a crossed immunoelectrophoresis against an antiserum specific for dissociated hemocyanin. Figure 5A is a control before the addition of the enzyme. It shows a single homogeneous peak. After a 30-min incubation (Figure 5B) at 37 °C with an enzyme/substrate ratio of 6×10^{-4} (w/w), two

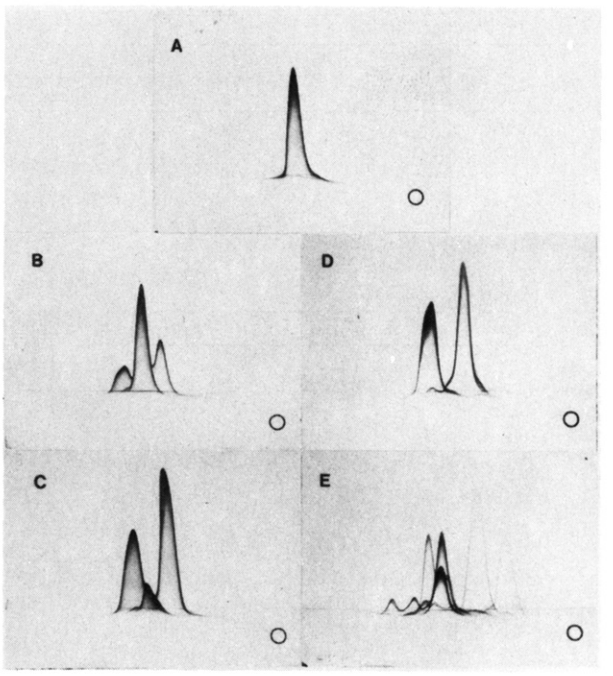


FIGURE 5: Kinetics of digestion of *O. dofleini* hemocyanin by elastase: (A) undigested hemocyanin, (B) after 30-min incubation at 37 °C with enzyme/substrate ratio of 6×10^{-4} (w/w), (C) after 2-h incubation, (D) after 5-h incubation, and (E) after 20-h incubation. The anode is on the left and the cathode on the right in each picture.

shoulders appear on both sides of the peak. These minor peaks are obviously antigenically deficient compared to the main peak. Actually, the left (anodic) foot of the right peak and the right (cathodic) foot of the left peak fuse with the right and left feet of the central main peak, respectively. This pattern is still more visible in Figure 5C after a 2-h incubation, except that the right and left peaks are now higher than the central one. After 5 h (Figure 5D) the central peak has completely disappeared, and after 20 h (Figure 5E) the dissociation pattern has become complex. The interpretation of Figure 5A–D is simple, the subunit has been split into two fragments, likely a three-domain fragment and a four-domain fragment. One of these fragments is more negatively charged than the whole subunit and the other more positively charged at the pH of the electrophoresis (8.6). The domain composition of the fragments can be determined with the domain-specific antisera. For example, Figure 6 shows how the domain composition was determined for the fragments obtained after a 2-h incubation with elastase at an enzyme/substrate ratio of 6×10^{-4} (w/w) (conditions of Figure 5C). Figure 6A shows the result obtained with the anti-*Od* 1 antiserum. Obviously the antiserum only precipitates two of the three peaks present in Figure 5C. A comparison of the electrophoretic mobilities of Figures 5C and 6A shows that these peaks are the cathodic and central peak of Figure 5C. Therefore, domain *Od* 1 is present in the fragment migrating on the right (cathodic) side of the residual subunit and of course in the residual subunit itself. The same method using antisera directed against domains *Od* 2, 3, 4, 5, and 6 indicates that domains *Od* 2, 5, and 6 are also present in the cathodic peak while domains *Od* 3 and 4 are present in the anodic peak. The present lack of an anti-*Od* 7 antiserum does not permit the determination of where domain *Od* 7 is located. Fortunately, another immunological method, the crossed line immunoelectrophoresis, allowed identification of domain *Od* 7 in the anodic fragment.

As stated above, domain *Od* 7 is very labile, and this fact has made the preparation of an anti-*Od* 7 antiserum impossible up to now. However, advantage was taken of this exceptional

Table II: Products of Limited Proteolysis of *O. dofleini* Hemocyanin Subunits

enzyme	fragments			
	6 domains	4 domains	3 domains	2 domains
subtilisin		(1,2,5,6), (3,4,6,7)	(1,2,5), (3,4,7)	(2,5), (4,7)
serine protease from *S. aureus*	(2,3,4,5,6,7)		(2,5,6), (3,4,7)	(2,5)
α-chymotrypsin		(1,2,5,6)	(2,5,6), (3,4,7)	
trypsin	(2,3,4,5,6,7)[a]	(3,4,6,7)	(1,2,5), (3,4,7)	(2,5), (3,4)[b]
elastase		(1,2,5,6)	(3,4,7)	
plasmin			(1,2,5), (3,4,7)	(2,5)
papain		(1,2,5,6)	(1,2,5), (3,4,7)	(2,5), (4,7)
thrombin[c]				

[a] Low yield. [b] Digest of isolated (3,4,7). [c] No fragment was obtained.

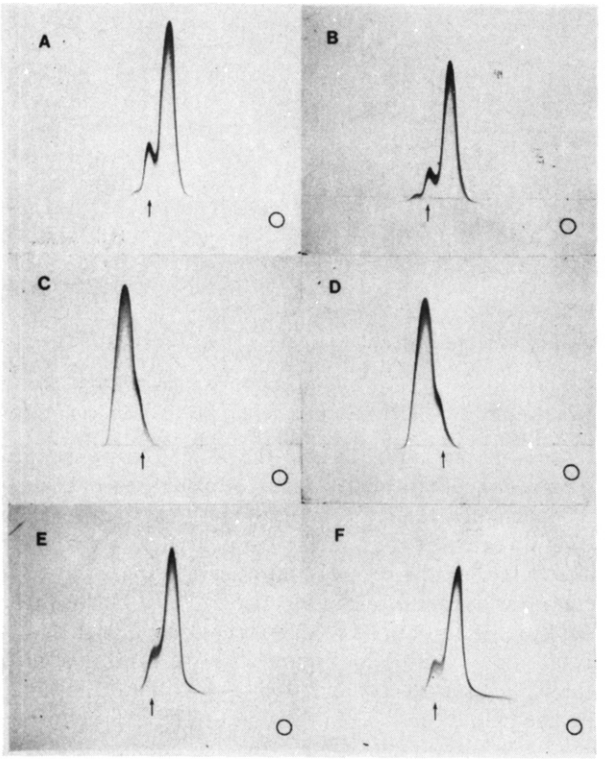


FIGURE 6: Domain composition of fragments of *O. dofleini* hemocyanin obtained following elastase digestion under conditions in Figure 5C [2 h at 37 °C with enzyme/substrate ratio of 6×10^{-4} (w/w)]. (A–F) Precipitation patterns against antisera for purified domains. The arrows point to the center peak of (B). (A) Anti-*Od* 1; (B) anti-*Od* 2; (C) anti-*Od* 3; (D) anti-*Od* 4; (E) anti-*Od* 5; (F) anti-*Od* 6. The anode is on the left and the cathode on the right in each picture.

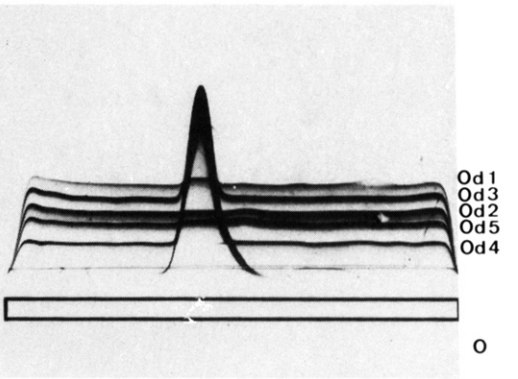


FIGURE 7: Crossed line immunoelectrophoresis demonstration that the three-domain anodic fragment from elastase digestion (Figure 6) contains domains *Od* 3, 4, and 7 (see text). The anode is on the left and the cathode on the right.

lability, especially in the presence of trypsin, to identify domain *Od* 7 in the proteolysis fragments. Actually, even under conditions of limited proteolysis [enzyme/substrate ratio = 2×10^{-2} (w/w), 15-h incubation], domains *Od* 7 and *Od* 6 are destroyed while all the other domains are still present. Figure 8 shows a crossed line immunoelectrophoresis vs. an anti-dissociated hemocyanin antiserum, which demonstrates the existence of domain *Od* 7. This immunoelectrophoresis contains two different antigens. In the deposit hole is the three-domain fragment purified by chromatography corresponding to the anodic peak of Figure 5B–D. This antigen produces the precipitation peak. In the trough a trypsin digest of the subunit devoid of domain *Od* 7 was deposited. This antigen produces the series of precipitate lines. The domain composition of the various lines was established with the various domain-specific antisera. Notice that trypsin also destroys domain *Od* 6, which is not present in the lines. From top to bottom, the various lines contain domains *Od* 1, 3, 2, 5, and 4, respectively. It is clear that three lines cross the peak. These lines correspond to domains *Od* 1, 2, and 5, meaning that these domains are *not* present in the antigen deposited in the hole. Conversely, the lines corresponding to domains *Od* 3 and 4 fuse with the peak, demonstrating that these domains are present in the antigen. All these aspects are in perfect agreement with the data of Figures 5 and 6. The crucial point is that the peak has two feet under the line corresponding to domain *Od* 4. These feet can only correspond to domain *Od* 6 and/or 7. However, we know from Figure 6F that domain *Od* 6 is not present in the anodic peak. Therefore, the foot demonstrates the presence in the peak of domain *Od* 7.

Together with the immunoelectrophoreses of Figures 5 and 6, the crossed line immunoelectrophoresis of Figure 7 demonstrates that elastase cuts the subunits into the three-domain fragment *Od* (3,4,7) and the four-domain fragment *Od* (1,2,5,6).

Other experiments of this type using different proteases have produced different proteolysis patterns for which the results are summarized in Table II. In some cases the fragments had to be purified by FPLC or by semipreparative polyacrylamide gel electrophoresis. This was specifically the case for the demonstration of the existence of the *Od* (4,7) and *Od* (3,4) fragments, two very important fragments that could be observed only after the purification of the *Od* (3,4,7) three-domain fragment. Indeed, the overlap between fragments *Od* (3,4) and *Od* (4,7) demonstrates that domain *Od* 4 is located between domains *Od* 3 and *Od* 7.

Table II also shows that, with the exception of trypsin and serine protease from *Staphylococcus aureus*, which can produce a six-domain fragment plus domain *Od* 1, all the proteases lyse the subunit into a four-domain and three-domain fragment. The four-domain fragment is in most cases *Od* (1,2,5,6), but trypsin produces an *Od* (3,4,6,7) fragment. Increasing the enzyme/substrate ratio or the incubation time usually results in a further degradation of the four-domain

Table III: N-Terminal Amino Acid Sequences of *O. dofleini* Hemocyanin Subunit and Domains

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Od 1	T	V	G	D	A	I	I	R	K	N	V	N	S	L	T	P	S	D	I	K	E	L	R	D	A	M	A
whole subunit[a]					N	L	I		K	N	V	(D)	A	L	(D)	(E)											
Od (4,7)[a]					N	L	I	(S)	K	D^b	V	D	A	L													
Od 2	X	P	P	S	N	E	D	A	D	I	D	T	P	L	N	H	I	R	(R)	N							
Od (2,5)	(S)	(E)	(E)	(G)	(N)	(Q E)	(Y)	(L)	(V)	(M)	(A)																
	(A G)	X	(S)	(A)	(P)	(S)	(L)	(L)	(G)	(R)	(K)																

[a] There appears to be some sequence homology between these two sequences and *Od* 1. They have been aligned to emphasize this homology. The sequences are not complete because they contain many glycosylated residues, which can hydrolyze spontaneously in the sequencer, causing modification and/or destruction of some residues. [b] Hydrolysis of Asn to Asp at this position could have resulted during the rather extensive manipulation of the two-domain fragment *Od* (7,4) during purification.

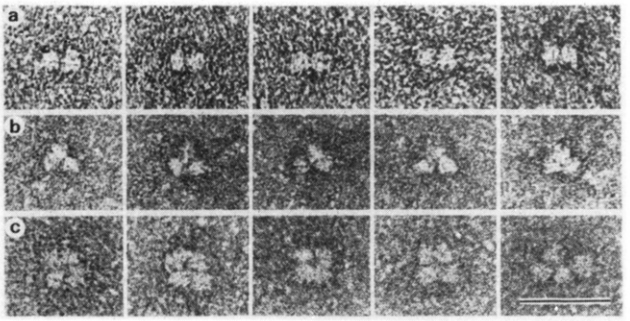


FIGURE 8: Selected gallery of electron micrographs fragments resulting from limited proteolysis of *O. dofleini* hemocyanin. The length of the bar is 30 nm. (A) Two-domain fragments; (B) three-domain fragments; (C) four-domain fragments.

fragment into a three-domain and a one-domain fragment. A detailed examination of the results suggests that there are two particularly labile interdomain areas located on both sides of domain *Od* 6. These labile areas are obviously responsible for the absence of five-domain fragments in the digests. The fragments resulting from digestions under various conditions are visible in micrographs of Figure 8. The four-domain fragments frequently appear as a block, corresponding perhaps to the block of four seen in micrographs of the whole subunit (Figure 3).

Amino Acid Sequencing. The enzymatic digestion method together with the immunological identification of the domains led to the identification of the three-domain fragment *Od* (3-4-7) and the four-domain fragment *Od* [1-(2,5)-6]. However, it was not clear which of these fragments was on the N-terminal side of the subunit nor which domain was on the N-terminal side of each fragment. To answer this question, several partial amino acid sequences of proteolytic fragments and of the whole subunit were determined. They are shown in Table III. The first problem was to find out which domain is located on the N-terminal side of the subunit: 1, 6, 3, or 7. We were greatly helped in the resolution of this problem by the results that Dr. Takagi of Tohuko University presented at the meeting on "Invertebrate Oxygen Carriers" held in August 1985 at Tutzing. Dr. Takagi presented the complete sequence of the C-terminal domain of *Paroctopus dofleini dofleini* hemocyanin. This domain had been prepared by trypsinolysis under conditions very close to those that produced domain *Od* 1 in *Octopus dofleini martini* hemocyanin. Immediately after the meeting, we isolated enough pure domain *Od* 1 to determine the N-terminal sequence of amino acids and found that this sequence was identical, with the exception of the first five residues, with that of Takagi's C-terminal domain. Later, it appeared that the apparent difference was not real (see discussion) so that domain *Od* 1 is clearly the C-terminal domain. This location implies that domain *Od* 6 cannot be the N-terminal domain of the subunit.

The second problem was the orientation of fragment *Od* (3-4-7) with respect to the N-terminal end of the subunit. The three-domain fragment could not be used for this purpose because it necessarily contains the N-terminal end. Domain *Od* 3 was purified in an immunologically pure form, but this material produced more than one residue at each degradation step. Probably the folded domain contained a cut so that upon unfolding two polypeptide chains were released. The sequence of neither fragment was similar to the subunit N-terminal sequence however. As domain *Od* 7 is very labile, it could not be purified. However, the two-domain fragment *Od* (4-7) could be isolated by a combination of fast protein liquid chromatography on Mono Q and gel permeation on Superose 12. This fragment had the same N-terminal sequence of amino acids as the whole subunit, with a single exception at position 10. The two-domain fragment contained Asp instead of Asn. This may be due to hydrolysis during purification or sequencing of the amide function of Asn, a possibility which may be related to a high concentration of carbohydrate. Thus *Od* 7 must be the N-terminal domain of the subunit. Heavy glycosylation of many residues in the N-terminal sequences of *Od* 7 and the subunit made these sequences difficult to obtain because they tended to hydrolyze spontaneously in the sequencer, greatly reducing resolution. However, this fact tends to support further the possibility that *Od* 7 is the N-terminal domain since similar effects of glycosylation were not observed in other domain sequences.

These data give the complete domain sequence with the exception of domains *Od* 2 and 5. Domain *Od* 2 and the two-domain fragment *Od* (2,5) were then purified in an immunologically pure state, and their amino acid sequences were determined. Unfortunately, although the sequence of domain *Od* 2 was unique, containing a very distinctive pair of prolines close to the N-terminus, the two-domain fragment produced at least two different sequences, none of them being similar to that of domain *Od* 2. The complete sequence of domains in *O. dofleini* hemocyanin remains as follows: N-7-4-3-6-(5,2)-1-C.

Discussion

Homogeneity of the Polypeptide Chain. The *O. dofleini* subunit is almost certainly a homogeneous polypeptide chain. Physical methods can indicate homogeneity of size but not of polypeptide chain composition. Although sedimentation equilibrium experiments (Miller & van Holde, 1982) and hydrodynamic behavior indicate homogeneity, there remained

the necessity to show that there is only one polypeptide chain. SDS gels can sometimes show heterogeneity among hemocyanins. In *Octopus* there is only one major band at approximately 350 kDa (Miller & van Holde, 1982). There is a minute amount of material from purified hemocyanin that runs on SDS gels as several very small faster bands. These bands are undetectable except on gels so heavily overloaded that the main protein band is unmeasurably nonlinear. In no case does any single minor band exceed 1% of the total protein. The precise nature of these proteins has not been conclusively demonstrated. Their regular spacing and positions on the SDS gel strongly suggest fragments of the native molecule lacking one or more domains. The domains themselves, excepting perhaps domains *Od* 6 and *Od* 7, are quite resistant to proteases, but the linker regions between domains are much more susceptible to protease digestion. Herskovits and Villanueva (1986) have observed similar small bands and concluded from experimental evidence that they were the result of proteolytic degradation. They found that the molecular weight gradually decreased when samples dissociated in 6 M guanidinium chloride were allowed to stand for several days. Dissociation can render the subunits more susceptible to the action of proteases. Whatever their precise nature, the small polypeptides were released from the native hemocyanin upon dissociation in EDTA and complicated the determination of the N-terminal sequence by increasing the background. If the subunits were repurified at this point, the resultant sample showed a single N-terminal sequence. It is quite likely that the microheterogeneity of the sample as observed in the SDS gels was partially responsible for the background problems. If the small bands result from proteins distinct from the subunit molecule but nevertheless part of the whole hemocyanin, they could potentially complicate the production of antibodies, but there is no evidence for more than one protein being released upon dissociation (Figure 2). There is a single immunoprecipitation peak against anti-dissociated hemocyanin antiserum, which indicates a single polypeptide chain. Gielens et al. (1986) have obtained similar results in *O. vulgaris* hemocyanin, which has been first purified to remove agglutinins and other small contaminants.

The situation for *Octopus* hemocyanin contrasts sharply with that for other molluscs. *Helix* (Lontie, 1983), *Murex* (Brouwer et al., 1978), and *Nautilus* (Bonaventura et al., 1981) all show heterogeneity of subunits, being composed of several distinct subunit types. Although there may still be some unresolved questions concerning the precise nature of the microheterogeneity of *Octopus* hemocyanin, it probably results from smaller fragments of hemocyanin subunit rather than from the existence of several different subunit types. There remains, of course, the possibility that a quite recent gene duplication may have generated small differences in amino acid sequence between the subunits, which were not detected with the available methods. The only way to demonstrate the absolute purity of the sample is with a complete amino acid sequence.

Number of Domains. The electron microscope shows (Figure 2) and immunological methods confirm the fact that *O. dofleini* hemocyanin contains seven oxygen binding domains (see Homogeneity of the Polypeptide Chain). This observation is in agreement with the results of Gielens et al. (1986) for *O. vulgaris*. It is also in agreement with the minimum molecular weight for molluscs of 50 000 per functional unit, on the basis of copper content, as well as the observation made by Miller and van Holde (1982) that the molecular weight of this hemocyanin (3.5×10^6) is less than that of other molluscan hemocyanins that contain eight domains (i.e., *Helix*, 4.5×10^6; *Loligo*, 3.85×10^6; *Sepia*, 3.8×10^6). The electron microscope cannot distinguish between a seven-domain subunit and a subunit with six functional domains plus a seventh, modified domain. Furthermore, the immunological tests used to demonstrate the existence of domain *Od* 7 were indirect, because of the difficulties encountered in purifying it. Nevertheless, the accumulated weight of evidence strongly suggests that it does exist and that the seven-domain structure for *Octopus* hemocyanin is a valid one. It would seem that the individual domain size in cephalopods must be somewhat smaller than that in gastropods to account for the smaller molecular weight of the eight-domain subunit in *Sepia* and *Loligo*. It would be interesting to know whether *Nautilus* hemocyanin at 3.5×10^6 daltons (Bonaventura et al., 1981), but with a somewhat higher sedimentation coefficient (57.9 S vs. 51 S for *Octopus*), has seven domains or eight.

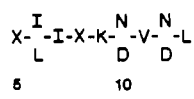
The question of the subunit size in cephalopods at the present time seems much better resolved than it is for gastropods. The domain size, on the basis of SDS gels of the products of limited proteolysis (Gielens et al., 1986), is close to 50 000 daltons, which is in good agreement with the copper content of 50 000 daltons per pair of copper atoms. The seven-domain subunit of *Octopus* hemocyanin is 7 times this, or 3.5×10^5 daltons. Ten subunits form a whole molecule of 3.5×10^6 daltons. In *Loligo* and *Sepia* the eight-domain subunits are somewhat larger, $\sim 3.8 \times 10^5$ daltons, forming whole molecules also proportionately larger. The situation for gastropods remains confusing. The domain size does seem somewhat larger (Lontie, 1983) at 55 000 daltons, but the measurements of copper content for gastropods is, like that of cephalopods, about 50 000 daltons. If one multiplies eight domains by 55 000 daltons, one obtains a subunit molecular weight of 4.4×10^5, which agrees well with a half-molecule size of 4.5×10^6 daltons but not with the published values for subunit size (van Holde & Miller, 1982), which are all closer to 3.5×10^5 daltons. Salavato and Zatta (1983) have found quite different values for the subunit molecular weight of molluscan hemocyanin. It may be necessary to have a complete amino acid sequence for a gastropod hemocyanin before this issue is fully resolved.

Orientation of the Polypeptide Chain. The results of the amino acid sequencing show that the N-terminal sequence of the C-terminal domain for *P. dofleini dofleini* (Takagi, 1986), which was prepared by limited trypsinolysis similarly to domain *Od* 1, is indeed the same as that for domain *Od* 1 of *O. dofleini martini*. At first there was some question concerning these sequences because the first five amino acids differed. A reexamination of this sequence (T. Takagi, personal communication) has demonstrated that the two sequences now agree totally up to residue 27. Although there is some uncertainty regarding the relationship between the two octopus species, this result indicates that the sequences have been determined for very closely related animals. Serine protease from *S. aureus* (Table II) yields domain *Od* 1 and the six-domain fragment *Od* (2,5,6,3,4,7). Gielens et al. (1986) obtain similar results for *O. vulgaris*. The fragment they term "Spd" is probably the same as *Od* 1. They too find that it is the C-terminal domain.

We have found that the N-terminal sequences of the whole subunit and fragment *Od* (7-4) are the same. They both begin with the amino acid asparagine. Gielens et al. (1986) report the same result for *O. vulgaris*. Takagi, on the other hand, finds leucine in this position. He gives no additional sequence information for the N-terminus, however. Without extensive

comparison of amino acid sequences, taxonomy, etc., there is at present no good explanation for this discrepancy.

Relatedness of the Domains and the Nature of Domain Od 7. The N-terminal amino acid sequences of the whole subunit and fragment *Od* (7-4) are probably the same, if one assumes a chemical modification by hydrolysis of Asn to Asp (see Amino Acid Sequencing). The heavy glycosylation observed in the amino acid sequencer of fragment *Od* (7-4) (the N-terminal fragment of the subunit) is interesting in itself, for it contrasts rather sharply with the negligible glycosylation of other residues. This is especially important because the sequence of domain *Od* 7 has some clear homologies to that of domain *Od* 1. The sequence (from position 5 through position 13) is



The homologies are most notable at positions 7, 9, 11, and 13 (see Table II) and to a lesser extent at positions 6, 10, and 12, where amino acid substitutions of leucine for isoleucine and aspartate for asparagine are among the commonest alterations of sequence. Perhaps the glycosylation of the N-terminal residue has some significance for the structure of the molecule. Also, since there is little or no immunological cross-reactivity between the domains, it is possible that heavy glycosylation of domain *Od* 7 diminishes its apparent immunological relatedness to domain *Od* 1, which has a similar N-terminal sequence. Of course, nothing is known of the rest of the sequence of domain *Od* 7. This observation of limited sequence homology between the N-terminal and C-terminal domains of *Octopus* hemocyanin is possibly significant.

The precise nature of *Od* 7 is still somewhat puzzling. Usually domains are much more resistant to proteases than the linker regions between them. This makes sense when one considers that proteolytic enzymes work more easily on unfolded regions of a polypeptide chain than on folded regions. In this case however, *Od* 7 is extremely labile. Perhaps its N-terminal position is important for the assembly or structure of the molecule in some way. It could be more heavily glycosylated or more loosely folded than the other domains for this purpose. Unfortunately, these modifications also make purification of *Od* 7 extremely difficult.

On the micrographs we see some difference between the two circular faces of *Octopus* hemocyanin, suggesting a slight asymmetry of structure although it is not clear what causes this difference. There is no clearly visible polarity of the side view caused by a collar structure as seen in gastropod hemocyanins. Gielens et al. (1979, 1980) have isolated the two-domain collar fragment of *Helix* β_c-hemocyanin. It occurs at the C-terminal end of the subunit, whereas domain *Od* 7, whether normal or not, occurs at the N-terminal end of the *Octopus* subunit. Thus the structural homologies between the collar domains of gastropods and *Octopus Od* 7 are not strong, at least from the point of view of the sequence of domains within the subunit.

Remaining Uncertainties Regarding the Domain Sequence. Although the sequence of domains in *Octopus dofleini* hemocyanin has been nearly resolved, there remains an uncertainty regarding the positions of domains *Od* 5 and *Od* 2. Domain *Od* 2 and the dimeric fragment (2,5) were both produced by the same tryptic digestion. Domain *Od* 2 is a single polypeptide with a unique N-terminal sequence, containing a highly unusual pair of prolines. Fragment (2,5) is heterogeneous in amino acid sequencing, containing at least two polypeptides, although it is immunologically pure. Nothing is known of the sizes of these two fragments. It appears that either domain *Od* 5 must be the more fragile of the two-domains or the linker region contains several possible tryptic digestion sites and that domain *Od* 2 must therefore be the limit tryptic digestion product. One could also argue that because the sequence of neither part of fragment (2,5) has any similarity to domain *Od* 2 with its unique prolines, the dimeric fragment must begin with something besides domain *Od* 2. If one accepts these arguments, the sequence must be 5-2. This is unfortunately somewhat inconclusive, negative evidence, but if it is true, the sequence of domains must be N-7-4-3-6-5-2-1-C.

ACKNOWLEDGMENTS

We thank Dr. A. W. Martin of the University of Washington for his assistance in obtaining *Octopus* and Solange Compin for technical assistance.

REFERENCES

Axelsen, N. H., Kroll, J., & Weeke, B. (1973) *Scand. J. Immunol.*, Suppl. 1.

Bonaventura, C., Bonaventura, J., Miller, K., & van Holde, K. E. (1981) *Arch. Biochem. Biophys. 211*, 589–598.

Brouwer, M., Ryan, M., Bonaventura, J., & Bonaventura, C. (1978) *Biochemistry 17*, 2810–2815.

Brouwer, M., Wolters, M., & Van Bruggen, E. F. J. (1979) *Arch. Biochem. Biophys. 193*, 487–495.

Davis, B. (1964) *Ann. N.Y. Acad. Sci. 121*, 404–427.

De Sadeleer, J., Gielens, C., Préaux, G., & Lontie, R. (1983) *Life Chem. Rep., Suppl. Ser. 1*, 133–134.

Ghiretti-Magaldi, A., Salvato, B., Tognon, G., Mammi, M., & Zanotti, G. (1981) in *Invertebrate Oxygen Binding Proteins* (Lamy, J., & Lamy, J., Eds.) pp 393–404, Dekker, New York.

Gielens, C., Préaux, G., & Lontie, R. (1977) in *Structure and Function of Haemocyanin* (Bannister, J. V., Ed.) pp 85–94, Springer-Verlag, Berlin.

Gielens, C., Lanckriet, M., & Lontie, R. (1979) in *Metalloproteins* (Weser, U., Ed.) Thieme, Stuttgart.

Gielens, C., Verschueren, L. J., Préaux, G., & Lontie, R. (1980) *Eur. J. Biochem. 103*, 463–470.

Gielens, C., Verschueren, L. J., Préaux, G., & Lontie, R. (1981) in *Invertebrate Oxygen Binding Proteins* (Lamy, J., & Lamy, J., Eds.) pp 295–303, Dekker, New York.

Gielens, C., Bosman, F., Préaux, G., & Lontie, R. (1983) *Life Chem. Rep., Suppl. Ser. 1*, 121–124.

Gielens, C., Benoy, C., Préaux, G., & Lontie, R. (1986) in *Invertebrate Oxygen Carriers* (Linzen, B., Ed.) pp 223–226, Springer-Verlag, Berlin and Heidelberg.

Herskovits, T., & Villanueva, G. (1986) *Biochemistry 25*, 931–939.

Laemmli, V. K. (1970) *Nature (London) 227*, 680–685.

Lamy, J., Lamy, J., & Weill, J. (1979a) *Arch. Biochem. Biophys. 193*, 140–149.

Lamy, J., Lamy, J., Weill, J., Bonaventura, J., Bonaventura, C., & Brenowitz, M. (1979b) *Arch. Biochem. Biophys. 196*, 324–339.

Lamy, J., Compin, S., & Lamy, J. (1983) *Arch. Biochem. Biophys. 223*, 584–603.

Lamy, J., Lamy, J. N., Leclerc, M., Compin, S., Miller, K. I., & van Holde, K. E. (1986) in *Invertebrate Oxygen Carriers* (Linzen, B., Ed.) pp 231–234, Springer-Verlag, Berlin and Heidelberg.

Lontie, R. (1983) *Life Chem. Rep., Suppl. Ser. 1*, 109–120.

Miller, K. I., & van Holde, K. E. (1982) *Comp. Biochem. Physiol., B: Comp. Biochem. 17B*, 1013–1018.

Pickford, G. E. (1964) *Bull. Bingham Oceanogr. Collect. 19*, 5-67.

Salvato, B., & Zatta, P. (1983) *Life Chem. Rep., Suppl. Ser. 1*, 139-140.

Takagi, T. (1986) in *Invertebrate Oxygen Carriers* (Linzen, B., Ed.) pp 259-262, Springer-Verlag, Berlin and Heidelberg.

van Bruggen, E. F. J., Schutter, W. G., van Breemen, J. F. L., Bijholt, M. M. C., & Wichertjes, T. (1981a) in *Electron Microscopy of Proteins* (Harris, J., Ed.) pp 1-38, Academic, London.

van Bruggen, E. F. J., Schutter, W. G., Wichertjes, T., & Keegstra, W. (1981b) in *Invertebrate Oxygen Binding Proteins* (Lamy, J., & Lamy, J., Eds.) pp 405-414, Dekker, New York.

van Holde, K. E., & Weischet, W. (1978) *Biopolymers 17*, 1387-1403.

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

van Holde, K. E., & Miller, K. I. (1985) *Biochemistry 24*, 4577-4582.

Weeke, B. (1973) *Scand. J. Immunol., Suppl. 1*, 1-35.

Wichertjes, T., Gielens, C., Schutter, W. G., Préaux, G., Lontie, R., & van Bruggen, E. F. J. (1986a) in *Invertebrate Oxygen Carriers* (Linzen, B., Ed.) pp 227-230, Springer-Verlag, Berlin and Heidelberg.

Wichertjes, T., Gielens, C., Schutter, W. G., Préaux, G., Lontie, R., & van Bruggen, E. F. J. (1986b) *Biochim. Biophys. Acta 872*, 183-194.

Use of Consensus Oligonucleotides for Detecting and Isolating Nucleic Acids Encoding Calcium Binding Domains of the Troponin C Superfamily†

Susan H. Hardin,‡ Matthew J. Keast,‡ Paul E. Hardin,§ and William H. Klein*

Department of Biochemistry and Molecular Biology, M. D. Anderson Hospital, University of Texas System Cancer Center, Houston, Texas 77030

Received November 17, 1986; Revised Manuscript Received January 23, 1987

ABSTRACT: Proteins belonging to the troponin C superfamily (troponin C, calmodulin, myosin light chains, and parvalbumin) are involved in a wide variety of cellular activities mediated by calcium ions. Most of these proteins bind ionic calcium, and all have calcium binding domains that are conserved to some extent at the nucleic acid level. We made use of the conservation in the third calcium binding domain to synthesize two consensus sequence oligonucleotide probes, one 43 bases and the other 25 bases long. By using cDNA and genomic clones encoding calmodulin, troponin C, parvalbumin, and the sea urchin Spec proteins, we show that these probes hybridize with nucleic acid sequences representing calcium binding domains. In an RNA gel blot analysis of embryonic RNA from the sea urchin *Strongylocentrotus purpuratus*, we show that transcripts which have previously been shown to encode troponin C like proteins hybridize with the consensus sequence probes. Screening sea urchin cDNA and genomic libraries with the 43-base consensus oligonucleotide shows that the probe can be used to isolate cloned nucleic acids. Two such genomic clones from a *Lytechinus pictus* library were isolated and characterized. One clone encodes part of an *L. pictus* calmodulin gene, and the other encodes a member of the superfamily that has not been characterized previously. The consensus oligonucleotides should be valuable probes in the diagnosis and isolation of nucleic acids encoding proteins of the troponin C superfamily.

Troponin C, calmodulin, alkali and regulatory myosin light chains, and parvalbumin belong to a group of proteins termed the troponin C superfamily by M. O. Dayhoff (1978). The diagnostic structural feature of the superfamily is the presence, in each protein, of conserved calcium binding domains that form a distinct helical structure called an EF hand (Kretsinger, 1980). Generally, these proteins function as regulators of cellular processes mediated by calcium ion, and because of this, much attention has been focused on their structure and mechanism of action (Means & Conn, 1987). X-ray crystallography of rat testes calmodulin (Babu et al., 1985) and turkey skeletal muscle troponin C (Herzberg & James, 1985) has recently been performed and has provided detailed structural information. Both proteins consist of two globular lobes containing the calcium binding sites connected by an exposed α helix.

The genes encoding several members of the troponin C superfamily have been isolated, and their nucleotide sequences have been determined (Falkenthal et al., 1985; Putkey et al., 1983; Nabeshima et al., 1984; Robert et al., 1984; Hardin et al., 1985; Salvato et al., 1986). These sequences and those obtained from protein sequencing (Kretsinger, 1980) show that there is weak conservation of amino acid residues among the various proteins. However, the nucleotide sequences encoding amino acids important for calcium ion interaction and for EF-hand helix structure are more strongly conserved.

Given the conservation of the calcium binding sites, it seemed likely that a consensus probe could be developed to detect nucleic acid sequences encoding members of the troponin C superfamily. Such a probe could then be used for isolating novel genes encoding troponin C related proteins that have not yet been characterized. The troponin C superfamily appears to be a much larger group of proteins than previously

† This work was supported by NIH Grant HD 22619. W.H.K. is the recipient of NIH Research Career Development Award HD735.

* Author whom correspondence should be addressed.

‡ Present address: Program in Molecular, Cellular and Developmental Biology, Indiana University, Bloomington, IN 47405.

§ Present address: Program in Genetics, Indiana University, Bloomington, IN 47405.

0006-2960/87/0426-3518$01.50/0