Fragmentation of a mollusc haemocyanin with plasmin and immunological identification of the fragments

William J. GULLICK,* Erica J. H. HEAD† and Edward J. WOOD

Department of Biochemistry, University of Leeds, 9 Hyde Terrace, Leeds LS2 9LS, U.K.

(Received 19 September 1980/Accepted 12 December 1980)

Haemocyanin from the gastropod mollusc *Lymnaea stagnalis* (pond snail) was partially digested with plasmin under a variety of experimental conditions and the products of digestion analysed by detergent/polyacrylamide-gel electrophoresis and crossed immunoelectrophoresis. Fragments were obtained corresponding to one, two, three, four and five oxygen-binding domains, one domain having a mol.wt. of approx. 50 000 and containing 2 ions of Cu. The fragments obtained after extensive digestion were non-identical immunologically, and summation of their molecular weights allowed a minimal mol.wt. of 413 000 to be calculated for the original, undigested, eight-domain polypeptide chain. The use of mild-digestion conditions allowed the time course and sequence of the digestion to be monitored. An initial cleavage gave a three-domain and a five-domain fragment. The three-domain fragment was resistant to further digestion. The five-domain fragment could be digested further to give, successively a four-domain, a three-domain, and finally a two-domain fragment, single-domain units being cleaved. These data form the basis for a proposed sequence for the different domains in the original chain.

The oxygen-transporting proteins of gastropod molluscs are copper-containing haemocyanins with mol.wts. of nearly 900 000 (Wood, 1980). It is well-established that one molecule of oxygen is carried per pair of copper atoms in a protein of mol.wt. about 50 000; the native circulating haemocyanin molecule therefore contains some 160 oxygen-binding sites. Complete dissociation of the molecule by 6 M-guanidinium chloride, 1% SDS, etc. yields a polypeptide chain that is believed to correspond to a one-twentieth molecule containing eight oxygen-binding sites (Gielens *et al.*, 1980). It has been concluded from proteolytic-digestion experiments that the oxygen-binding sites in the one-twentieth molecule reside in compact, globular protein domains (mol.wt. approx. 50 000), which are linked by relatively more exposed stretches of polypeptide chain (Gullick *et al.*, 1979; Gielens *et al.*, 1980). The one-twentieth molecule therefore resembles a string of beads with the linking stretches

of polypeptide chains being especially susceptible to attack by proteolytic enzymes, and indeed electron micrographs have been published that confirm this structure (Siezen & Van Bruggen, 1974).

In previous work we obtained fragments of the polypeptide chain of *Lymnaea stagnalis* (pond snail) haemocyanin apparently containing one, two or three oxygen-binding domains (Gullick *et al.*, 1979). These could be separated by gel filtration, and an examination of their absorption and circular-dichroic spectra provided good evidence for the domains in the polypeptide chain being non-identical. In the present work we have extended these studies in two directions. By observing the time course with the proteolytic enzyme plasmin, we have been able to obtain some information on the ordering of the domains in the polypeptide chain, and we have used immunological techniques to confirm that the fragments generated during digestion are distinct. A preliminary report of the latter part of the work has appeared (Wood & Gullick, 1980).

Abbreviations used: SDS, sodium dodecyl sulphate; IgG, immunoglobulin G.

* Present address: The Salk Institute for Biological Studies, San Diego, CA 92138, U.S.A.

† Present address: Marine Ecology Laboratory, Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada.

Materials and methods

Haemocyanin was obtained from the pond snail (*Lymnaea stagnalis*) as described previously, pre-

cautions being taken throughout the isolation and purification procedure to avoid proteolysis by bacteria, etc. (Gullick *et al.*, 1979; Terwilliger *et al.*, 1978). Porcine plasmin (fibrinolysin) and thyroglobulin and the other molecular-weight markers used (see below) were obtained from Sigma (London) Chemical Co., Poole, Dorset, U.K. All other chemicals were obtained from BDH Chemicals Ltd., Poole, Dorset, U.K. Digestion of haemocyanin and separation of the fragments by gel filtration was performed essentially as described previously (Gullick *et al.*, 1979), any modifications to the procedure being noted in the text.

The time course of digestion with plasmin was monitored by withdrawing portions of the reaction mixture at intervals, mixing them with an equal volume of SDS/polyacrylamide-gel sample buffer [50mM-Tris/HCl buffer, pH 6.8, containing 1% (w/v) SDS, 1% (v/v) 2-mercaptoethanol and 10% (v/v) glycerol] and immediately heating to 100°C for 3 min to stop the proteinase action. The samples were then analysed by SDS/polyacrylamide-gel electrophoresis on slabs formed of a 4–15% gradient of polyacrylamide in the buffer system of Laemmli (1970). Gels were stained with Coomassie Blue R250, and molecular weights were estimated from mobilities by comparison with proteins of known molecular weight (thyroglobulin, 330 000; β -galactosidase, 130 000; bovine serum albumin, 68 000; ovalbumin, 45 000; cytochrome *c*, 12 700). It should be noted that the calibration plot [log(molecular weight) against relative mobility] is non-linear in the upper molecular-weight range (see Neville, 1971).

Antibodies to intact, native haemocyanin and to its proteolytic fragments were raised in rabbits by using the immunization schedules described by Wood & Gullick (1980). The IgG was purified by $(\text{NH}_4)_2\text{SO}_4$ precipitation and DEAE-cellulose chromatography (Harboe & Inghild, 1973). Crossed immunoelectrophoresis was performed in 110mm $\times$ 90mm glass plates under the conditions described by Booth *et al.* (1978).

Results

Immunological properties of proteolytic fragments

Digestion of *L. stagnalis* haemocyanin under the conditions described previously (Gullick *et al.*, 1979) yields a mixture of fragments that were separable into three fractions by gel filtration on Sephadex G-150. Fig. 1(a) shows that the starting material, i.e. undigested haemocyanin, one-tenth molecules at the pH used in this experiment, is immunologically homogeneous and consists of a single component. Fig. 1(b) shows the effect of extensive digestion with plasmin; all of the digestion products reacted against antiserum to whole haemocyanin. This corresponds

Fig. 1. Crossed immunoelectrophoresis with *L. stagnalis* haemocyanin and its plasmin digestion products against antiserum to whole haemocyanin

(a) Undigested one-tenth haemocyanin; (b) un-separated digestion mixture. The digestion mixture was separated on a column of Sephadex G-150 to give three peaks (see Gullick *et al.*, 1979) and material from the three peaks was run in crossed immunoelectrophoresis. (c) First peak to emerge, corresponding to material of mol.wt. 150 000, i.e. three-domains; (d) intermediate peak, consisting of material of mol.wt. 100 000 (i.e. two-domain) and collar (smaller peak on left-hand side); (e) third peak corresponding to material of mol.wt. 50 000 (i.e. single domains). [Reproduced from Wood & Gullick (1981) by permission of Marcel Dekker Inc., New York.]

to the material applied to the Sephadex G-150 column and contains some six to eight components. Most of these have been identified by using crossed immunoelectrophoresis as follows.

The first fraction to emerge from the Sephadex G-150 column had a mol.wt. of approx. 150 000, and its immunological reaction to antiserum to whole haemocyanin is shown in Fig. 1(c). Although this material is not completely homogeneous, a single major component, a three-domain fragment, is readily identified. The second fraction to emerge from the column contained two distinct immunological components (Fig. 1d). It was already suspected

from analysis on polyacrylamide gels that these two components were the so-called ‘collar’ domain (Mellema & Klug, 1971) and a two-domain fragment (mol.wt. approx. 100000). The smaller precipitin arc in Fig. 1(d) has been unambiguously identified as collar fragment by its specific reaction with anti-collar antiserum (Wood & Gullick, 1980). The third and final fraction to emerge from the gel-filtration column also contained two immunological components (Fig. 1e), one with a high relative mobility. Both of these fragments had a mol.wt. of about 50000 and are thought to be single oxygen-binding domains.

On the basis that the starting material is immunologically homogeneous and that the identifiable products of plasmin digestion are immunologically distinct (and therefore do not contain overlapping sequences), it is reasonable to combine the molecular weights of the fragments to arrive at a minimal molecular weight for the starting material. By using the molecular-weight values obtained previously (Gullick *et al.*, 1979) a minimal mol.wt. of 413 000 was calculated (Table 1), which agrees well with that obtained by SDS/polyacrylamide-gel electrophoresis, making allowance for the fact that the calibration is non-linear in this region (Neville, 1971). Taking a value of approx. 50 000 per domain, this suggests that the polypeptide chain contains eight domains.

Time course of digestion with plasmin

The fragments described above represent those resulting from ‘complete’ digestion with plasmin, since further exposure to this enzyme did not result in any further cleavages. By using much lower plasmin/haemocyanin ratios, or by performing the digestion at 0°C rather than at 20°C, however,

sequential series of proteolytic cleavages could be observed. By controlling the conditions of digestion in this way, it was possible to stop the reaction at various stages, allowing intermediate products to be isolated for further study.

When samples for SDS/polyacrylamide-gel-electrophoretic analysis were removed from a digestion with 3% (w/v) plasmin proceeding at 0°C, the pattern of bands shown in Fig. 2 was observed. This may be interpreted in terms of an initial single cleavage of the polypeptide chain. If it is accepted that the starting material for digestion contained eight domains, then the pair of products from this first cleavage have molecular weights corresponding to five- and three-domain portions (Fig. 2a). It will be shown below that this three-domain fragment is stable to further digestion with plasmin; for the moment, however, we will concentrate on the fate of the five-domain piece. Fig. 2 shows that this later made a transitory appearance, and as digestion proceeded was cleaved next into a four-domain and a single-domain unit. The single-domain unit had a mol.wt. of about 50 000 by SDS/polyacrylamide-gel electrophoresis and in fact ran as a closely spaced doublet.

The four-domain fragment also had a transitory existence if digestion was allowed to proceed further

Table 1. Summation of molecular weights of haemocyanin plasmin fragments

The column headed ‘Fraction’ shows the fragments as obtained by gel filtration on Sephadex G-150 as described by Gullick *et al.* (1979). Fraction P2 contain both a two-domain fragment and collar, and Fraction P3 contain two types of single-domain fragment, on the basis of immunological evidence.

Fraction	Molecular weight	No. of O ₂ -binding domains
P1	157 000	3
P2	100 000	2
Collar	~60 000	1
P3	51 000	1
	45 000	1
	413 000	8

Fig. 2. Time course of digestion of *L. stagnalis* haemocyanin with plasmin

In each case the plasmin/haemocyanin ratio was 3% (w/w), but series (a) digestions were performed at 0°C and series (b) at 20°C; the buffer was 50 mM-Tris/HCl, pH 8.2, containing 50 mM-NaCl. The first track (a1, far left) contains calibration markers, and the calibration is shown on the side; the arrow shows the position of undigested haemocyanin. Portions were removed from the digestion mixture and mixed with SDS/polyacrylamide-gel buffer at the following times after adding plasmin: (a 2–9) 0, ½, 1, 3, 5, 15, 30, 60 min; (b 1–7) 3, 5, 7, 10, 15, – (nothing applied), 60 min.

(Fig. 2b), being cleaved to give another single-domain fragment and a three-domain fragment that was different from the stable one described in the preceding paragraph. The one-domain fragment cleaved at this stage (again running as a closely spaced doublet) had a molecular weight by SDS/polyacrylamide-gel electrophoresis of about 65 000, and was shown to be collar by virtue of its electrophoretic behaviour and immunological properties.

Finally, the three-domain fragment produced as a result of the above cleavage, and which may be referred to as 'unstable', undergoes a further cleavage. The products of this cleavage are a two-domain and a one-domain fragment. The two-domain fragment is stable to further digestion and corresponds to the 100 000-mol.wt. band shown in Fig. 1(d). The one-domain fragment resulting from this final cleavage has not been identified with any degree of certainty, and may itself be susceptible to further digestion to give fragments of mol.wt. < 50 000. Scheme 1 summarizes this tentative series of cleavages.

After analytical digestion, runs were performed on a larger scale and attempts were made to isolate and purify some of the intermediate cleavage products. The initial products of plasmin cleavage, that is to say, five-domain and three-domain (stable) pieces, were separated from one another and from undigested haemocyanin by adoption of the following procedure (Fig. 3). The three-domain fragment was separated from the mixture by gel chromatography on Sepharose CL6B at pH 8.2 (Fig. 3a). In this system the undigested haemocyanin and the five-domain fragment eluted together as a single broad peak. There is some evidence that the five-domain fragment self-associates (Gielens *et al.*, 1980). The mixture of five-domain fragment and undigested haemocyanin were then separated from each other by a second gel-filtration step on Sepharose 4B in the

presence of 25 mM-CaCl₂ and at pH 7.5. The addition of Ca²⁺ and the lowering of the pH caused the undigested haemocyanin, but not the five-domain fragment, to reassociate to give whole, native molecules (mol.wt. approx. 9 000 000) and which were therefore easily separated from the five-domain fragment (mol.wt. approx. 250 000, Fig. 3b). Although only small amounts of material have been isolated in this way, it was possible to show that (a) this three-domain fragment is resistant to further digestion by plasmin, and that (b) all the other fragments observed arose from cleavage of the five-domain fragment (Scheme 1).

Separation of two types of three-domain fragment

By the choice of the appropriate conditions,

Fig. 3. Separation of multi-domain fragments of *L. stagnalis* haemocyanin by gel filtration

(a) Digestion with plasmin (0.003%) at 20°C for 1 h, followed by fractionation on a column (142 cm × 1 cm) of Sepharose CL6B equilibrated in 50 mM-Tris/HCl buffer, pH 8.2, containing 50 mM-NaCl. The flow rate was 7 ml/h, and the fraction size was 2.6 ml. The first peak to emerge contained a mixture of five-domain fragments plus undigested haemocyanin; the second peak contained three-domain fragments that were stable to further plasmin digestion. (b) First peak from (a) concentrated, dialysed into 100 mM-Tris/HCl buffer, pH 7.5, containing 25 mM-CaCl₂, and re-run on Sepharose CL6B, equilibrated in the same buffer. The first peak (void volume) contained undigested haemocyanin polypeptide chains, and the second peak five-domain fragments.

Scheme 1. Fragmentation of *Lymnaea* haemocyanin polypeptide chain by plasmin

The native polypeptide chain contains eight oxygen-binding domains. Numbers underlined show fragments that are stable to further digestion under the conditions employed.

Fig. 4. Fractionation of plasmin digestion products of *L. stagnalis* haemocyanin of DEAE-Sephacel CL6B. The column, of volume 30ml, was equilibrated with 10 mM-Tris/HCl buffer, pH 7.5. The initial plasmin/haemocyanin digestion ratio was 3% (w/w). The fraction size was 10ml; 200ml of buffer containing 200 mM-NaCl was run through before starting a gradient of 0.2–0.5 M-NaCl (----).

Fig. 5. Separation of two types of three-domain fragments of *L. stagnalis* haemocyanin on SDS/polyacrylamide gels and their digestion with plasmin

Gel (a) shows a set of marker proteins (track 1) with the corresponding molecular weights on the extreme left; tracks 2, 3 and 4 are samples from the three peaks, (III, II, I respectively, see Fig. 4) obtained by fractionation of the plasmin digest of haemocyanin on DEAE-Sephacel CL6B. Gel (b) shows the result of further digestion of the fractions with plasmin. The strong band from track (a)2 was digested and then run on track (b)1; the strong band from track (a)3 was digested and then run on track (b)2. In the latter case, it is clear that little digestion has taken place thus showing that the two three-domain fragments differ in their susceptibility to proteolysis.

especially the time of digestion, it was possible to obtain a mixture of products that included both types of three-domain fragment, stable and unstable. These could be separated from one another and from the other products by ion-exchange chromatography on DEAE-Sephacel CL6B, with an [NaCl] gradient as eluent (Fig. 4). The fractionation procedure produces three peaks, and samples from each of these were analysed by SDS/polyacrylamide-gel electrophoresis (Fig. 5a). The first peak to emerge from the column contained a three-domain fragment, but was heterogeneous and was not investigated further. Peaks I and II (Fig. 4) both contained three-domain fragments that had slightly different mobilities on SDS/polyacrylamide-gel electrophoresis. These were further purified by gel filtration and then subjected to digestion with plasmin. The result is shown in Fig. 5(b); the faster-moving three-domain fragment is the 'unstable' one (track 2 in Fig. 5a, and peak III in Fig. 4) and was digested to give a two-domain fragment that had an apparent mol.wt. of about 100 000 and a single-domain fragment with an apparent mol.wt. of 45 000 (doublet). The more slowly-moving three-domain fragment was resistant to further digestion, and is therefore one of the initial products of cleavage of the native, eight-domain polypeptide chain. The molecular weights of the two three-domain fragments were determined in a sedimentation-equilibrium experiment by using the meniscus-depletion procedure. The values obtained, at initial concentrations of about 0.3 mg/ml, were 127 000 (unstable) and 157 000 (stable to plasmin digestion).

Discussion

It may be concluded from the present experiments that *L. stagnalis* haemocyanin has a structure closely similar to, but not identical with, that of the β_c -haemocyanin of *Helix pomatia* (Roman snail) (Lontie & Gielens, 1980). This haemocyanin also appears to be composed of eight-domain polypeptide chains, and a good deal of information on its structure has been obtained by a study of its limited proteolysis (Gielens *et al.*, 1980). On the basis of the cleavage pattern, the properties of the fragments obtained, and some *N*-terminal analyses, Lontie and his co-workers (Gielens *et al.*, 1980) have proposed a sequence or order for the domains along the polypeptide chain of the one-twentieth molecule. It is instructive to compare the results from plasmin cleavage of *L. stagnalis* haemocyanin with this proposed structure. It is not, however, possible to go as far in the deductions about the structure of *L. stagnalis*, since we have, despite many attempts, failed to identify the *N*-terminal amino acid(s) of this haemocyanin.

The present work provides good evidence that the fragments obtained by plasmin cleavage are non-identical, and this may be added to the evidence from spectral studies already reported (Gullick *et al.*, 1979). Lontie & Gielens (1980) have designated the domains of the native polypeptide chain *a...h*, and domain *a* carries the *N*-terminus. In *L. stagnalis* haemocyanin this may be blocked. It has been proposed that the so-called collar of the molecule is at the *C*-terminal end, on the basis that selective removal of collar domains does not change the *N*-terminus. If this is accepted, then the stable three-domain fragment which initially appears in the plasmin cleavage of *L. stagnalis* haemocyanin, corresponds to the *N*-terminal region *a-b-c*; all subsequent cleavages are in the section *d-h*, which was demonstrated to carry the collar domain. As the second cleavage (i.e. $5 \rightarrow 4 + 1$) does not release collar domain, the one-domain fragment would correspond with domain *d*, collar fragment *h* being released in the next step (i.e. $4 \rightarrow 3 + 1$). The three-domain fragment, which was susceptible to further plasmin digestion, is then *e-f-g*; it is impossible to say at present which bond is cleaved finally to yield the two-domain piece. This proposed sequence of cleavages is summarized in scheme 2. Although it is a reasonable scheme on the basis of the comparison with *Helix pomatia* β_c haemocyanin, it can only be tentative until further evidence is obtained. Unlike *H. pomatia*, *L. stagnalis* appears to have only a single type of haemocyanin in the haemolymph.

The present work lends additional support to the idea that the polypeptide chain of gastropod haemocyanin is an enormous structure of mol.wt. > 400000 and containing eight domains, although our previous evidence led us to doubt this (Gullick *et al.*,

Scheme 2. Digestion of *L. stagnalis* haemocyanin with plasmin

The domains in the original polypeptide chain are designated (a)–(h), and the order of the cleavages is shown above. It is assumed that *h* is the collar domain and is at the *C*-terminus. The fourth cleavage may be between *f* and *g* rather than between *e* and *f* (see Lontie & Gielens, 1979).

1979). The supporting arguments may be summarized as follows: (1) undigested haemocyanin had a lower mobility than thyroglobulin ($M_r > 330000$) on SDS/polyacrylamide-gel electrophoresis, and (2) digestion of the haemocyanin with proteolytic enzymes yielded immunologically distinct fragments about whose molecular weights there is no controversy, as they fall within the normal calibration range. Therefore it is justifiable to add up the weights of all the products of digestion to arrive at a minimal molecular weight for the undigested material. The analytical-ultracentrifuge results of Quittner *et al.* (1978) with *Murex fulvescens* (whelk) haemocyanin seemed to point to a mol.wt. of about 300000 for this haemocyanin. It is possible that despite precautions taken, some proteolysis had occurred; it is clear from Fig. 2 that proteolysis will continue at low concentrations of enzyme and at 0°C.

It is difficult to relate the proposed structure for the polypeptide chain of gastropod haemocyanin to the structure of the native circulating molecule and its function. It may be presumed that the immunological non-identity of most of the domains as shown in the present work is providing information largely about the outside surface of the domain regions. The function of these topological areas is likely to be connected with controlling the assembly and final organization of the twenty polypeptide chains in the native molecule. Possibly the internal regions differ less from one another. Our peptide 'maps' (Gullick & Wood, 1979) suggested that there were regions of homology within the polypeptide chains of the various domains. This would be expected, if, as is highly probable, the domains arose by gene duplication. On the other hand, the spectral data for the different domains (Gullick *et al.* 1979; Gielens *et al.*, 1980) suggest that there are differences in the amino acids in the proximity of the copper–oxygen binding site. Such differences may of course be very subtle ones, but may nonetheless be important in deciding the oxygen-binding properties of the different domains. Finally there must also be interactions, presumably through surface amino acid residues, that are responsible for mediating the co-operativity in oxygen-binding observed with these haemocyanins and accounting for the Bohr effect. We know little about these interactions and nothing about the contacts by which they arise.

This work was supported by grants from the Science Research Council. We thank Professor R. Lontie, Katholieke Universiteit te Leuven, Belgium, for valuable discussions.

References

- Booth, A. G., Hubbard, L. M. L. & Kenny, A. J. (1978) *Biochem. J.* **179**, 397–405

- Gielens, C., Vershueren, L. J., Préaux, G. & Lontie, R. (1980) *Eur. J. Biochem.* **103**, 463–470
- Gullick, W. J. & Wood, E. J. (1979) *Biochem. Soc. Trans.* **7**, 395–396
- Gullick, W. J., Herries, D. G. & Wood, E. J. (1979) *Biochem. J.* **179**, 593–602
- Harboe, N. & Inghild, A. (1973) in *Manual of Quantitative Immunelectrophoresis* (Axelsen, N. H., Kroll, J. & Weeke, B., eds.), pp. 161–164, Universitets forlaget, Oslo
- Laemmli, U. K. (1970) *Nature (Lond.)* **229**, 145–150
- Lontie, R. & Gielens, C. (1979) in *Metalloproteins: Structure, Molecular Function and Clinical Aspects* (Weser, U., ed.), pp. 62–72, Georg Thieme, Stuttgart and New York
- Mellema, J. E. & Klug, A. (1972) *Nature (London)* **239**, 145–150
- Neville, D. M. (1971) *J. Biol. Chem.* **246**, 6328–6334
- Quittner, S., Watts, L. A. & Roxby, R. (1978) *Anal. Biochem.* **89**, 187–195
- Siezen, R. J. & Van Bruggen, E. F. J. (1974) *J. Mol. Biol.* **90**, 77–89
- Terwilliger, R. C., Terwilliger, N. B., Bonaventura, C. & Bonaventura, J. (1977) *Biochem. Biophys. Acta* **494**, 416–425
- Wood, E. J. (1980) *Essays Biochem.* **16**, 1–47
- Wood, E. J. & Gullick, W. J. (1981) in *Structure, Active Site and Function of Invertebrate Oxygen Carriers* (Lamy, J., ed.), Marcel Dekker, New York, in the press