Evidence for a Repeating 3,4-Dihydroxyphenylalanine- and Hydroxyproline-containing Decapeptide in the Adhesive Protein of the Mussel, *Mytilus edulis* L.

J. Herbert Waite

From the Orthopaedics Research Laboratory, University of Connecticut, Farmington, Connecticut 06032

(Received for publication, September 16, 1982)

Previous work has shown that the permanent adhesive of the marine mussel *Mytilus edulis* is a protein containing large amounts of hydroxyproline (13%) and 3,4-dihydroxyphenylalanine (Dopa, 11%). The protein also known as the polyphenolic protein is produced and stored in the exocrine phenol gland of the mussel and deposited onto marine surfaces by the animal's foot during the formation of new adhesive plaques. The adhesive protein has been purified by a combination of ion exchange on sulphonylpropyl-Sephadex and gel filtration on low surface energy chromatographic media. Polyacrylamide gel electrophoresis of the protein at acidic pH shows it to consist of two components having a molecular weight of about 130,000. Treatment of the protein with clostridial collagenase reduced the molecular weight by less than 10%. The collagenase-resistant fragment contains most or all of the Hyp and Dopa. Trypsin treatment of the polyphenolic protein results in extensive degradation. The major tryptic peptide (80%) contains 10 amino acids including Hyp and Dopa and was shown by sequence analysis to be HZN-Ala-Lys-Pro-Ser-Tyr-Hyp-Hyp-Thr-Dopa-Lys-COOH. Calculations suggest that this and related sequences may be repeated as often as 75 times in the polyphenolic protein.

Permanently sessile invertebrates in the sea have necessarily evolved adhesive strategies to resist the impact of waves (1, 2). The common mussel *Mytilus edulis* secures itself to solid substrates through a complex array of plaque-tipped collagenous byssal threads. On glass, the attachment plaques display an average adhesive tensile strength of 10^6 newton-meter^-2 although maximal values often exceed 10^7 newton-meter^-2 (2). The substance in the plaque mediating adhesion between the collagenous threads and the substrate is the polyphenolic protein (3–6). The polyphenolic protein is attracting much attention as an adhesive since, unlike most synthetic adhesives, its performance, polymerization, and longevity are not adversely affected by the presence of water. Although the reason for this resistance to water remains unknown, it is very likely related to the unusual chemical composition of the polyphenolic protein. Waite and Tanzer (6) reported that prior to polymerization the polyphenolic protein consists of a rather large polypeptide chain (M_r = 130,000) in which seven amino acids, lysine, hydroxyproline, alanine, serine, threonine, tyrosine, and Dopa, account for about 80% of all the residues. The occurrence of Dopa and 3- and 4-hydroxyprolines seems particularly odd. Dopa is only rarely encountered as a component of naturally occurring proteins (7), and hydroxyprolines are primarily associated with collagens in which every third residue is glycine (8). The very low glycine content (3%) of the polyphenolic protein tends to militate against a resemblance to collagen.

The objective of this study was to describe a purification of the polyphenolic protein and to characterize a Dopa-containing peptide prepared from the polyphenolic protein by enzymatic digestion. The results suggest that the adhesive polyphenolic protein may consist in large part of a repeated Dopa-containing sequence.

MATERIALS AND METHODS

The purification of the polyphenolic protein from dissected phenol glands is outlined in Table I. The rather low solubility of the polyphenolic protein at neutral pH can be exploited by first extracting extraneous proteins with large amounts of neutral salt buffer followed by gentle centrifugation and re-extraction of insoluble materials with dilute acetic acid. The neutral salt buffer contains various protease inhibitors to prevent premature degradation of polyphenolic protein and cyanide to avoid enzymatic oxidation of Dopa residues. The gentle first centrifugation is critical to prevent the irreversible coalescence of insoluble proteins in the pellet. Developing a high yield purification of polyphenolic protein has been difficult due to extensive adsorption of the protein to surfaces. Considerable improvements are necessary. Ion exchange on sulphonylpropyl-Sephadex (Fig. 1) provides the single most effective step; however, up to 70% of the applied polyphenolic protein is not recovered (even at 6 M guanidine hydrochloride). Gel filtration of polyphenolic protein using a variety of chromatographic materials (Bio-Gel P, Sepharose, Sephacryl-S and Fractogel-Merck) and buffers has resulted in very low or

---

*This work was supported by Grant DE 05956 from the National Institutes of Health, Grant PCM-8206463 from the National Science Foundation, and Grant 35-084 from the Connecticut Research Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

---

1 The abbreviations and trivial name used are: Dopa, 3,4-dihydroxyphenylalanine; GuHCl, guanidine hydrochloride; PMSF, phenylmethylsulfonyl fluoride; NEM, N-ethylmaleimide.

2 Portions of this paper (including "Materials and Methods," Figs. 1–3 and 6–10, Tables II–IV, and a scheme) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9600 Rockville Pike, Bethesda, MD 20814. Request Document No. 82M-2555, cite the author, and include a check or money order for $6.40 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.
Repeating Dopa- and Hyp-containing Decapeptide

negligible yields. Yields are improved on Sephadex if an elution buffer with low pH (2-4) and a cationic detergent is used (Fig. 2). Recovery from phenyl-Sepharose 4B is excellent although the limited fractionation range of this material is not particularly helpful (Fig. 3). Using the Dopa/protein ratio as an index of purity, the highest Dopa/protein ratios observed were 0.155. As indicated in Table 1, this corresponds to about 110 residues of Dopa/1000 residues of amino acids.

As noted earlier (6), extracts of polyphenolic protein from the phenol gland contain two species (A and B) which migrate closely on polyacrylamide gels (Fig. 4). These two species are co-purified using all of the methods presented. In extractions from individual mussel glands, the ratio of A to B seems to vary (with B > 50%) without known reason from one individual to the next. A separation of the two has not yet been achieved for comparison. The apparent molecular weight of the polyphenolic protein as determined by polyacrylamide gel electrophoresis in the presence of cetylpyridinium bromide is estimated to be 130,000 ± 10,000 (Fig. 5). The amino acid composition is shown in Table I.

Collagenase treatment of polyphenolic protein results in a limited degradation (Fig. 6). Only about 8% of the protein was attacked, leaving entirely intact a fragment with $M_r = 120,000$. This fragment has an amino acid composition similar to that of the original protein but is noticeably reduced in glycine and proline (Table II). Clostridial collagenase selectively cleaves

TABLE I

| Step                  | Total protein* | Total dopamine* | Dopa/protein | Yield | Fold |
|-----------------------|----------------|-----------------|--------------|-------|------|
| Neutral salt homogenate | 4080           | 4.9             | 0.0012       | 100   | 1    |
| Acetic acid homogenate | 132            | 3.7             | 0.028        | 75    | 23   |
| Peak SP-Sepharose      | 17             | 1.6             | 0.095        | 33    | 79   |
| Sephadex G-200         | 11             | 1.4             | 0.127        | 28    | 106  |
| Phenyl-Sepharose 4B    | 10             | 1.4             | 0.155        | 28    | 129  |

* As measured by Hartree (9).
* As measured by Waite and Tanzer (10).

Trypsin digestion of polyphenolic protein is extensive and rapid. Under the conditions used, polyphenolic protein completely disappeared within 5 min of the addition of trypsin as determined by gel electrophoresis. Fractionation of the tryptic digest was achieved using gel filtration on Sephadex LH-60 (Fig. 7), which removes trypsin from the peptides, followed by ion exchange on Sephadex SP-25 (Fig. 8) with a pyridine acetate gradient (the latter method being recommended for basic aromatic peptides (16)). Fig. 8 illustrates that 75-80% of the ninhydrin-positive material eluting from SP-Sepharose can be ascribed to the Dopa-containing peak. Moreover, nearly 95% of the Dopa and 80% of the protein as measured by Waite and Tanzer (10) and Hartree (9), respectively, originally applied to SP-Sepharose are recovered in the major Dopa-rich peak. This material was further purified by passage through Sephadex LH-20 (Fig. 9). The tryptic peptide resembles the polyphenolic protein in containing the same group of amino acids that predominate in the latter, namely Hyp, Thr, Ser, Pro, Ala, Dopa, Tyr, and Lys. In the tryptic peptide, however, Dopa and Hyp are significantly enriched, whereas Lys, Pro, and Tyr are lower (Table II). The tryptic peptide was homogeneous on 12% acrylamide gels in 3 M urea and 5% acetic acid but was visualized by Dopa staining since it could not be fixed for protein staining. Molecular weight in cetylpyridinium bromide gel electrophoresis was estimated to be about 6500 (Fig. 10). Peptide homogeneity was also suggested by thin layer chromatography on cellulose in 1.5% formic acid and thin layer electrophoresis in 5% acetic acid (Table III). Since borate strongly complexes Dopa at pH 7-9, it has the property of introducing additional negative charges into the peptide (17). Heterogeneity of the Dopa peptide (2 spots) in borate (Table III) suggests variation in the degree of Tyr to Dopa conversion.

Sequenator analysis (35 cycles) of the tryptic Dopa peptide (Table IV) revealed it to be a decapeptide (molecular weight of about 1400), in contrast to expectations based on the estimated molecular weight of 6500. The amino acid sequence is given in Fig. 11, and features 2 lysines, 2 hydroxyprolines, 2 (Dopa + tyrosine), 1 proline, 1 serine, 1 threonine, and 1 alanine. Lys 2 is clearly protected from tryptic proteolysis by the presence of Pro 3. Dopa is detected mostly penultimate to the carboxyl terminus although significant amounts are also

Fig. 4. Acidic polyacrylamide gel electrophoresis of polyphenolic protein-containing samples. Lane a, 0.8 M acetic acid extract of polyphenolic protein from phenol glands (30 µg stained with Serva blue R-250); lane b, same as lane a (60 µg stained for Dopa (6)); lane c, polyphenolic protein after purification on phenyl-Sepharose 4B (10 µg stained with Serva blue R-250). In each case, electrophoresis was extended for 1 h following elution of the tracker methyl green from the gels.

Fig. 5. Clostridial collagenase digestion of polyphenolic protein (species A and B) as detected by polyacrylamide gel electrophoresis. Purified polyphenolic protein (1 mg/ml) dissolved in 0.1 M sodium borate (pH 8.0) was treated with protease-free collagenase (4 µg/ml) at 35 °C. Digestion times for lane a, 0 min; lane b, 15 min; lane c, 45 min; lane d, 1.5 h; lane e, 4 h; lane f, 8 h; and lane g, 24 h.
FIG. 11. Amino acid sequence of the tryptic Dopa peptide. Residues marked by an asterisk contain both hydroxylated (Hyp/Dopa) and prehydroxylated (Pro/Tyr) species. Amino acid shown is present with Tyr 5. The 2 Hyps are located next to one another at positions 6 and 7; additional Hyp is present with Pro 3. 3-Hyp as well as 4-Hyp occurs at position 7.

DISCUSSION

The adhesive polyphenolic protein of the mussel M. edulis contains a large number of hydroxylated amino acids (500 residues/1000). Among these, Dopa and 3-4-hydroxyproline are presumably derived post-translationally from tyrosine and proline, respectively. Although hydroxyproline is erroneously assumed to be unique to collagenous proteins, in the polyphenolic protein also from Mytilus, also from Periostracin, the collagenase-labile moiety may contain a collagenase-resistant fragment. The collagenase-labile moiety may contain a collagenase-resistant fragment. The collagenase-labile moiety may contain a collagenase-resistant fragment. The collagenase-labile moiety may contain a collagenase-resistant fragment. The collagenase-labile moiety may contain a collagenase-resistant fragment.

Perhaps future structural studies with synthetic analogues of the Dopa peptide will shed some light on what conformational analyses are adopted in solution and how these conformations might be related to the adhesive function of the polyphenolic protein.

ACKNOWLEDGMENTS—I am grateful for the technical expertise of Tim Housley, George Korza, and Ida Shteyngard. Dr. Juris Ozols helped me interpret sequenator analysis. I thank Dr. Marv Tanzer for his continued interest in this work.

REFERENCES

1. Walker, G. (1981) J. Adhesion 12, 51-58
2. Young, G. A., and Crisp, D. J. (1982) Adhesion 6, 88-39
3. Brown, C. H. (1983) J. Microsc. Sci. 9, 487-562
4. Tamarin, A., Lewis, P., and Aoki, J. (1970) J. Morphol. 139, 196-202
5. Wain, J. H., and Tanzer, M. L. (1980) Biochem. Biophys. Res. Commun. 79, 1504-1561
6. Wain, J. H., and Tanzer, M. L. (1981) Science (Wash. D.C.) 212, 1038-1040
7. Wain, J. H., Saudelli, A. M., and Andersen, S. D. (1976) J. Comp. Physiol. 110, 301-307
8. Yrres, D. R. (1980) Science (Wash. D.C.) 205, 115-1122
9. Hartree, F. E. (1972) Anal. Biochem. 48, 425-427
10. Wain, J. H., and Tanzer, M. L. (1981) Anal. Biochem. 112, 133-136
11. Panxyin, S., and Chilka, H. (1989) Arch. Biochem. Biophys. 130, 337-346
12. Marjanned, L. A., and Hyne, J. B. (1974) Biochem. Biophys. Acta 371, 436-450
13. Sekhon, M. (1975) Anal. Biochem. 63, 345-349
14. Weber, R., and Osborn, M. (1960) J. Biol. Chem. 246, 4400-4412
15. Moore, S. R., and Stein, W. H. (1949) J. Biol. Chem. 176, 367-377
16. Kasper, C. B. (1955) Mol. Biol. Biophys. Biophys. 8, 114-131
17. Michl, H. (1952) Monatsh. Chem. 83, 735-747
18. Simpson, R. J., Neuberger, M. R., and Liu, T.-Y. (1976) J. Biol. Chem. 251, 1936-1940
19. Spackman, D. H., Stein, W. H., and Moore, S. (1958) Anal. Chem. 30, 1596-1596
20. Onishi, J., and Heinemann, F. S. (1983) Biochem. Biophys. Acta 764, 163-173
21. Oulou, J., Gerard, C., and Nobrega, F. G. (1976) J. Biol. Chem. 271, 6677-6677
22. Pekoszakal, B. (1962) Enzym. Insect Chem. 82A, 451-457
23. May, C., and Rosenthal, T. L. (1981) Biochemistry 20, 2810-2817
24. Bhattacharyya, S. N. (1981) Biochem. J. 192, 447-457
25. Wain, J. H. (1983) in Biochemistry of Molluscs (Hochachka, P. W., ed) Academic Press, New York, in press
26. Soria, M. F., and Hubbard, A. T. (1982) J. Am. Chem. Soc. 104, 2753-2742
27. Davies, H., and Frahn, J. L. (1977) J. Chem. Soc. (Lond.) 30, 2296-2297
28. Piekarszak, J., Rautten, J., Vastamok, M., Lamplab, K., Kari, A., and Kuklen, K. (1986) Eur. J. Biochem. 4, 555-559
29. Smyth, J. D. (1946) Q. J. Microsc. Sci. 95, 136-152
30. Vitellin-Zarcellaro, L. (1981) Tissue Cell 13, 701-713
31. Gray, W. R., Sandberg, L. B., and Foster, J. A. (1973) Nature (London) 241, 466-466
32. Gasp, J. P. (1982) Trends Biochem. Sci. 7, 105-108
33. Bennick, A. (1982) Mol Cell Biochem. 45, 81-99
34. Choa, P. Y., and Faasen, G. D. (1976) Advan. Res. Biochem. 47, 251-276
35. Yamamoto, H., and Hasayawa, T. (1978) Polymer 19, 11883-11103
36. Catt, J. W., Ellis, J. H., and Infecta, J. (1975) Planta 131, 195-197
37. Toszywolsky, K. J. (1976) Planta (Berl.) 138, 235-240
38. Ashton, D., and Nothegger, W. (1980) Trends Biochem. Sci. 5, 245-248
39. Usui, K., Yoshida, K., and San Clemente, C. L. (1981) Curr. J. Microbiol. 27, 955-956
40. Tana, S., and Nakai, K. (1981) J. Biol. Chem. 236, 11397-11400
41. Langer, P. T., A. Kata, L., and Herig, S. (1973) Biochem. J. 133, 129-131

Repeating Dopa- and Hyp-containing Decapeptide
Peptide sequencing was done on a Beckman 890 sequencer with the 8.5 M urea protocol (20). The phenylthiocarbamyl (PTC) derivatives were identified by high performance liquid chromatography (HPLC). The amino acid analysis was performed on a Beckman amino acid analyzer (model 6330). The protein samples were hydrolyzed in 6 N hydrochloric acid at 110°C for 24 h. The hydrolysates were analyzed on a Beckman amino acid analyzer following hydrolysis of the PDC derivatives in hydroxylamine (21).

The hydroxyproline-FPLC was used to completely degrade new hydrolysates.

| TABLE I | AMINO ACID COMPOSITION OF THE POLYPEPTIDE SOLID PHASE CHROMATOGRAPHY | YIELD | AMINO ACID IDENTIFIED BY FPLC | AFTER HYDROLYSIS |
|---------|-------------------------------------------------|-------|-------------------------------|-----------------|
| Glycine | 8.3                                            | 1     | Glycine                       | 1               |
| Alanine | 12.2                                           | 1     | Alanine                       | 1               |
| Serine  | 13.2                                           | 1     | Serine                        | 1               |
| Proline | 10.6                                           | 1     | Proline                       | 1               |
| Tyrosine| 7.5                                            | 1     | Tyrosine                       | 1               |
| Lysine  | 11.5                                           | 1     | Lysine                         | 1               |
| Arginine| 13.9                                           | 1     | Arginine                       | 1               |
| Aspartic acid | 12.6                                           | 1     | Aspartic acid                       | 1               |
| Glutamic acid | 13.2                                           | 1     | Glutamic acid                       | 1               |
| Cysteine | 5.3                                            | 1     | Cysteine                       | 1               |
| Methionine | 7.8                                            | 1     | Methionine                       | 1               |
| Tryptophan | 11.2                                           | 1     | Tryptophan                       | 1               |
| Phenylalanine | 12.7                                           | 1     | Phenylalanine                       | 1               |
| Isoleucine | 10.9                                           | 1     | Isoleucine                       | 1               |
| Leucine  | 15.9                                           | 1     | Leucine                         | 1               |
| Valine   | 12.9                                           | 1     | Valine                          | 1               |
| Threonine| 11.9                                           | 1     | Threonine                       | 1               |
| Histidine| 13.1                                           | 1     | Histidine                       | 1               |
| Serine  | 13.2                                           | 1     | Serine                        | 1               |
| Aspartic acid | 12.6                                           | 1     | Aspartic acid                        | 1               |
| Glutamic acid | 13.2                                           | 1     | Glutamic acid                       | 1               |
| Cysteine | 5.3                                            | 1     | Cysteine                       | 1               |
| Methionine | 7.8                                            | 1     | Methionine                       | 1               |
| Tryptophan | 11.2                                           | 1     | Tryptophan                       | 1               |
| Phenylalanine | 12.7                                           | 1     | Phenylalanine                       | 1               |
| Isoleucine | 10.9                                           | 1     | Isoleucine                       | 1               |
| Leucine   | 15.9                                           | 1     | Leucine                         | 1               |
| Valine    | 12.9                                           | 1     | Valine                          | 1               |
| Threonine | 11.9                                           | 1     | Threonine                       | 1               |
| Histidine | 13.1                                           | 1     | Histidine                        | 1               |

*Estimated composition for peptide with 40 residues in parentheses. Values for Ser, Thr, and Cys were corrected for loss due to hydrolysis.
Repeating Dopa- and Hyp-containing Decapeptide

Fig. 1 Gel filtration of the Dopa-rich peak from SP Sephadex on Sephadex G-200. Elution buffer is 0.1 M ammonium acetate (pH 5.1) with 0.01% TPB. Elution buffer used volume and bar denotes fractions pooled for further purification.

Fig. 3 Gel filtration of Dopa-rich fractions on Sephadex G-200. Elution buffer is 0.2 M acetic acid. Fractions under bar were pooled and concentrated. Fractions were assayed for ultraviolet light (---) and o-dihydroselenoacetic acid (-----).

Fig. 5 Molecular weight estimation of the polyphenolic proteins by polyacrylamide gel electrophoresis in the presence of CPB on 0.5-M acetate buffer. A linear relationship between mobility and molecular weight is obtained by using double log coordinates. Mobility is measured in cm.

Fig. 7 Gel filtration of tryptic-digested polyphenolic proteins on Sephadex G-200. Elution buffer is 0.02 M acetic acid. Fractions under bar were pooled and concentrated. Fractions were assayed for ultraviolet light (-----) and o-dihydroselenoacetic acid (-----).

Fig. 9 Gel filtration of tryptic-digested polyphenolic proteins on Sephadex G-200. Elution buffer is 0.1 M acetic acid. Fractions were assayed for ultraviolet light (-----) and o-dihydroselenoacetic acid (-----). Fractions under bar were pooled and flash evaporated.

Fig. 11 Molecular weight estimation of trypsin-digested proteins by polyacrylamide gel electrophoresis in the presence of CPB on 0.5-M acetate buffer. A linear relationship between mobility and molecular weight is obtained by using double log coordinates. Mobility is measured in cm.
Evidence for a repeating 3,4-dihydroxyphenylalanine- and hydroxyproline-containing decapeptide in the adhesive protein of the mussel, *Mytilus edulis* L.

J H Waite

*J. Biol. Chem.* 1983, 258:2911-2915.

Access the most updated version of this article at [http://www.jbc.org/content/258/5/2911](http://www.jbc.org/content/258/5/2911)

Alerts:
- When this article is cited
- When a correction for this article is posted

[Click here](http://www.jbc.org/content/258/5/2911.full.html#ref-list-1) to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/258/5/2911.full.html#ref-list-1](http://www.jbc.org/content/258/5/2911.full.html#ref-list-1)