Marine Biology (1999) 134: 729–734

S. C. Warner · J. H. Waite

Expression of multiple forms of an adhesive plaque protein in an individual mussel, *Mytilus edulis*

Received: 11 August 1998 / Accepted: 15 April 1999

Abstract Individual blue mussels, *Mytilus edulis* L., can express at least 20 variants of a small protein known as *M. edulis* foot protein 3 or Mefp3. Mefp3 has been shown to be a component of the adhesive plaque of the byssus, the structure securing mussels to solid substrata. The cDNAs and deduced fp3 protein sequences display more variation at the carboxy-terminus than at the N-terminus, although there is some variation present throughout the protein. This indicates that there most likely are multiple copies of the gene encoding this protein. Each protein sequence contains a signal peptide, 24 to 25 residues in length, and a mature protein sequence of 44 to 54 residues. Gly is the most common amino acid in the mature protein at 20 to 25 mol%. Tyr and Arg follow closely at 20 to 23 and 16 to 21 mol%, respectively. Both of these amino acids were previously shown to be post-translationally modified to 3,4-dihydroxyphenylalanine (Dopa) and 4-hydroxyarginine, respectively, in this protein. MALDI-TOF (matrix-assisted laser desorption ionization with time-of-flight) mass spectrometric analysis of the underside of adhesive plaques reveals the presence of Mefp3-like proteins. Curiously, only four or five out of 20 possible fp3 variants are detectable in plaques deposited on glass or plastic. This would suggest that selection of protein variants for deposition onto surfaces is determined at the level of translation.

Communicated by J.P. Grassle, New Brunswick

S.C. Warner
College of Marine Studies,
University of Delaware,
Newark, Delaware 19716, USA

J.H. Waite (✉)
Department of Molecular,
Cell and Developmental Biology,
University of California,
Santa Barbara, California 93106, USA

Introduction

Like other sessile organisms in the intertidal zone, marine mussels rely on the durability and strength of their holdfasts for secure attachment. The mussel holdfast or byssus is a peculiar extracorporeal bundle of collagenous threads (Waite et al. 1998). Byssal thread attachment to a hard substratum is mediated distally by flattened adhesive plaques. Each new thread and plaque is made in the ventral groove of the mussel's foot after it has explored available surfaces using a sensory patch near the tip (Mahéo 1970). Due to factors such as water turbulence, emersion, intraspecific competition for space and predators, mussels must be expeditiously opportunistic in bonding to hard surfaces. The plaque–substratum interface is consequently the most unpredictable and, perhaps, vulnerable link in the chain of molecules that comprises the holdfast. Fueled by antifouling and biomimetic research, there has been a growing interest in the molecular adaptations that facilitate opportunistic adhesive bonding in marine organisms.

Recent biochemical analyses of adhesive plaques indicate that these are composed of at least four prominent protein families abbreviated "fps" (i.e. foot proteins): fp1, fp2, fp3, and fp4, all of which contain the post-translationally modified amino acid 3,4-dihydroxyphenylalanine (Dopa) (Diamond 1993; Papov et al. 1995). Mefp1 is a large protein, ~110 kDa, with more than 75 tandem repeats of a decapeptide that contains *trans*-4-hydroxyproline and *trans*-2,3-*cis*-3,4-dihydroxyproline as well as Dopa (Waite 1983; Taylor et al. 1994). Mefp2 also has a repetitive primary structure: 11 tandem repeats of an epidermal growth factor motif (Inoue et al. 1996) with Dopa confined to the N- and C-termini (Rzepecki et al. 1992). Mefp3 comprises a family of small, nonrepetitive proteins ranging from 5 to 7 kDa in mass and with 4-hydroxyarginine as well as Dopa (Papov et al. 1995; Inoue et al. 1996). A fourth protein, Mefp4, has a mass of 80 kDa and may consist of tyrosine-rich octapeptide repeats (Weaver 1998). The

catalog of byssal plaque proteins is still incomplete, and little to nothing is known about the function and location of each plaque protein. There is some evidence that Mefp1 is not limited to the plaques, but forms a protective sealant around all exposed portions of the byssus (Rzepecki et al. 1992). Mefp2, 3, and 4 are specific to the plaques where fp2 is the most abundant protein, but distribution patterns within the plaque are not known.

Recent use of MALDI-TOF mass spectrometry has indicated for the first time that proteins the size of Mefp3 are present on the surface of the adhesive side of the plaque (Ross et al. 1995). Given that previous studies have shown that the ultraviolet lasers used in MALDI-TOF penetrate no deeper than 300 nm into a variety of porous membranes (Strupat et al. 1994), this is a revealing result. Byssal plaques are highly porous materials (Tamarin et al. 1976), and detection of Mefp3s within 300 nm of the interface has important implications for adhesion. Consistent with its localization in the plaque portion of the byssus, fp3 was shown to be transcribed in the phenol gland at the tip of the foot by in situ hybridization in a closely related species, *Mytilus galloprovincialis* (Inoue et al. 1996). Previous studies of fp3 in both *M. edulis* and *M. galloprovincialis* have pooled samples of feet from multiple mussels, and it has therefore been unclear whether detection of multiple forms reflects allelic variation (with two variants per individual mussel) or is due to the simultaneous production of nonallelic fp3 variants by individual mussels.

The present study probed fp3 distribution at the plaque interface using a laser desorption method, and determined the extent of fp3 polymorphism in individual *Mytilus edulis*.

Materials and methods

Mussels

Mytilus edulis L. used in these experiments were collected from submerged pilings at the mouth of the Broadkill River and from the fishing pier at Cape Henlopen State Park in Delaware, USA. Mussels (2 to 4 cm) were transferred to aquaria containing stationary or running sea water (30‰) at 8 to 15 °C. Those used for laser desorption experiments were fed weekly with *Isochrysis galbana* cultures.

Plaque collection

Getting mussels to deposit adhesive plaques exclusively onto glass or acrylic plates (Plexiglas) was configured using specially cut plates ($10 \times 10 \times 0.3$ cm) which were tightly wrapped by two to three turns of nylon fishing line (diameter 1 mm). Mussels were tethered to these wrappings using 6 cm long rubber bands that were loop-tied at a distance two-thirds from one end. The larger loop was slip-knotted through the fishing line, while the smaller loop was dorsoventrally wrapped around the mussel. The mussel/plate combinations were then draped from a clothes line suspended across the top of the aquarium using small noncorroding spring clips. Mussels began exploring the plate surface within 1 h of immersion, and plaque deposition commenced shortly afterwards. To avoid extensive protein cross-linking in the byssal plaques the following conditions were required: low sea water temperatures (8 °C), minimal water agitation, and plaque harvest within 24 h of deposition. Plaques were harvested under the stereomicroscope by cutting the byssal threads a few millimeters above the plaque and, then, by shearing the plaques from the plate surface with a new, single-edge razor blade. Each collected plaque was washed with two changes of 10 ml Q-water.

Laser desorption mass spectrometry

Mass spectrometric analysis was done by matrix-assisted laser desorption ionization with time-of-flight (MALDI-TOF) using a PerSeptive Voyager instrument with delayed extraction (Perkin-Elmer, Framingham, Massachusetts). Byssal adhesive plaques with "top" or "bottom" sides facing up were mounted onto a gold specimen plate using double-sided tape (3 M), coated with matrix ($1\ \mu l \times 2$) and air-dried. Matrix was sinapinic acid (Sigma; 10 mg ml^{-1}) dissolved in 0.1% (v/v) trifluoroacetic acid in a 50:50 mixture of Q-water and HPLC-grade acetonitrile. The samples were evacuated to at least 5×10^{-7} mmHg in the vacuum chamber of the spectrometer and irradiated at 337 nm with a nitrogen laser at a power setting of 2100 to 2300. Counts of molecules desorbed, ionized, and detected in the positive ion mode were averaged over 255 accumulated scans. Other conditions included an accelerating voltage of 20 kV, grid voltage at 94%, a guide wire voltage of 0.02% and low mass gate of 800. Although MALDI-TOF has a mass accuracy that is generally better than 0.01%, some runs included a monoisotopic standard (r-hirudin, 6964.5 $[M+H]^+$) for internal calibration.

cDNA

Total RNA was isolated from the foot of individual mussels by the modified method of Chomczynski and Sacchi (1987) (Eddington 1996). Initially, thermal rapid amplification of cDNA ends (RACE) in the 3′ and 5′ directions was performed according to the method of Frohman (1993) with total RNA as the template. Nested primers were not used due to the variable nature of the protein and the small size of the gene. A poly-T primer comprised of 5′-$GGGGGGGT_{17}$-3′ was used for both the reverse transcription of the 3′ RACE and the subsequent PCR. Gene-specific primers, Mefp3-1 and Mefp3-1as (Table 1), were initially designed from a conserved portion of *Mytilus galloprovincialis* fp3 sequence (Inoue et al. 1996) that was known to code for the same amino acids in the *M. edulis* version of fp3 (Papov et al. 1995) (Fig. 1). An additional primer (Mefp3-9) that would enable the complete coding sequence to be obtained with 3′ RACE was designed from the preliminary 5′ end partial cDNA sequences obtained. All primers were obtained from Operon Technologies (Alameda, California). Control reactions with all components except the reverse transcriptase were done to determine if the subsequent PCR products were due to genomic contamination of the RNA preparations. The PCR products were then cloned into the plasmid pBluescript (Stratagene, La Jolla, California) which was cut with *Eco*RV, and a T-overhang was added by incubation with *Taq* polymerase in the presence of 2 m*M* dTTP (Hadjeb and Berkowitz 1996). Plasmids were transformed into competent *Escherichia coli* strain XL1-Blue-MRF' (Stratagene). Transformants were plated on LB plates with 100 μg ml^{-1} carbenicillin (Sigma Chemical Co., St. Louis, Missouri). Colonies were initially screened for inserts in the presence of isopropyl-thiogalactopyranoside

Table 1 Gene-specific primers used in this study

Name	Sequence
Mefp3-1	5′-GGYCCTCCAAGRCGYTAC
Mefp3-1as	5′-GTARCGYCTTGGAGGRCC
Mefp3-9	5′-AACAATGAATAATWYCAGTGTCTCAGTG

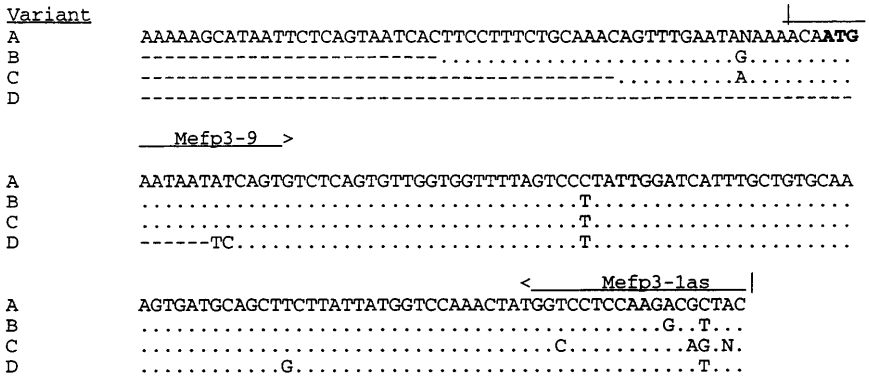

Fig. 1 *Mytilus edulis*. Preliminary 5′ end sequences from one mussel obtained by 5′ RACE using an antisense primer (Mefp3-1as) designed from the conserved region of previously published sequences of *M. galloprovincialis* (Inoue et al. 1996) which also codes for the same sequence of amino acids in *M. edulis* (Papov et al. 1995). Primer locations (for the antisense primer to obtain these sequences) and the subsequent 3′ RACE primer (Mefp3-9) (for obtaining the complete coding region) are as indicated above the sequences. The primer Mefp3-1 was in the same location as Mefp3-1as, but in the opposite direction. The start codon (ATG) is in bold

(IPTG) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal). Plasmids were grown up in Terrific Broth with 50 μg ml^{-1} ampicillin, and the DNA isolated using an alkaline lysis miniprep procedure. Inserts were sequenced using the FS dye-terminator sequencing kit (Applied Biosystems, Inc., Foster City, California) and universal primers (KS and SK) for the vector and sequences run on an ABI Model 373 sequencer. Analysis of the resulting traces was done with Sequencher (Gene Codes Corp., Ann Arbor, Michigan).

Results

Laser desorption is a sensitive tool for the analysis of biomolecular materials such as byssal attachment plaques. Proteins were definitely desorbed from the "bottom" or adhesive side of attachment plaques by laser irradiation. Using three different matrices, we were unable to detect anything above 20 kDa in the overall range of 1 to 100 kDa. The most abundant desorbed proteins had monoprotonated masses of 5 to 7 kDa and exhibited the same pattern of clustered peaks separated by 16 Da which was evident in Mefp3 variants isolated from the foot (Papov et al. 1995).

At least four variants were desorbed from plaques deposited on glass and plexiglass, all of these showed incremental clustering $\Delta m = 16$ Da, e.g. note especially peaks clustered at $[M+H]^+$ 5757.85 (Fig. 2A) and 5764.74 (Fig. 2B). The two surfaces had three proteins in common: (1) 5304.8 (glass) and 5312.5 (plexiglass), (2) 5470.12 (glass) and 5440.95 (plexiglass), and (3) the clusters between about 5750 and 5860 on both surfaces. Although these $[M+H]^+$ values were not identical, they corresponded to the similar variants modified to different degrees by addition of oxygen (16 Da) and dehydrogenation (−2 Da probably per Dopa residue oxidized to Dopaquinone). Unique fp3s appear to be $[M+H]^+$ 5629.7 (glass) and 4996.36 and 5506.2 (both on plexiglass). If the unique peak clusters, in fact, represent different variants of Mefp3, then multiple variants observed earlier in preparations from pooled mussels (Papov et al. 1995) could not have arisen through multiple alleles alone. Instead, each individual mussel foot may express many different Mefp3 variants. To test this hypothesis, we turned to RT-PCR of cDNA from a single mussel using gene-specific primers.

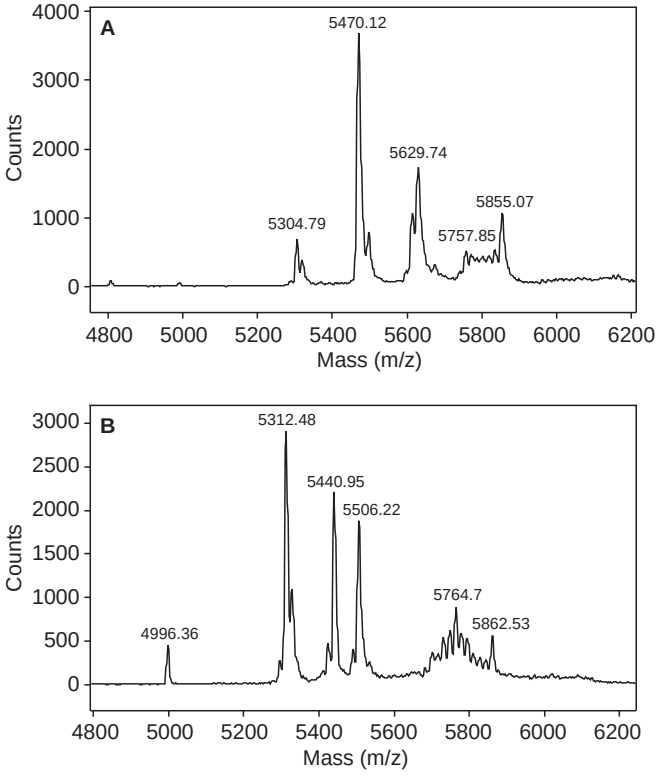

Fig. 2 *Mytilus edulis*. MALDI-TOF mass spectrometric analyses of plaques that were deposited by mussels on glass or plexiglass. **A** Proteins desorbed from the underside, i.e. adhesive side, of a plaque deposited on glass. *m/z* clusters at: **5304.79** + 5320.51; **5470.12** + 5486.02; **5629.74** + 5612.41; **5757.85** + 5774.11 + 5790.23 + 5806.31 + 5824.04 + 5839.5 + **5855.07**. **B** Proteins desorbed from the underside of a plaque deposited on plexiglass. *m/z*: 5295.36 + **5312.48** + 5327.5; 5424.27 + **5440.95**; 5490.42 + **5506.22** + 5522; 5732.64 + 5749.16 + **5764.7** + 5779.89 + 5796 + 5812 + 5828 + 5844; **5862.53**. Bold *m/z* values denote peaks identified in Fig. 2A and B. Related members vary by $\Delta m \approx 16$ Da increments

Initially the 3′ and 5′ RACE of one mussel yielded six 3′ products (not shown) and four 5′ products (Fig. 1) from a central sequence (5′-GGYCCTCCAAGRCGY-TAC-3′) used as a primer in both the sense and antisense directions. The 5′ RACE sequence information was used to synthesize another primer to encompass the whole coding region with 3′ RACE. Twenty-two different cDNA sequences of the complete coding region were obtained from another individual mussel with 3′ RACE. No PCR products were seen in the control reactions in which there was no reverse transcriptase added to the first strand reaction, indicating that the PCR products were not the result of genomic DNA contamination in the RNA preparation. Three of these cDNA transcripts coded for the same amino acid sequence (i.e. mutations were silent), thereby reducing the variants to 20 protein sequences (Fig. 3). Although there was variation throughout the protein, the 5′ end appeared to be more conserved than the 3′ end.

Each transcript contained a signal peptide sequence based on the description of von Heijne (1985) and Nielson et al. (1997) and predicted with the "SignalP" program based on Nielson et al. (1997). The signal peptide comprised the first 24 amino acids in 19 cases and the first 25 in the remaining protein sequence (No. 16) (Fig. 3). Although basic residues were absent from the N-termini, a hydrophobic core made up of Val, Leu, Ala, Ile, Gly, Ser, and Phe was evident. This hydrophobic core (beginning with Val-8) was 13 amino acids in length in 19 cases and 14 amino acids long in the remaining sequence (Variant 16). The C-termini of the signal peptides contained typical polar residues. When the signal peptides are removed, the resulting "mature" proteins varied in length between 44 and 54 amino acids. Molecular masses of the mature proteins range from 5303 to 6387 Da. Calculated isoelectric points are between 10.1 and 10.9. Gly is the most common amino acid, followed closely by Tyr and Arg. These latter two amino acids were shown to be modified co- or post-translationally to Dopa and 4-hydroxyarginine, respectively, in the variant fp3F, and the same trend was apparent in the compositions of the other variants (Papov et al. 1995). Papov et al. (1995) also demonstrated that the modification of fp3 variants isolated from *Mytilus edulis* was incomplete, and more so in the case of Arg than Tyr. If hydroxylation of Tyr and Arg were complete, then the molecular mass range of the variants would be shifted up from a low of 5607 (e.g. Variant 7) to a high of 6707 Da (Variant 2) (Fig. 3). Asn is also common in fp3 (between 5 and 7 residues variant^{-1}). The remaining amino acids consist of varying but small amounts of Ala, Pro, Ser, Trp, Thr, Asp, Lys, and Ile. His (Variant 10) and Glu (No. 19) occurred but once in the sequences of the mature variants. The C-terminal and, often, the penultimate residues were aromatic amino acids.

Discussion

The MALDI-TOF analysis of the adhesive plaque surfaces suggests (1) that Mefp3s are the only consistently desorbed proteins; (2) that each mussel plaque contains more than one variant of Mefp3 with multiple hydroxylation states ($\Delta m = 16$ Da); and (3) perhaps, that the variants selected for deposition are determined by the surface being adhered to. The first conclusion is well-supported given the accuracy of MALDI-TOF and the

Fig. 3 *Mytilus edulis*. Amino acid sequences of the 20 protein variants obtained in this study from one individual mussel. The proposed signal peptide and mature protein regions are indicated. Amino acids that are the same as in Variant 1 are indicated by ".". Alignment gaps are indicated by "–". Mature protein masses represent the calculated masses of the mature protein without any hydroxylation (minimum) and fully hydroxylated (maximum)

	Signal peptide | Mature protein	Mature protein Mass Min	Max
#1	MNNFSVS-VLVALVLIGSFAVQSDAADYYGPNYGPPRRYGGGNYNRYNRYGRR---YGGYKGWNNGWNRG--RRGKYW-	5831	6135
#2	...I.....A.....S.....S.....GGKGG.....R..SWG.R..NY	6387	6707
#3	.....V.....E....A.....--..G....A.G.....NR...R.....WRG.R..Y.	5773	6093
#4	.....S.....I...S...R.....GGKGG.....G..SWG.R..NY	6287	6591
#5	...S.....S.....S...R.....GGKGG.....R..SWG.R..NY	6387	6707
#6	.....EV.....E....A.....--..G....A.G.....NR...R.....WRG.R..Y.	5772	6092
#7	...S.....V.....R....A.....--.....A.G.....----.T.....WRG.R..Y.	5303	5607
#8	.....LV.....A.....--.....A.G.....----.T.....WRG.R..Y.	5303	5607
#9	...S.....LV.....A.....--.....A.G.....----.T.....WRG.R..Y.	5303	5607
#10	...S.....G.....H...Y.....K.....	5868	6172
#11	...T.....	5831	6135
#12	...S.....Y.....K.....	5894	6214
#13	...I.....	5831	6135
#14	...S.....V.....R....A.....--.....A.G.....----.T.....WRG.R..Y.	5303	5607
#15	...S.G.....A....S.....S...R.....GGKGG.....R..SWG.R..NY	6387	6707
#16	.....VL..V.....E....A.....--..G....A.G.....NR...R.....WRG.R..Y.	5772	6092
#17	...T.....Y.....W.....K..--.....	6024	6344
#18	...S.....LV.....S..E....A.....--..G....A.G.....NR...R.....WRG.R..Y.	5772	6092
#19	...S.....Y.....I.....K..--..E....	5924	6228
#20	...T.....V.....E....A.....--..G....A.G.....NR...R.....WRG.R..Y.	5772	6092

masses and hydroxylation patterns detected. Similar results have been obtained with Mefp3s isolated from mussel foot tissue (Papov et al. 1995). Additional confirmation will require a detailed structural characterization. This conclusion does not exclude the presence of other proteins that may not have been ionized or desorbed under the conditions used. However, since Mefp1, 2, and 4 yield readily to MALDI-TOF analysis in purified or combined form (Waite unpublished results), their presence at the interface (unless cross-linked) does not seem likely. The second conclusion is strengthened by the molecular evidence for fp3 variant expression in individual mussels. The existence of multiple fp3 variants was evident from earlier studies (Papov et al. 1995; Inoue et al. 1996), but the extent of individual variation was unclear since protein and mRNA were both extracted from pooled mussel feet. The third observation is the most exciting in that it suggests that mussels may tailor their adhesion to the properties of a detected surface. It remains speculative until large numbers of plaques on different surfaces can be examined by MALDI-TOF. The present data, however, predict at least three types of chemical adaptation to surfaces by mussels: (1) different fp3 variants for different surfaces, (2) different proportions of a limited set of fp3 variants on different surfaces, and (3) different hydroxylation patterns in fp3s.

The present results indicate extensive polymorphism of Mefp3 mRNA in an individual mussel. Whether all varieties of this mRNA are equally and simultaneously translated into protein remains to be shown. The uncharacteristically short poly-A tail of mRNA of Mefp3 (K. Inoue personal communication) combined with the large number of mRNA variants (present study) and the handful of fp3 protein variants detected on the adhesive plaques suggests that the selection of secreted protein variants may be controlled at a translational level. This would result in faster production of appropriate variants by circumventing transcription and the processing of the mRNA. The extent of the polymorphism cannot be due to allelic variation alone. Additional variation could be due to alternative splicing and RNA editing (Benne 1993; Patton et al. 1997). This may include alternative splicing of both the 5′ and 3′ ends. Protein polymorphism based on multiple gene copies and alternative splicing is well documented in other systems, the adaptive function being increased output or functional modulation, respectively. Ocean pout use multiple gene copies to increase the output of antifreeze proteins (Hew et al. 1988), while multiple gene copies are also used by parasites to provide variation in the egg capsule proteins (Rice-Ficht et al. 1992; Waite and Rice-Ficht 1992) and in prokaryotes to provide antigenic variation (Palmer et al. 1994; Alleman et al. 1997). Squid use RNA editing to create different forms of a protein involved in potassium channels (Patton et al. 1997), and eight different isoantigens of BGP, a cellular adhesive molecule, derived by differential alternative splicing of a single transcriptional unit, provide functional diversity in adhesion (Barnett et al. 1993). Multiple copies of an Mefp3 gene may be necessary to express sufficient amounts of this protein for repeated thread formation – particularly during periods of high water flow. A mussel can produce a new byssal thread in 5 min or less using proteins stockpiled in the holocrine glands of the foot (Waite 1992). A sustained demand, however, would require high output gene expression. This would pertain to all byssal proteins, not just Mefp3.

The preliminary analyses of adhesive plaques on glass and plastic by MALDI-TOF mass spectrometry point to a possible functional adaptation or tailoring of protein to surface. The Mefp3-like variants desorbed from the plaques bonded to the two surfaces were not identical. There are two obvious but different pathways to form a range of fp3 variants (Fig. 2): (1) variants with different primary amino acid sequences can be synthesized by the mussel from the different gene copies, alternative splicing, or RNA editing; and (2) differential hydroxylation can occur as shown in Fig. 2. If both strategies worked concurrently then an enormous array of functionally different molecules would result. Both pathways may be responsive to cues from the surface being bonded.

As mature proteins, Variants 1, 11, and 13 (Fig. 3) are essentially identical and correspond to Variant F, a protein whose sequence was directly determined in an earlier study (Papov et al. 1995). These three variants differ only in the signal peptide portion of the protein. The calculated isoelectric point for the mature protein is 10.4. If all of the Tyr and half (five of nine residues) of the Arg residues are modified to Dopa and 4-hydroxyarginine, respectively, as indicated from the studies of the protein (Papov et al. 1995), then the molecular mass of the mature protein would be 6071 kDa. If all Tyr and Arg residues are hydroxylated, then the calculated molecular mass is 6135.3 Da, which is in agreement with the largest mass (i.e. fully hydroxylated) obtained from the MALDI-TOF mass spectra of Variant F (Papov et al. 1995).

The modified amino acid, L-Dopa, is believed to be involved in cross-linking (Waite 1985; Burzio 1996) and surface bonding (Waite et al. 1992). Many, but not all, of the Tyr residues (to be converted to Dopa) appear to be conserved in the sequences. This may reflect their importance to the adhesive function of this protein. The Arg residues are not as conserved as Tyr, and it is unknown if those that are conserved are more likely to be post-translationally hydroxylated or involved in the cross-linking.

The frequency of amino acid substitutions in the N- and C-terminal halves of fp3 (Fig. 3) is noticeably asymmetric. Compared with Sequence 1, which corresponds to Mefp3F of Papov et al. (1995), the N-terminal half of the mature sequence (A-1 to Y-25) contains an average of about two substitutions per 25 residues in other variants; the C-terminal half, in contrast, has more than eight. This apparent asymmetry is intriguing and worthy of modest speculation. If fp3 variants collectively serve as something akin to an ''adhesive primer'' at the plaque–substratum interface, then molecular asymmetry

in fp3s may serve to couple the structurally dissimilar adherends. One adherend, the plaque, presumably has a biochemistry that is both familiar to the mussel and independent of the other adherend, i.e. the substratum. Thus, the more conserved N-terminal half may be anchored in the plaque matrix, whereas the more variable C-terminal half might be associated with the substratum. Future work should explore this possibility.

Acknowledgements The authors would like to thank Drs. J.H. McDonald and A. Hunt for their assistance with the DNA sequencing and Chengjun Sun for assistance in setting up the mussels for MALDI-TOF analysis. This work was funded by PHS Grant DE 10042 and ONR Grant N0001496-0812.

References

Alleman AR, Palmer GH, McGuire TC, McElwain TF, Perryman LE, Barbet AF (1997) *Anaplasma marginale* major surface protein 3 is encoded by a polymorphic, multigene family. Infection Immunity 65: 156–163

Barnett TR, Drake L, Pickle W (1993) Human biliary glycoprotein gene: characterization of a family of novel alternatively spliced RNAs and their expressed proteins. Molec cell Biol 13: 1273–1282

Benne R (1993) RNA editing, an overview. In: Benne R (ed) RNA editing. The alteration of protein coding sequences of RNA. Ellis Horwood, New York

Burzio LA (1996) Catechol oxidases associated with byssus formation in the blue mussel *Mytilus edulis*. MS thesis, University of Delaware, Newark

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. Analyt Biochem 162: 156–159

Diamond TV (1993) Dopa-containing components in the byssal plaque of *Mytilus edulis* L. MS thesis, University of Delaware, Newark

Eddington ND (1996) Partial oligonucleotide sequence of a mussel byssal precursor protein, *Dreissena polymorpha* foot protein 2. MS thesis, University of Delaware, Newark

Frohman MA (1993) Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: thermal RACE. Meth Enzym 218: 340–356

Hadjeb N, Berkowitz GA (1996) Preparation of T-overhang vectors with high PCR cloning efficiency. BioTechniques 20: 20–22

Hew CL, Wang N-C, Joshi S, Fletcher GL, Scott GK, Hayes PH, Buettner B, Davies PL (1988) Multiple genes provide the basis for antifreeze protein diversity and dosage in the ocean pout, *Macrozoarces americanus*. J biol Chem 263: 12049–12055

Inoue K, Takeuchi Y, Miki D, Odo S, Harayama S, Waite JH (1996) Cloning, sequencing and sites of expression of genes for the hydroxyarginine-containing adhesive-plaque protein of the mussel *Mytilus galloprovincialis*. Eur J Biochem 239: 172–176

Mahéo R (1970) Étude de la pose et de l'activité de sécrétion du byssus de *Mytilus edulis* L. Cah Biol mar 11: 475–483

Nielson H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10: 1–6

Palmer GH, Eid G, Barbet AF, McGuire TC, McElwain TF (1994) The immunoprotective *Anaplasma marginale* major surface protein 2 is encoded by a polymorphic multigene family. Infection Immunity 62: 3808–3816

Papov VV, Diamond TV, Biemann K, Waite JH (1995) Hydroxyarginine-containing polyphenolic proteins in the adhesive plaques of the marine mussel *Mytilus edulis*. J biol Chem 270: 20183–20192

Patton DE, Silva T, Bezanilla F (1997) RNA editing generates a diverse array of transcripts encoding squid Kv2 K^+ channels with altered functional properties. Neuron 19: 711–722

Rice-Ficht AC, Dusek KA, Kochevar GJ, Waite JH (1992) Eggshell precursor proteins of *Fasciola hepatica*. I. Structure and expression of vitelline protein B. Molec Biochem Parasit 54: 129–142

Ross MM, Taylor SW, Waite JH (1995) Mass spectrometry of marine biomolecules. Proc 43rd ASMS Conference on Mass Spectrometry and Allied Topics. American Society of Mass Spectrometry, Atlanta, Georgia, p 298

Rzepecki LM, Hansen KM, Waite JH (1992) Characterization of a cystine-rich polyphenolic protein family from the blue mussel *Mytilus edulis* L. Biol Bull mar biol Lab, Woods Hole 183: 123–137

Strupat K, Karas M, Hillenkamp F, Eckerskorn C, Lottspeich F (1994) Matrix-assisted laser desorption ionization mass spectrometry of proteins electroblotted after polyacrylamide gel electrophoresis. Analyt Chem 66: 464–470

Tamarin A, Lewis P, Askey J (1976) The structure and formation of the byssus attachment plaque in *Mytilus*. J Morph 149: 199–222

Taylor SW, Waite JH, Ross MM, Shabanowitz J, Hunt DF (1994) *Trans*-2, 3*cis*-3,4-dihydroxyproline, a new naturally occurring amino acid, is the sixth residue in the tandemly repeated consensus decapeptides of an adhesive protein from *Mytilus edulis*. J Am chem Soc 116: 10803–10804

von Heijne G (1985) Signal sequences: the limits of variation. J molec Biol 184: 99–105

Waite JH (1983) Evidence for a repeating 3,4-dihydroxyphenylalanine- and hydroxyproline-containing decapeptide in the adhesive protein of the mussel, *Mytilus edulis* L. J biol Chem 258: 2911–2915

Waite JH (1985) Catechol oxidase in the byssus of the common mussel, *Mytilus edulis* L. J mar biol Ass UK 65: 359–371

Waite JH (1992) The formation of mussel byssus: anatomy of a natural manufacturing process. In: Case ST (ed) Results and problems in cell differentiation. Vol. 19. Springer-Verlag, Berlin, pp 27–54

Waite JH, Jensen RA, Morse DE (1992) Cement precursor proteins of the reef-building polychaete *Phragmatopoma californica* (Fewkes). Biochemistry 31: 5733–5738

Waite JH, Qin XX, Coyne KJ (1998) The peculiar collagens of mussel byssus: a minireview. Matrix Biol 17: 93–106

Waite JH, Rice-Ficht AC (1992) Eggshell precursor proteins of *Fasciola hepatica*. II. Microheterogeneity in vitelline protein B. Molec Biochem Parasit 54: 143–152

Weaver JK (1998) Isolation, purification, and partial characterization of a mussel byssal precursor protein, *Mytilus edulis* foot protein 4. MS thesis, University of Delaware, Newark