

Linking Adhesive and Structural Proteins in the Attachment Plaque of *Mytilus californianus**

Received for publication, May 8, 2006, and in revised form, July 13, 2006. Published, JBC Papers in Press, July 14, 2006, DOI 10.1074/jbc.M604357200

Hua Zhao and J. Herbert Waite¹

From the Department of Molecular, Cellular & Developmental Biology and the Department of Chemistry & Biochemistry, Marine Science Institute, University of California, Santa Barbara, California 93106

The byssal attachment of California mussels *Mytilus californianus* provides secure adhesion in the presence of moisture, a feat that still eludes most synthetic polymers. Matrix-assisted laser desorption ionization mass spectrometry was used to probe the footprints of byssal attachment plaques on glass cover slips for adhesive proteins. Besides the abundant mcfp-3 protein family (Zhao, H., Robertson, N. B., Jewhurst, S. A., and Waite, J. H. (2006) *J. Biol. Chem.* 281, 11090–11096), two new proteins, mcfp-5 and mcfp-6, with masses of 8.9 kDa and 11.6 kDa, respectively, were identified in footprints, partially characterized and completely sequenced from a cDNA library. mcfp-5 resembles mcfp-3 in its basic pI and abundant 3,4-dihydroxyphenyl-L-alanine (Dopa; 30 mol %), but is distinct in two respects: it is more homogeneous in primary sequence and is polyphosphorylated. mcfp-6 is basic and contains a small amount of Dopa (<5 mol %). In contrast to mcfp-3 and -5, tyrosine prevails at 20 mol %, and cysteine is present at 11 mol %, one-third of which remains thiolate. Given the oxidative instability of Dopa and cysteine at pH 8.2 (seawater), we tested the hypothesis that thiols serve to scavenge dopaquinones by adduct formation. Plaque footprints were hydrolyzed and screened for cysteine dopaquinone adducts using phenylboronate affinity chromatography. 5-S-CysteinylDopa was detected at nearly 1 mol %. The results suggest that mcfp-6 may provide a cohesive link between the surface-coupling Dopa-rich proteins and the bulk of the plaque proteins.

Mussels inhabit wind and wave swept rocky seashores. Such habitats are deathtraps for exposed organisms lacking a secure attachment. Accordingly, mussels have evolved a robust holdfast known as the byssus, which is essentially a specialized 4–5-cm long bundle of collagenous fibers that is proximally rooted in the mussel and distally bonded to foreign surfaces underwater by flattened attachment plaques (1).

Given that engineering durable adhesive bonds between minerals and organic polymers in the presence of moisture remains a serious technological challenge, fundamental insights into the mechanism of holdfast adhesion in mussels and other sessile marine organisms represent a potential data base of bio-inspired solutions to the moisture problem (1). One popular technique for improving “wet” adhesion on siliceous substrates involves the application of surface-coupling agents or adhesion promoters (2). Organosilanes are the best known synthetic adhesion promoters and typically designed with specific moieties for silica ligation at one end and reactivity toward the organic polymer at the other (2).

The use of surface-coupling agents to promote adhesion resonates with the adhesive biochemistry of byssal plaques made by mussels. A recent investigation of plaque footprints in *Mytilus californianus* has revealed a family of protein variants (mcfp-3) with a 3,4-dihydroxyphenyl-L-alanine (Dopa)² content that approaches 25 mol % (3, 4). With its many Dopa residues, mcfp-3 has been compared with a multifunctional surface-coupling agent (4). By forming bidentate coordination complexes with metal centers in metal oxide and mineral surfaces, the Dopa/surface interaction is stronger than any noncovalent interaction (5) and, unlike noncovalent interactions, not diminished by the large dielectric constant of water (6). If multiple Dopa side chains represent the surface-ligating moiety in mcfp-3, then there must also be another moiety that is specialized for reactivity with other plaque proteins.

The present investigation was undertaken to determine whether there is a sidedness to the reactivity of mussel adhesive proteins with surfaces. In other words, if Dopa moieties of adsorbed proteins interact with the mineral surface, what interactions define their binding to other proteins in the adhesive plaque? Of the two new proteins detected in the footprints (mcfp-5 and mcfp-6), mcfp-6 is a thiol-rich protein that may mediate coupling of the surface proteins with those in the plaque by cysteinylDopa cross-links.

MATERIALS AND METHODS

MALDI-TOF Mass Spectrometry—MALDI-TOF examination of footprints and purified proteins was performed using a Voyager DE mass spectrometer (Applied Biosystems, Foster City, CA). The MALDI matrix was prepared by dissolving sina-

* This research was supported by in part by grants (to J. H. W.) from NASA (University Research Engineering and Technology Institute on Bio-Inspired Materials under Award No. NCC-1-02037) and the National Institutes of Health (DE015415). 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) DQ351537, DQ351538, DQ351539, and DQ444853.

¹ To whom correspondence should be addressed: Marine Science Institute, University of California, Santa Barbara, CA 93106. Tel.: 805-893-2817; Fax: 805-893-7998; E-mail: waite@lifesci.ucsb.edu.

² The abbreviations used are: Dopa, 3,4-dihydroxyphenylalanine; Mfps, *Mytilus* foot proteins; mcfp-5 and mcfp-6, *M. californianus* foot protein 5 and 6; NBT, nitroblue tetrazolium; MALDI-TOF, matrix-assisted laser desorption and ionization with time-of-flight; GdnCl, guanidine hydrochloride; RACE, rapid amplification of cDNA ends.

pinic acid (10 mg/ml) in 50% acetonitrile. Footprints of byssal plaques were screened for proteins by MALDI-TOF mass spectrometry. The technique as currently practiced involves the following: glass cover slips were collected shortly after plaque deposition. The cover slips were cleaned of slime and debris by washing thoroughly with Q-water after which the plaques were scraped off using a clean single-edge razor. The glass surface with the footprint residue was dried and mounted onto a MALDI sample plate with double stick tape. Two microliters of matrix mixture (30% sinapinic acid) were applied to the footprint and air-dried before subjecting to pulsed laser irradiation (4). Singly and doubly charged peaks for apomyoglobin from horse (16,952.56, and 8476.78, respectively, for average masses) were used as internal calibrants.

The purified footprint proteins derived thereof were dissolved in this matrix solution to give a final concentration between 1 and 10 pmol/ μ l. About 1 μ l of this solution was applied to the target plate and allowed to evaporate. The sample spots were irradiated using an N₂ laser (LSI, Inc., Cambridge, MA) with a wavelength of 337 nm, pulse width of 8 ns and operated at a repetition rate of 5 Hz. MALDI ionization generates protonated singly and doubly charged ions for the footprint proteins, which were accelerated using either 20 or 25 kV accelerating voltage.

Protein Isolation from Adhesive Plaques—*M. californianus* were collected locally from Goleta pier at Santa Barbara, CA and immediately transferred to shallow tanks in the laboratory and maintained with running sea water. To collect the plaques for protein isolation, mussels were tethered onto plastic plates. After a couple of hours, mussels started deposition of plaques on the plastic plate. New plaques deposited within a 24-h period were delaminated using a clean single edge razor blade and rinsed extensively with Milli-Q water to remove any salts. Collected plaques either were used immediately or stored at -80°C for future protein isolation.

Once 2 g of plaques were accumulated, they were homogenized on ice with 20 ml of 5% acetic acid and 8 M urea with protease inhibitors using a small hand-held tissue grinder (Kontes, Vineland, NJ). Supernatant was harvested by centrifugation for 40 min at 20,000 $\times g$ and 4 $^{\circ}\text{C}$ using SS-34 rotor. Ammonium sulfate was then added to give a final concentration of 20% for protein precipitation. After centrifugation, the supernatant was collected and dialyzed against 4 liters of Milli-Q water overnight at 4 $^{\circ}\text{C}$. Dialysis resulted in mcfp-3 precipitation, which was removed by centrifugation. The resulting supernatant was dialyzed against 5% acetic acid overnight and concentrated by ultrafiltration (Amicon) to 1–2 ml. For each run, only 200 μ l of the concentrated crude extract subjected to gel filtration on a Shodex-803 column (5 μ m, 8 \times 300 mm). The column was equilibrated and eluted with 5% acetic acid in 0.2% trifluoroacetic acid, and monitored at 280 nm (7).

When purifying mcfp-5 from the adhesive plaques, around 1 gram of the adhesive plaques collected was homogenized in 10 ml of 5% acetic acid buffered with 4 M GdnCl. Supernatant was harvested by centrifugation and filtered through a 0.22- μ m microfilter from Millipore. The resulting supernatant was fractionated by reverse phase C8 HPLC using a 260 \times 7 mm RP-300

Aquapore (Applied Biosciences) column eluted with a linear gradient of aqueous acetonitrile.

Protein Isolation from the Mussel Feet—The phenol gland in the mussel foot where the adhesive precursor proteins are produced and stockpiled was also used for protein isolation. Live mussels were shucked, following which the foot was carefully dissected, arrayed on glass plates, and frozen at -80°C . In preparation for protein extraction, the mussel feet were removed from the freezer and partially thawed. The outer, pigmented epithelium in mussel feet was flayed away with a scalpel. Underlying phenol glands were easily visualized from the ventral side and dissected. Dissected phenol glands were homogenized to a puree on ice in 5% acetic acid with protease inhibitors (10 μ M leupeptin and pepstatin) and centrifuged at 20,000 $\times g$ and 4 $^{\circ}\text{C}$. The pellet (P1) was saved for isolation of mcfp-5, while supernatant (S1) containing mcfp-1, -2, and -6 was acidified with 70% perchloric acid to a final concentration of 1.5%. After centrifugation at 20,000 $\times g$, pellet (P2) was discarded and supernatant (S2) was decanted into a small beaker. Ammonium sulfate was then slowly added to a final concentration of 20% (w/v). The mixture was stirred for 40 min at room temperature and centrifuged at 20,000 $\times g$ and 4 $^{\circ}\text{C}$ for 30 min, and the resultant supernatant (S3) was collected and dialyzed against 4 liters of 5% acetic acid overnight using dialysis tubing with a molecular weight cutoff 1000 (Spectrum Industries, Los Angeles) with two changes of dialysis buffer, then freeze-dried. The lyophilized crude extract was resuspended in 2 ml of 5% acetic acid and run on a Shodex-803 column (5 μ m, 8 \times 300 mm), which was equilibrated and eluted with 5% acetic acid in 0.2% trifluoroacetic acid. Eluant was monitored at 280 nm. Fractions containing pure mcfp-6 were pooled and desalted by reversed phase C8 HPLC (260 \times 7 mm RP-300 Aquapore, Applied Biosciences) column, which was eluted with a linear gradient of aqueous acetonitrile. Eluant was monitored continuously at 220 and 280 nm, and collected 1-ml fractions were assayed by amino acid analysis and electrophoresis following freeze-drying.

To isolate mcfp-5, the pellet (P1) was homogenized with 5% acetic acid and 8 M urea. After centrifugation, supernatant was discarded and pellet was extracted again with 5% acetic acid containing 4 M GdnCl (8). The resulting supernatant was harvested by centrifugation and dialyzed against 4 liters of Milli-Q H₂O. Dialysis resulted in proteins precipitated out that contained crude mcfp-5 and was harvested by centrifugation followed by redissolving in 2 ml of 5% acetic acid and 8 M urea. Because it was partially dissolved, the resultant crude extract was filtered through a 0.22- μ m microfilter before loading onto a C8 column (260 \times 7 mm RP-300 Aquapore, Applied Biosciences) by reverse phases HPLC eluted with a linear gradient of aqueous acetonitrile.

Electrophoresis—Routine electrophoresis was done on polyacrylamide gels (7.5% acrylamide and 0.2% *N,N'*-methylenebisacrylamide) containing 5% acetic acid and 8 M urea (9). After electrophoresis, gels were stained with Coomassie Blue R-250 (Serva Fine Chemicals) and for Dopa-containing proteins by a redox-cycling method with nitroblue tetrazolium (NBT) in 2 M glycinate buffer (10). To estimate the apparent

Adhesive Footprint Proteins in Mussels

molecular weight, purified footprint proteins were run on 15% SDS-PAGE.

Amino Acid Analysis and Sequencing—Purified protein was hydrolyzed in 6 M HCl with 5% phenol *in vacuo* at 110 °C for 24 h. The hydrolysate was evaporated at 50 °C under vacuum and to dryness with a small volume of Milli-Q water and followed by methanol. Amino acid analysis was performed according to conditions described earlier with Beckman System 6300 Auto Analyzer (11). *O*-Phosphoserine was identified by amino acid analysis and estimated after correcting for losses because of hydrolysis (12). The N-terminal sequence of purified footprint proteins was determined by automated Edman degradation on a Porton Instruments Microsequencer (Model 2090, Porton, CA).

Molecular Cloning—Total RNA was extracted from the phenol gland in *M. californianus* foot tip using the RNase Plant Mini kit from Qiagen (Valencia, CA). Briefly, one freshly dissected foot tip was used, and Qiagen's protocols were followed after initial tissue disruption under liquid nitrogen with a mortar and pestle. Following that, mRNA was purified from total RNA with the Oligotex mRNA Mini kit from Qiagen. With purified mRNA, a cDNA library was constructed using the CloneMiner™ cDNA Library Construction Kit from Invitrogen. This cDNA library was a readily available source of cDNA.

Based on the known N-terminal sequence of mcfp-5 and -6 (Y**D*GY*SDGY*Y*P and GGGNY*RGY*, in which Y* denotes Dopa), degenerate primers forward, 5'-TAY GAY GGN TAY AGY/TCN GAY GGN TAY TAY CC-3', and 5'-GGN GGN GGN AAY TAY CGN/AGR GGN TA-3', were designed and coupled with an vector-encoded T7 universal primer to amplify the cDNA sequences of mature protein from cDNA library, respectively.

PCRs were carried out in 25 μ l of 1 \times Buffer B (Fisher) and 5 pmol of each primer, 5 μ mol of each dNTP, 1 μ l of first-strand reaction, and 2.5 units of TaqDNA polymerase (Fisher) for 32 cycles on a Robocycler (Stratagene). Each cycle consisted of 30 s at 94 °C, 30 s at 52 °C, and 40 s at 72 °C, with a final extension of 5 min. The PCR products were subjected to 1% agarose gel electrophoresis, purified, and cloned into a pCR4-TOPO vector with TOPO TA Cloning kit from Invitrogen and transformed into competent Top10 cells for amplification, purification, and sequencing.

For 5'-end information, the GeneRacer kit from Invitrogen was used to obtain sequence information from full-length transcripts by a 5'-RACE strategy. PCR was performed with gene-specific primers (antisense 5'-ATT TAA CAC GTG TGA CTA ACT GCT ACC-3' and 5'-AAT ATA GGC TCG CCT TTA GTA ACC-3', which reversely primes C terminus of mcfp-5 and -6, respectively) and a GeneRacer 5'-primer from Invitrogen (sense 5'-CGA CTG GAG CAC GAG GAC ACT GA-3').

Cysteine Modification—Cysteine in mcfp-6 was modified by alkylation with iodoacetate (13). In brief, purified mcfp-6 was reconstituted in 50 ml of 50 mM ammonia bicarbonate and alkylated with 15 μ l of 500 mM iodoacetate at room temperature for 40 min in the dark. Alkylated mcfp-6 was subjected to amino acid analysis post-hydrolysis. Cysteine was detected as carboxymethylcysteine, which ran at 9.5 min by amino acid analysis.

Isolation of CysteinylDopa Cross-links—*M. californianus* footprints on glass coverslips were hydrolyzed in 6 M HCl with 5% phenol *in vacuo* at 110 °C for 1.5 h. Hydrolysates were flash-evaporated to dryness at 50 °C under vacuum with a small volume of Milli-Q water followed by methanol. The flash-evaporated hydrolysate was taken up in 100 mM sodium phosphate buffer (pH 7.5), microcentrifuged for 10 min at 15,000 $\times g$ to remove insolubles, and applied to a pre-equilibrated phenylboronate column (Affi-Gel Boronate, Bio-Rad). Bound ligands on phenylboronate column were washed extensively with 100 mM phosphate buffer, and desalted by washing with 2.5 mM NH₄HCO₃ and Milli-Q water to facilitate subsequent amino acid analysis and electrospray ionization mass spectrometry described below. Fractions eluted with 5% acetic acid were lyophilized and subjected to a modified ninhydrin based amino acid analysis. Pure authentic 2- and 5-*S*-cysteinylDopa standards were donated by K. Wakamatsu (Fujita Health University) and detected by ion exchange amino acid analysis (see "Amino Acid Analysis and Sequencing") with elution times at 47 and 54 min, respectively (14, 15).

Cross-linking in Vitro—About 1 mg of mcfp-6 was resuspended in 400 μ l of 100 mM phosphate buffer (pH 7.5) and mixed with mushroom polyphenol oxidase (Sigma, T7755, $\geq 2,000$ units/mg) at an enzyme to protein at a ratio of 1:10 by weight (15, 16). The mixture was incubated at room temperature under constant stirring for 1 h. The reaction was stopped by adding 100 μ l of 6 M HCl. The resulting mixture was hydrolyzed in 6 M HCl with 5% phenol at 110 °C for 2 h *in vacuo*. The hydrolysate was flash-evaporated to dryness, redissolved in 1 ml of 50 mM phosphate buffer, and applied to phenylboronate column. Fractions eluted with 5% acetic acid were subjected to amino acid analysis for 5-*S*-cysteinylDopa. A control reaction was run using boiled enzyme (2 min).

Phospholipase A2 Activity Assay—Detection of the phospholipase activity was performed following nondenaturing electrophoresis using a 5.5% polyacrylamide gel containing a lecithin emulsion at 33 mg of lecithin per ml of gel (17, 18). Purified mcfp-6 was resuspended in 1 mM Tris-HCl, 10% glycerol, and 2 mM EDTA at pH 7.4, and then loaded onto the lecithin-containing gel. Electrophoresis was carried out at a constant current of 20 mA at 4 °C for 2.5 h with 2 mM EDTA, 5 mM Tris, and 38 mM glycine pH 8.9. Phospholipase activity in gels was visualized as described by Shier & Trotter (17). Briefly, gels were incubated overnight at 37 °C with gentle shaking in a bath containing 100 ml of 0.1 M Tris-HCl, 20 mM CaCl₂, and 5 μ g/ μ l melittin diluted by 10% by volume with a 0.12% aqueous solution of rhodamine 6G. Afterward, the gels were extensively washed with Milli-Q water to remove the excess dye. The positive control was cobra phospholipase A2 (Sigma P6139); negative controls were the molecular weight standards (Bio-Rad).

RESULTS

MALDI-TOF mass spectrometry was recently exploited for *in situ* detection of proteins in the adhesive footprints of *M. californianus* (3). In the present study we have extended this strategy to analyze footprint proteins in addition to mcfp-3. Mussels readily deposit adhesive plaques onto glass coverslips.

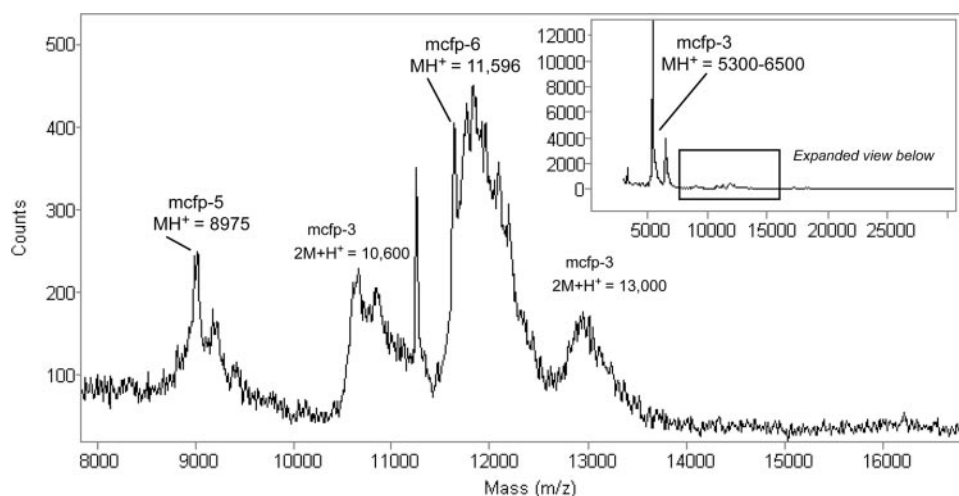


FIGURE 1. MALDI-TOF mass spectrum of the footprint of *M. californianus* on a glass coverslip. Inset shows the full mass and signal intensity range of the method used. The dominant peaks in the inset represent monoprotonated forms of mcfp-3 (Fp3 $[M+H]^+$). Monoprotonated dimers of Fp3 $[2M+H]^+$ appear in the expanded field. Footprint proteins, mcfp-5 and mcfp-6, are centered at m/z values of 8,974, and 11,674, respectively.

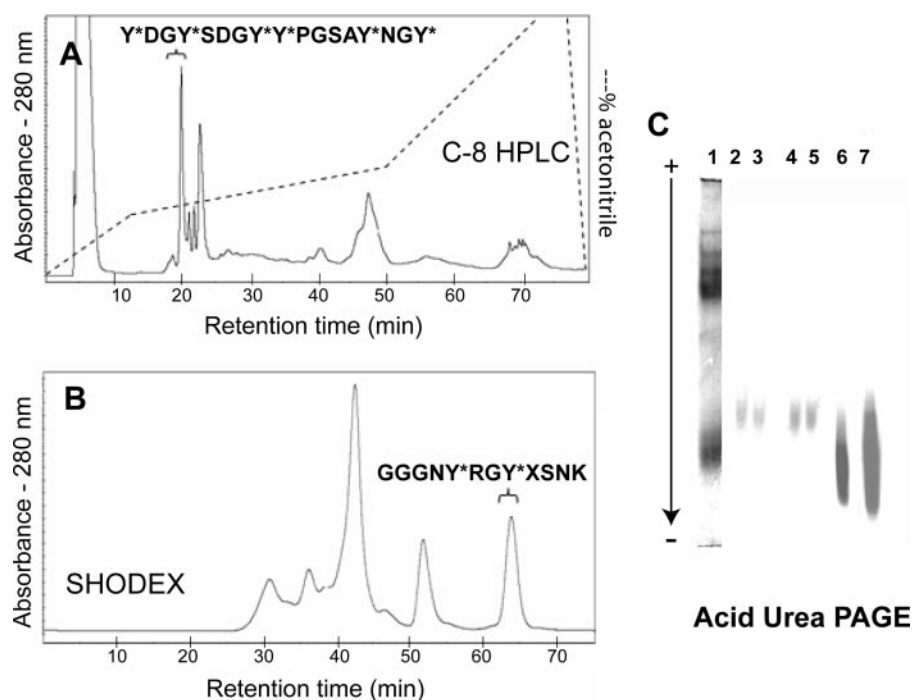


FIGURE 2. Purification of footprint proteins from plaques. A, C8 reverse phase HPLC of crude mcfp-5 obtained from plaques extracted with 5% acetic acid/4 M GdnCl (inserted sequences denote the N termini by Edman analysis). B, gel filtration chromatography on Shodex-803 of crude mcfp-6 obtained from plaques extracted with 5% acetic acid/8 M urea. Bracketed peaks denote pooled fractions. C, acetic acid/urea PAGE. Lane 1, unfractionated proteins contained in acetic acid/urea extracted plaques; lanes 2 and 3 contain single and double aliquots of the pooled fractions in A stained with CBR250; lanes 4 and 5 same as previous, but stained with NBT; lanes 6 and 7 were loaded with aliquots of the pooled fractions from B and stained with CBR250 and NBT, respectively.

Following removal of the adhesive plaques, the matrix-impregnated footprints were directly subjected to MALDI-TOF mass spectrometry. Although footprint spectra are dominated by the signal intensity of the mcfp-3 protein family characterized earlier, two new proteins were apparent by zooming in on the m/z range of 8000 to 13,000. The larger of the new proteins, mcfp-6, has a $[M+H]^+$ of 11.6 kDa that falls between the two mcfp-3 dimers $[2M+H]^+$ centered at 10.6 and 13 kDa, respectively (Fig. 1). The dense cluster of peaks centered at m/z 12,000 accu-

culated gradually after the peak at 11,600 formed, possibly suggesting matrix adducts. Another smaller protein, mcfp-5, was detected below these at 8.9 kDa (Fig. 1).

Footprint Protein Characterization from Plaques—To more fully characterize the new footprint proteins, it was first necessary to isolate them from the plaques and, for more material, from the phenol gland in the mussel foot. Collected plaques were extracted sequentially with acid-urea and acid-GdnCl, respectively, according to details outlined under “Materials and Methods.” After precipitation by ammonium sulfate (20% w/v), the acid-urea extract was dialyzed against dilute acetic acid, concentrated, and fractionated by gel filtration chromatography (Fig. 2B). Pure mcfp-6 fractions were identified by their mass on MALDI and pooled for biochemical characterization. In contrast, acid-GdnCl plaque extracts were enriched in mcfp-5, which could easily be purified by passage through C8 HPLC, eluting at 20 min (Fig. 2A). The composition of footprint proteins isolated from plaques was determined by amino acid analysis following hydrolysis (Table 1). Accordingly, nearly three-quarters of all residues in mcfp-5 are represented by glycine (24%), lysine (15%), and Dopa (30 mol %). In contrast, mcfp-6 has much less Dopa (<5 mol %), but elevated levels of glycine (14%), aspartate (14%), and tyrosine (20%) are noteworthy; cystine is present at 3.4% in mcfp-6. Both mcfp-5 and -6 contain some phospho-*O*-serine estimated between 3 and 5 mol %.

N-terminal sequences (Figs. 3 and 5) for mcfp-5 and -6 are distinct, but Dopa occurs in both. The mobilities and staining properties with redox cycling on polyacrylamide gels of plaque-derived mcfp-5 and -6 are illustrated in Fig. 2. Strong redox cycling with NBT is typical for Dopa/Topa-containing proteins (10, 11).

Footprint Protein Isolation from Feet—With mass, composition, N terminus, and electrophoretic mobility established for plaque-derived mcfp-5 and -6, protein counterparts from foot tissue could be prepared from considerably more complex extractions. mcfp-6 was readily extracted with dilute acetic acid from phenol glands. Most co-extracted contami-

TABLE 1

Amino acid composition of footprint proteins isolated from plaques and foot tissue of *M. californianus* given in mol % (Res/100 Res)

Predicted compositions are based on cDNA-deduced sequences. Tabulated values for the foot and plaque represent the means of three analyses with a S.D. that is about $\pm 15\%$ of each mean.

Residue	Mcfp-5			Mcfp-6		
	Foot	Plaque	Predicted ^a	Foot	Plaque	Predicted ^a
	mol %	mol %	mol %	mol %	mol %	mol %
pSer ^b	4.6	4.8	0	3.5	2.8	0
CM-Cys ^b	0	0	0	2.9	3.4	0
Asx	3.5	3.6	4.2	13.5	13.4	14.2
Thr	1.2	1.5	1.4	3.1	3.0	2.0
Ser	1.2	1.2	8.3	3.6	4.3	7.1
Glx	4.0	0.6	0	1.7	2.3	0
Pro	2.7	3.6	2.8	4.4	4.9	3.0
Gly	21.2	19.6	20.8	14.4	13.7	14.2
Ala	2.9	2.7	2.8	2.7	2.9	3.0
Cys/2	0	0	0	3.5	2.9	5.5
Val	0	0	0	1.1	1.6	1.0
Met	0	0	0	0	0	0
Ile	0	0	0	0.6	1.0	1.0
Leu	1.3	0.8	1.4	2.3	2.2	2.0
Dopa	30.4	30.4	0	3.7	3.2	0
Tyr	0.5	0.2	27.8	20.0	19.2	20.2
Phe	0	0	0	3.1	3.1	4.0
His	3.1	4.8	5.6	0	0	0
Hyls	2.6	0	0	0	0	0
Lys	18.8	19.8	20.8	9.2	9.8	9.2
Arg	3.0	3.1	4.2	6.6	6.5	8.1
Total	100	100	100	100	100	100

^a Compositions predicted from the mcfp-5 α and mcfp-6 β sequences, respectively.

^b Determined independently after modification or timed hydrolysis.

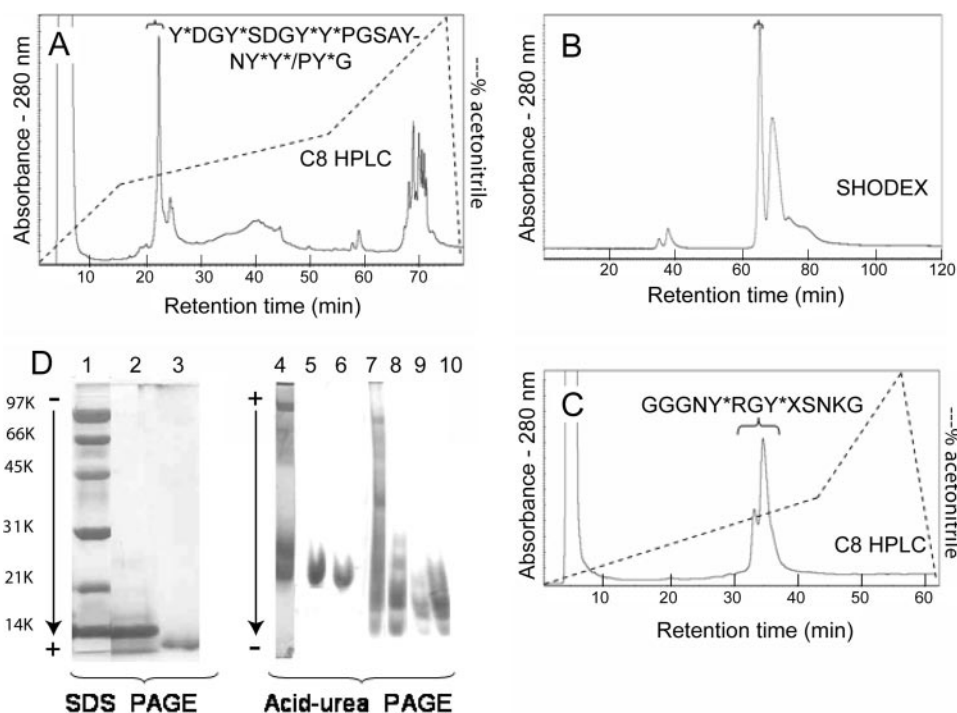


FIGURE 3. Purification of footprint proteins from the mussel foot. *A*, elution profile for mcfp-5 subjected to C8 reverse phase HPLC. *B*, gel filtration chromatography on Shodex-803 of acetic acid-extracted mussel feet for mcfp-6 purification, followed by further C8 reverse phase HPLC for more pure mcfp-6 preparation (*C*). (Inserted sequences denote the N termini by Edman analysis.) *Bracketed peaks* denote pooled fractions. *D*, polyacrylamide gel electrophoresis of key steps in protein purification from mussel foot. *Lanes 1–3* are SDS-PAGE and stained with CBR-250. *Lane 1* contains the low molecular weight markers (5 μ g each): phosphorylase B (97,400), bovine serum albumin (66,200), ovalbumin (45,000), carbonic anhydrase (31,000), soybean trypsin inhibitor (21,500), and lysozyme (14,400). *Lanes 2 and 3* contain purified mcfp-6 and -5, respectively. *Lanes 4–10* are acid-urea PAGE. *Lane 4* contains crude mcfp-5 obtained by GdnCl extraction. *Lanes 5 and 6* contain the pooled fractions following C8 reverse phase HPLC (*3-A*) and are stained with CBR-250 and NBT, respectively. *Lane 7* contains the crude extract of mcfp-6 following dialysis against 5% acetic acid. *Lane 8* illustrates the mcfp-6 containing fractions pooled following chromatography with Shodex-803 column. *Lanes 9 and 10* contain mcfp-6 following C8 reverse phase HPLC and are stained with CBR-250 and NBT, respectively.

nants were precipitated by adding perchloric acid to 1.5% (v/v). Subsequent protein precipitation with ammonium sulfate led to the removal of remaining contaminants including mcfp-1 and -2. High purity mcfp-6 was thus easily prepared by gel filtration chromatography of the protein soluble in ammonium sulfate (Fig. 3*B*). Reversed phase HPLC confirmed the homogeneity of mcfp-6 (Fig. 3, *C* and *D*). In contrast to mcfp-6, isolation of mcfp-5 from the phenol gland required unusually aggressive conditions including sequential extraction with acetic acid, acid-urea, and acid-GdnCl. The sequential treatment removed most unwanted proteins, which simplified the chromatographic purification of mcfp-5 to a single step: C8 HPLC (Fig. 3*A*). Foot-derived mcfp-5 and -6 were evaluated by several criteria to assess their relationship to plaque proteins: They have highly comparable amino acid compositions, notably with respect to Dopa and phosphoserine (Table 1). N-terminal sequences (Figs. 2 and 3), and masses according to MALDI-TOF mass spectroscopy are the same: 8989 Da for mcfp-5 and 11,596 for mcfp-6 (Fig. 4). Finally, the foot- and plaque-derived proteins show similar electrophoretic behavior: on acid-urea PAGE; both run with the same mobility and stain light blue with Coomassie Blue R-250 and dark purple by redox-cycling stain with NBT (Fig. 3*D*). Although mcfp-5 ran as a single homogeneous band on SDS and acid-urea PAGE,

mcfp-6 exhibited multiple bands on acid-urea PAGE, whereas on SDS-PAGE there was a single sharp band (Fig. 3*D*). The multiple bands suggest charge heterogeneity because of variation in sequence and/or phosphorylation.

End-to-end Sequence Deduced from cDNA—The unambiguous N-terminal sequences of mcfp-5 and mcfp-6 enabled the design of degenerate oligos and, coupled with T7 universal primer encoded by the vector, to PCR-amplify the cDNA of mature proteins from cDNA library. With 5'-RACE strategy, complete sequences including signal peptides and 5'-untranslated regions were deduced (Fig. 5). Predicted signal peptide cleavages in the cDNA-deduced mcfp-5 and -6 sequences appear 16 and 21 residues after the initiating Met, respectively ("SIGNALP", ExPaSY), and are consistent with the observed N termini detected by Edman degradation with the purified proteins.

The deduced sequence of mcfp-6 reveals it to be dominated by Tyr (20.2 mol %), Gly (14.1 mol %), and Cys (11 mol %), which was in excellent agreement with the amino acid analysis of the purified

TABLE 2

Calculated and observed masses for mcfp-5 and -6

Calculated masses are based on cDNA-deduced sequences (M_{cDNA}), and selected levels of Dopa (M_{dopa}), and phosphoserine (M_{dopa+P}); m/z_{obs} values correspond to $[M+H]^+$ observed by MALDI TOF mass spectrometry.

Mcfp-5	Mass _{cDNA}	Dopa	Mass _{dopa}	pSer	M _{dopa+P}	m/z _{obs}
α	8494.5	20	8814.5	2	8974.5	8974
		20	8814.5	1	8894.5	8894
		19	8798.5	5	9198.5	9195
		18	8782.5	5	9182.5	9190
β	8536.5	20	8856.5	7	9416.5	9412
		19	8840.5	3	9080.5	9083

Mcfp-6	Mass _{cDNA}	Dopa	Mass _{dopa}	pSer	M _{dopa+P}	m/z _{obs}	m/z _{obs}
α	11,579.8	1	11,595.8	6	12,075.8		
β	11,562.9	2	11,594.9	6	12,074.9	11,596	12,072
γ	11,550.9	3	11,598.9	6	12,078.9		
δ	11,549.9	3	11,597.9	6	12,077.9		
ε	11,475.8	3	11,523.8	6	12,003.8		

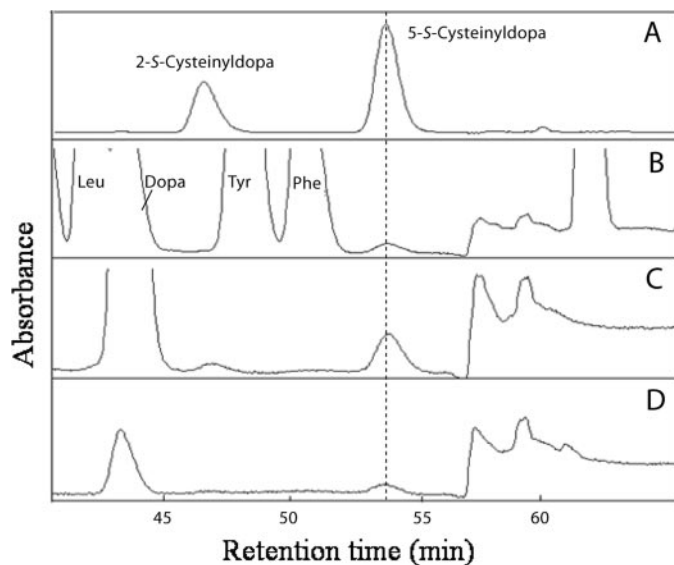


FIGURE 6. Amino acid analysis of cysteinylDopa cross-links in the footprint. A, authentic 2- and 5-S-cysteinylDopa. B, footprint hydrolysate (1.5 h hydrolysis at 110 °C). C, footprint hydrolysate after phenylboronate chromatography. D, hydrolysate of mcfp-6 after *in vitro* oxidation with mushroom tyrosinase and phenylboronate chromatography. Separation is by cation exchange chromatography and absorbance is 570 nm following reaction with ninhydrin.

o-quinones (19), we speculated that reduced cysteines might be effective as quinone scavengers in the footprint area. Because thiol-quinone adducts are relatively acid-stable (15, 16), plaque footprints were hydrolyzed and screened for cysteinyl-Dopa cross-links. Using ninhydrin-based amino acid analysis calibrated with standards (Fig. 6), the 5-S-cysteinylDopa isomer was detected at levels of 10% of the Dopa content in footprints or between 5–10 res/1000 res overall. Phenylboronate-coupled agarose is an effective affinity column for binding cysteinyl-Dopa adducts. By combining this step with ion exchange chromatography (same as that used in amino acid analysis but without the ninhydrin reaction) (Fig. 6), the purification and desalting of peak fractions was adequate to identify monoprotonated 5-S-cysteinylDopa (m/z 317.1) in the 54-min peak by electrospray ionization and tandem mass analysis (Table 3). Following collision-induced decomposition in tandem mass spectrometry, the fragments at 300.0, 237.0, 228, and 182 are considered the most diagnostic of 5-S-cysteinylDopa (15). 5-S-

TABLE 3

A comparison of fragment masses for standard 5-S-cysteinylDopa and the unknown amino acid from footprint hydrolysates at 54 min

Masses were determined using tandem mass spectrometry following collision induced decomposition. Chemical structures of fragment ions (bold) were cataloged in Ref. 15. Italicized m/z represents an impurity.

Ion	Standard		AA 54 min	
	m/z	%	m/z	%
Parent	317.12		317.12	
Fragment	300.06	5	300.10	11
1	—	—	<i>289.13</i>	5
2	271.09	4	271.06	7
3	254.05	5	254.05	12
4	—	—	<i>244.11</i>	13
5	236.02	9	237.02	9
6	228.03	11	228.03	20
7	208.05	3	208.07	5
8	182.03	11	182.03	12
9	—	—	<i>166.06</i>	3
10	155.01	100	155.02	100
11	—	—	<i>110.07</i>	13
12	74.03	5	74.03	4

CysteinylDopa was also formed *in vitro* by incubating a solution of mcfp-6 with mushroom tyrosinase at pH 8 (Fig. 6).

DISCUSSION

The adhesive footprints of *M. californianus* contain several distinct proteins that are detected *in situ* by MALDI-TOF mass spectrometry. Although a few mcfp-3 variants dominate the footprint spectra (3), others such as mcfp-3-1, -5, and -6 are detectable at less than a tenth of the intensity. It is tempting to assume that the intensity of the m/z signals reflects the relative abundance of each protein in the footprint, but we would caution against this interpretation. Cross-linking or surface chemisorption, for example, could significantly diminish the ability of certain abundant molecules to desorb and ionize in MALDI.

If a high Dopa content and proximity to the substrate surface is suggestive of a surface coupling role for mcfp-3 (3), then this role is even better suited to mcfp-5. mcfp-5 resembles the more abundant mcfp-3 variants in having a low mass, abundant Dopa, and a high pI. It is distinct, however, from mcfp-3 in its phosphorylation and relative lack of heterogeneity: the two deduced cDNA sequences differ by a single amino acid. A homologous protein, mcfp-5, with 76% identity to mcfp-5, was previously isolated from *Mytilus edulis* and detected in plaques (Fig. 5) (8, 20). Tyrosine conversion to Dopa in mcfp-5 approaches 100%. This is higher than in any of the mcfp-3 variants and intriguing because the Dopa-containing sequences are quite variable: roughly two-thirds of the Dopa residues occur in clusters with flanking Lys or Arg on one or both sides. The remainder (clustered toward the N terminus) are flanked by small neutral residues such as Gly, Pro, Asn, or Ala.

mcfp-6 resembles other Mfps in some respects (high pI and presence of Dopa) but it is peculiar in others. Except for the Dopa residues at the N terminus, the high tyrosine content (20 mol %) is rather inefficiently modified to Dopa (<5%) even though the tyrosine-containing sequences resemble those in mcfp-3 and -5 (Fig. 5). The Dopa-rich termini and the relatively high cystine content (5.5%) are reminiscent of mcfp-2, which is another, larger plaque protein (21, 22). Up to a third of the total Cys of mcfp-6 is in the thiolate form, which was previously observed only in *Perna fps* (15). mcfp-5 and -6 are phosphoryl-

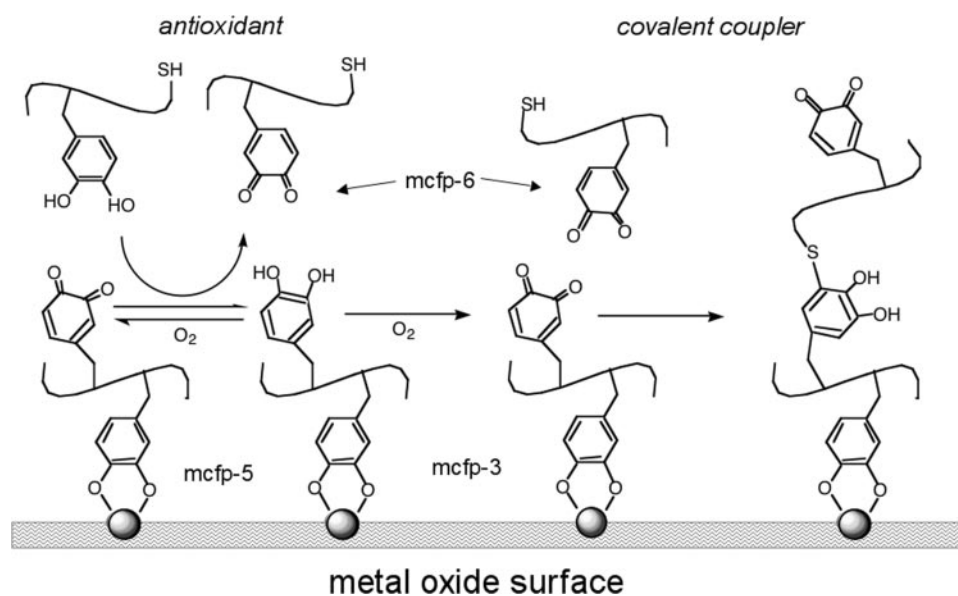


FIGURE 7. Model proposing roles of mcfp-3, -5, and -6 in plaque adhesion. Dopa-rich variants of mcfp-3 (α and β) and mcfp-5 are shown as surface-coupling agents, whereas mcfp-6 is a heterobifunctional protein cross-linker with both thiol and Dopa groups. Thiols rapidly form adducts with *o*-quinones, which are likely to form from any Dopa residues in mcfp-3 and -5 not adsorbed to the surface as well as Dopa in mcfp-6.

ated though apparently less so than mcfp-5 (13) and Pc-3, a cement protein in the sandcastle worm (14). Phosphorylation imparts a potential for both cohesive (by Ca²⁺ bridging) and adsorptive contributions to the glue (14, 23).

Given the ability of Dopa in synthetic polymers to coordinate to surface oxides (5), and that having more Dopa means stronger adhesion (24, 25), it makes sense that proteins with enhanced Dopa should be concentrated in the plaque footprints. The Dopa contents of the mcfp-3-1 and mcfp-5, for example, are at 24 and 30 mol %, respectively, the highest among all the byssal precursor proteins in *M. californianus*. In mcfp-6, however, the Dopa level barely reaches 4 mol %, which begs the question as to what its role in the footprint might be. Two hypotheses were pursued to address this. The first has nothing to do with Dopa but was suggested by the 31% identity with phospholipase A2 (Fig. 5C), *i.e.* that mcfp-6 functions to hydrolyze membrane phospholipids in surface-associated microbes. Despite testing a wide range of assay conditions, no phospholipase A activity could be detected in mcfp-6 (results not shown). Indeed none was anticipated because mcfp-6 is missing a His needed for the catalytic triad of phospholipase (26).

The second hypothesis was that mcfp-6 is involved in cross-linking footprint proteins. With eight ϵ -amino and three thiol equivalents/mol, the protein is well endowed with reactive side chains. Thiols are up to 4000 times more reactive than amines in the nucleophilic attack of quinones (19, 27) formed by the oxidation of peptidyl-Dopa. The participation of thiols in protein cross-linking was explored by hydrolyzing plaque footprints and screening for Dopa adducts by affinity chromatography on phenylboronate. In a previous study of whole plaques in *M. edulis*, only 5,5'-didopa cross-links were detected (28). The significant amounts of 5-*S*-cysteinylDopa detected here are presumably associated with insoluble footprint material. The cysteine in cysteinylDopa may be contributed by mcfp-6

because it is the only known thiol-containing plaque protein, but this awaits characterization of cross-linked peptides. The Dopa moiety, in contrast, could come from many sources including mcfp-3, -5, or -6. A previous study has determined that footprints maintain a strongly reducing environment for most of the Dopa in the mcfp-3 variants (3) so it is of considerable interest that, notwithstanding this, some Dopa residues are specifically targeted for oxidation and cross-linking.

The presence of cysteinylDopa cross-links in the adhesive footprints with mcfp-6 as the possible thiol donor suggests the following scenario (Fig. 7): mcfp-3 and -5 are concentrated near the surface to optimize Dopa binding. Perhaps only some of the Dopa (Y) sequences bind while others occur

in sequences directed away from the substratum *e.g.* GYG *versus* KYK. Unbound Dopa residues are more prone to oxidize to dopaquinones and become available for protein-protein interactions. Reactants such as mcfp-6 with Dopa and thiolates have three reaction options with such dopaquinones: 1) sacrifice their own Dopa for a 2-electron reduction of the dopaquinone back to Dopa; 2) also for a 1-electron reduction leading to a pair of semiquinones (and probable ring coupling), or 3) use their cysteine residues to scavenge the *o*-quinones by adduct formation (Fig. 7). The last of these is the fastest and surest of the three (19) and would provide an important link between the proteins at the surface and interior of the plaque.

Acknowledgments—We thank Dong-Soo Hwang for helping construct the expression library of *M. californianus*. Drs. K. Wakamatsu and S. Ito of Fujita Health University, Japan, generously provided the 2- and 5-*S*-cysteinylDopa.

REFERENCES

1. Waite, J. H., Holten-Andersen, N., Jewhurst, S. A., and Sun, C. J. (2005) *J. Adhesion* **81**, 297–317
2. Walker, P. (1987) in *Surface Coatings* (Wilson, A.D., Nicholson, J.W., and Prosser, H.J., eds) vol. 1, pp 189–232, Elsevier Applied Science, New York
3. Zhao, H., Robertson, N. B., Jewhurst, S. A., and Waite, J. H. (2006) *J. Biol. Chem.* **281**, 11090–11096
4. Papov, V. V., Diamond, T. V., Biemann, K., and Waite, J. H. (1995) *J. Biol. Chem.* **270**, 20183–20192
5. Dalsin, J. L., Lin, L., Tosatti, S., Vörös, J., Textor, M., and Messersmith, P. B. (2005) *Langmuir* **21**, 640–646
6. Israelachvili, J. N. (1985) *Intermolecular and Surface Forces*, pp. 12–23, Academic Press, London
7. Ohkawa, K., Nishida, A., Yamamoto, H., and Waite, J. H. (2004) *Biofouling* **20**, 101–105
8. Waite J. H., and Qin, X. (2001) *Biochemistry* **40**, 2887–2893
9. Waite, J. H., and Benedict, C. V. (1984) *Methods Enzymol.* **107**, 397–413
10. Paz, M., Flückinger, R., Boak, A., Kagan, H. M., and Gallop, P. M. (1991) *J. Biol. Chem.* **266**, 689–692

Adhesive Footprint Proteins in Mussels

11. Waite, J. H. (1995) *Methods Enzymol.* **258**, 1–19
12. Stewart, R. J., Weaver, J. C., Morse, D. E., and Waite, J. H. (2004) *J. Exp. Biol.* **207**, 4727–4734
13. Zhang, J. G., Matthews, J. M., Ward, L. D., and Simpson, R. J. (1997) *Biochemistry* **36**, 2380–2389
14. Zhao, H., Sun, C. J., Stewart, R. J., and Waite, J. H. (2005) *J. Biol. Chem.* **280**, 42938–42944
15. Zhao, H., and Waite, J. H. (2005) *Biochemistry* **44**, 15915–15923
16. Burzio, L. A., and Waite, J. H. (2000) *Biochemistry* **39**, 11147–11153
17. Shier, W. T., and Trotter, J. T. (1978) *Anal. Biochem.* **87**, 604–611
18. Durkin, J. P., Pickwell, G. V., Trotter, J. T., and Shier, W. T. (1981) *Toxicol.* **19**, 535–546
19. Sternson, A. W., McCreery, R., Feinberg, B., and Adams, R. N. (1973) *Electroanal. Chem. Interf. Electrochem.* **46**, 313–321
20. Waite, J. H. (2002) *Integr. Comp. Biol.* **42**, 1172–1180
21. Rzepecki, L. M., Hansen, K. M., and Waite, J. H. (1992) *Biol. Bull.* **183**, 123–137
22. Inoue, K., Takeuchi, Y., Miki, D., and Odo, S. (1995) *J. Biol. Chem.* **270**, 6698–6701
23. Lee, B. P., Chao, C. Y., Nunalee, F. N., Notan, E., Shull, K. R., and Messersmith, P. B. (2006) *Macromolecules* **39**, 1740–1748
24. Yu, M., Hwang, J., and Deming, T. J. (1999) *J. Am. Chem. Soc.* **121**, 5825–5826
25. Thompson, J. B., Kindt, J. H., Drake, B., Hansma, H. G., Morse, D. E., and Hansma, P. K. (2001) *Nature* **414**, 773–776
26. Hiraoka, M., Abe, A., and Shayman, J. A. (2005) *J. Lipid Res.* **46**, 2441–2447
27. Tse, D. C. S., McCreery, R. L., and Adams, R. N. (1979) *J. Med. Chem.* **19**, 37–40
28. McDowell, L. M., Burzio, L. A., Waite, J. H., and Schaefer, J. (1999) *J. Biol. Chem.* **274**, 20293–20295

