Proteins in Load-Bearing Junctions: The Histidine-Rich Metal-Binding Protein of Mussel Byssus[†,‡]

Hua Zhao[§] and J. Herbert Waite*[,§,||]

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

Received August 17, 2006; Revised Manuscript Received September 26, 2006

ABSTRACT: Building complex load-bearing scaffolds depends on effective ways of joining functionally different biomacromolecules. The junction between collagen fibers and foamlike adhesive plaques in mussel byssus is robust despite the strikingly dissimilar connected structures. mcfp-4, the matrix protein from this junction, and its presecreted form from the foot tissue of *Mytilus californianus* were isolated and characterized. mcfp-4 has a mass of ∼93 kDa as determined by MALDI-TOF mass spectrometry. Its composition is dominated by histidine (22 mol %), but levels of lysine, arginine, and aspartate are also significant. A small amount of 3,4-dihydroxyphenyl-L-alanine (2 mol %) can be detected by amino acid analysis and redox cycling assays. The cDNA-deduced sequence of mcfp-4 reveals multiple variants with highly repetitive internal structures, including ∼36 tandemly repeated His-rich decapeptides (e.g., HVHTHRVLHK) in the N-terminal half and 16 somewhat more degenerate aspartate-rich undecapeptides (e.g., DDHVNDIAQTA) in the C-terminal half. Incubation of a synthetic peptide based on the His-rich decapeptide with Fe^{3+}, Co^{2+}, Ni^{2+}, Zn^{2+}, and Cu^{2+} indicates that only Cu is strongly bound. MALDI-TOF mass spectrometry of the peptide modified with diethyl pyrocarbonate before and after Cu binding suggests that histidine residues dominate Cu binding. In contrast, the aspartate-rich undecapeptides preferentially bind Ca^{2+}. mcfp-4 is strategically positioned to function as a macromolecular bifunctional linker by using metal ions to couple its own His-rich domains to the His-rich termini of the preCOLs. Ca^{2+} may mediate coupling of the C-terminus to other calcium-binding plaque proteins.

The robustness of biological load-bearing scaffolds is defined partly by the mechanical properties of the various components present in the scaffold and partly by how these components are joined to one another. The musculoskeletal system, for example, consists largely of striated muscle, tendon, bone, and cartilage that are interconnected through junctions such as the enthesis (*1*). Quite a lot is known about structure−property relationships within bone (*2*), muscle (*3*), tendon (*4*), and cartilage (*5*), but how these merge with one another especially at a molecular level is still largely unexplored. The characterization of entheses is crucial for understanding how loads are distributed in natural scaffolds and, in particular, for engineering superior multifunctional biomedical materials.

The byssus or holdfast of marine mussels offers a convenient model for studying some aspects of molecular joinery particularly with regard to the load bearing function. Each byssal thread contains no living cells, readily lends itself to mechanical characterization, and consists of relatively few matrix proteins that are quickly and reproducibly put together. The focus of this study is the critical junction between the distal portion of the thread and the adhesive plaque (Figure 1). From an architectural perspective, this is an intriguing junction. The thread consists of densely and anisotropically packed collagenous microfibrils, whereas the plaque is a solid foam with a porosity ranging from 1 to 5 μm in pore diameter (*6*−*8*). Microscopic inspection of sections of the thread−plaque junction has revealed that the collagenous microfibrils undergo extensive splaying where they meet the foam, not unlike a tree deeply rooted in the earth. Apart from the large amount of surface area created by the splaying, the molecular aspects of interaction have not been explored.

The plaques and threads produced by the California mussel (*Mytilus californianus* Conrad) are sufficiently large to enable a dissection and molecular characterization of junction-specific proteins. Previous research has determined that the distal thread of *Mytilus* species consists primarily of two collagens and small amounts of a thread matrix protein tmp-1 (*9*, *10*). The collagens, known as preCOL-D and preCOL-NG, are peculiar in being kinked and having a linear block domain structure that includes silklike, Gly-rich, collagen, and His-rich domains (*11*, *12*). The plaque is composed of three parts: the adhesive footprint which contains mcfp-3, -5, and -6 (*13*, *14*) and the structural foam which is

[†] 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 Grant NCC-1-02037) and the National Institutes of Health (DE015415).

[‡] Data were deposited as GenBank accession numbers DQ351535 and DQ351536.

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

[§] Molecular, Cellular and Developmental Biology Department.

[||] Marine Science Institute and Chemistry and Biochemistry Department.

10.1021/bi061677n CCC: $33.50
Published on Web 11/02/2006

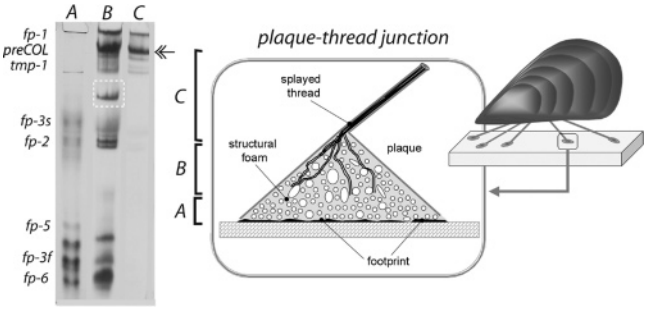

FIGURE 1: Schematic drawing of the mussel byssus illustrating a close-up of the junction between the thread and plaque portions and the extensive interface between the splayed collagen fibrils and foam proteins of the plaque (right). Gel electrophoresis (acid urea−PAGE) of the proteins extracted from three microdissected portions of freshly secreted byssal threads (left) [(A) footprint region, (B) junction, and (C) distal thread]. The gels have been stained for Dopa with glycinate and nitroblue terazolium (NBT). Identification of known protein families is according to previous studies (*9*−*16*). The boxed band is mcfp-4, and the double-headed arrow marks the position of an internal calibrant, the α chain of type I collagen.

dominated by mfp-2, a 45 kDa protein with 11 tandem repeats of an epidermal growth factor-like domain (*15*, *16*), and a large transitional region in which the foam and collagen fibrils interpenetrate one another. The aim of this study was to explore the plaque−thread junction for proteins capable of mediating the contact between the fibrous preCOLs and foam proteins. mcfp-4[1] is a junction-specific protein with two different but well-defined metal-binding domains.

MATERIALS AND METHODS

Dissection of the Thread−Plaque Junction. M. californianus was collected locally from Goleta Pier in Santa Barbara, CA, immediately transferred to the lab, and maintained in shallow tanks with running seawater. Because cross-linking in naturally secreted byssal threads is too rapid to allow efficient extraction of proteins, thread formation "on demand" was induced by KCl injection, a technique shown in an earlier study to lead to byssal secretions that were essentially indistinguishable from the noninduced ones (*6*). Briefly, 0.3 mL of 0.56 M KCl was injected once or twice into the base of the foot with a syringe (18 gauge needle). Within 30 min, a cream-colored byssal thread was assembled within the ventral groove of the foot. This was removed with microforceps and placed on a microscope slide frozen at −80 °C. When 10 induced threads were accumulated, they were dissected into three parts [footprint (A), junction (B), and distal thread (C); see Figure 1] with the help of a Nikon stereomicroscope and a #12 scalpel blade. Corresponding parts from 10 threads were pooled and minced in extraction buffer (8% acetic acid with 8 M urea and 100 mM dithiothreitol) using a glass rod and ceramic test plate (Coors #60427) as a mortar and pestle. The homogenate was microfuged (Eppendorf model 5415C) for 5 min at 17000*g* and ∼7 °C, and 5−10 μL aliquots of the supernatant were subjected to acid urea−polyacrylamide gel electrophoresis (AU PAGE) and a redox cycling stain described below.

Isolation of Bulk Protein from Adhesive Plaques. To collect plaques for protein isolation, mussels were tethered onto plastic plates and left to deposit plaques for 24 h. After this interval, plaques were scraped free using a clean single-edge razor blade and rinsed extensively with Milli-Q water to remove salts. Collected plaques were immediately extracted or stored at −80 °C for future use.

Once ∼500 plaques or 2 g of wet weight was accumulated, the plaques were homogenized on ice with 20 mL of 5% acetic acid and 8 M urea using a small hand-held tissue grinder (Kontes, Vineland, NJ). The supernatant, which contained crude mcfp-4, was harvested by centrifugation for 40 min at 20000*g* and 4 °C using an SS-34 rotor. The crude extract of mcfp-4 from adhesive plaques was adjusted to pH 7.5 with 6 M NaOH. During the adjustment of the supernatant pH, 1 mg/mL Na_2BO_4 powder was included for protection of Dopa residues. Partial purification of mcfp-4 from plaques was accomplished by immobilized metal affinity chromatography (IMAC) using a 1 mL chelating column from Amersham-Pharmacia charged with Zn^{2+}. Equilibration buffer contained 20 mM Tris, 10 mM Na_2BO_4, 0.5 M NaCl, and 8 M urea (pH 7.5), whereas for elution, the same buffer acidified to pH 3.0 was used. After the crude extract of mcfp-4 had been slowly loaded onto a pre-equilibrated IMAC chelating column, the column was flushed extensively with equilibration buffer and then eluted with 8 column volumes of the acidic elution buffer. Elution was monitored at 280 nm with an UV monitor, and 1 mL fractions were collected. Peak absorbing fractions were pooled, dialyzed against 5% acetic acid overnight, and lyophilized. The lyophilized, partially purified mcfp-4 was reconstituted in a small volume of 5% acetic acid. Purification of mcfp-4 was achieved by gel filtration chromatography using a Shodex-803 column (5 μm, 8 mm × 300 mm) which was equilibrated and eluted with 5% acetic acid in 0.2% TFA. The sample load was limited to 200 μL/run, and the eluted volume was monitored at 280 nm.

Isolation of Protein from the Mussel Feet. All plaque proteins are stockpiled in the phenol gland of the mussel foot, which is an excellent alternative source of protein when larger quantities are needed (see the Supporting Information). Mussels were shucked by severing the adductor muscles, after which the feet were carefully amputated, individually arrayed on glass plates, and frozen at −80 °C. While the feet remained adhesively frozen to the plates, the pigmented outer integument was quickly stripped away with a scalpel, whereupon the underlying phenol glands became visible and harvested by dissection. Phenol gland collections (0.2−0.3 g wet weight) were homogenized to a puree on ice in 5% acetic acid with protease inhibitors (10 μM leupeptin and pepstatin) and centrifuged at 20000*g* and 4 °C. Supernatant S1 was set aside, and the pellet (P1) was rehomogenized in 5% acetic acid with 8 M urea. Supernatant S2 was harvested by centrifugation at 20000*g* and 4 °C and set aside as well. The pellet (P2) was rehomogenized in 5% acetic acid containing 4 M guanidine HCl. After centrifugation at 20000*g*, the pellet (P3) was discarded and the supernatant (S3) was decanted into a small beaker. Ammonium sulfate was slowly added to a final concentration of 25% (w/v). The mixture was stirred for 40 min at room temperature. The

[1] Abbreviations: DEPC, diethyl pyrocarbonate; Dopa, 3,4-dihydroxyphenyl-L-alanine; MALDI-TOF, matrix-assisted laser desorption and ionization with time of flight; mcfp-4, *Mytilus californianus* foot protein 4; NBT, nitroblue tetrazolium; PAGE, polyacrylamide gel electrophoresis; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcriptase polymerase chain reaction.

precipitate was removed by centrifugation at 20000*g* and 4 °C for 30 min, and the supernatant (S4) was collected and dialyzed overnight against 4 L of 0.5% perchloric acid using dialysis tubing having a 1000 Da cutoff (Spectrum Industries, Los Angeles, CA) (*17*).

Dialysis resulted in visible turbidity, which was clarified by centrifugation at 20000*g* for 30 min at 4 °C. Redialysis of the supernatant against 0.1 M sodium borate at pH 8 precipitated most of the mcfp-4 and small amounts of fp-3 and fp-5, which were harvested by centrifugation for 30 min at 20000*g* and resuspended in a small volume of 5% acetic acid with 8 M urea. Residual particulate matter was removed by spinning the samples at 14K rpm for 20 min in an Eppendorf microfuge. Pure mcfp-4 was obtained by reverse phase HPLC (RP-300 Aquapore device with dimensions of 260 mm × 7 mm from Applied Biosciences Inc., Foster City, CA) using a linear gradient of aqueous acetonitrile for elution. Eluant was monitored continuously at 220 nm, and collected 1 mL fractions were assayed by amino acid analysis and electrophoresis following freeze-drying.

Tryptic Peptides. Approximately 0.5 mg of mcfp-4 purified by reverse phase HPLC was digested with trypsin at an enzyme:substrate ratio of 1:50. The digestion was performed overnight in 50 mM Tris and 2 M urea with 5 mM Ca^{2+} at pH 7.5 and room temperature (*18*). The digest was directly subjected to HPLC with a C18 column (260 mm × 7 mm, Applied Biosciences Inc.) using a linear gradient of 0 to 30% acetonitrile in 0.1% TFA over 40 min.

Electrophoresis. Because mcfp-4 precipitated in the presence of sodium dodecyl sulfate (SDS), SDS−polyacrylamide gel electrophoresis (PAGE) was avoided, and routine electrophoresis was performed on 7% polyacrylamide gels with 5% acetic acid and 8 M urea (*19*). In contrast to SDS−PAGE, acid urea−PAGE lacks the convenient correlation between mobility and mass. The method, however, exhibited good resolution of mfps, and the use of an internal protein marker (rabbit skin type I collagen; Sigma) allowed correction for day-to-day run variability. 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 (*20*). Parallel gels were also stained with Pauly's reagents for histidine-rich proteins exactly according to a published procedure (*21*). Pauly stains orange-red initially but fades to bright yellow with time.

Amino Acid Analysis and Sequencing. Purified protein was hydrolyzed in 6 M HCl with 5% phenol in vacuo at 110 °C for 24 h. Hydrolysate was flash-evaporated at 50 °C under vacuum and washed to dryness with a small volume of Milli-Q water, followed by a methanol wash. Amino acid analysis was performed under conditions described previously with a Beckman System 6300 Auto Analyzer (*22*). The amino acid sequence of mcfp-4 and tryptic peptides was determined by automated Edman degradation on a Porton Instruments (Beckman Coulter) model 2090 microsequencer. PTH-Dopa elutes at 10 min between PTH-Ala and PTH-His (*22*).

Mass Spectrometry. The mass of Mcfp-4 was determined by matrix-assisted laser desorption and ionization with time of flight (MALDI-TOF) using a PerSeptive Biosystems Voyager DE model (AB Biosystems, Foster City, CA) in the positive ion mode with delayed extraction. The MALDI matrix buffer was prepared by dissolving sinapinic acid (10 mg/mL) in 30 vol % acetonitrile. The purified mcfp-4 was dissolved in this matrix solution to produce a final concentration between 1 and 10 pmol/μL. Approximately 1 μL of this solution was applied to the target plate and allowed to evaporate. The sample spots were irradiated using a N_2 laser (LSI, Inc., Cambridge, MA) with a wavelength of 337 nm, a pulse width of 8 ns, and a frequency of 5 Hz. Singly and doubly charged analyte ions were accelerated using an accelerating voltage of either 20 or 25 kV. The singly and doubly charged peaks of an IgG mass standard from Applied Biosystems were used as two-point molecular mass calibrants (148 500 and 74 249, respectively, for average masses).

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 the Qiagen protocol was followed after initial tissue disruption under liquid nitrogen with a mortar and pestle. Then, mRNAs were purified from total RNA using the Oligotex mRNA Mini Kit from Qiagen.

With purified mRNAs, a cDNA library was constructed by using the CloneMiner cDNA library construction kit from Invitrogen. This cDNA library served as a readily available source of cDNA. Alternatively, first-strand cDNA was synthesized from total RNA using Superscript II reverse transcriptase with an adapter primer, 5′-GGC CAC GCG TCG ACT AGT ACT $(T)_{16}$-3′ (Invitrogen). The product of the room-temperature reaction was used as an alternative template for PCR.

On the basis of the known tryptic peptide sequence of mcfp-4, HQVLHGHVHT, a degenerate forward primer, 5′-CAY CAR GTN TTR CAY GGN CAY GTN CAY AC-3′, was designed and coupled with an abridged universal amplification primer (antisense, 5′-GGC CAC GCG TCG ACT AGT AC-3′) from Invitrogen to amplify the carboxyl terminus of mcfp-4 from RT template using DyNAzyme EXT DNA polymerase from New England Biolabs for the long PCR product (catalog no. F-505S).

After the carboxyl-terminal sequence of mcfp-4 had been cloned (~2 kb with a 3′-untranslated region and poly-A tail), a reverse primer (5′-G TTT ATG TAG AAC ACG ATG CAT G-3′), which was designed according to the known carboxyl-terminal sequence, and a degenerate forward primer (5′-CCN AGY GGN TAY GCN AAY ATH GGN CA-3′ based on the N-terminal sequence of mature mcfp-4 by Edman degradation) were used to amplify the amino terminus of mcfp-4 from the cDNA library. The complete sequence of mature mcfp-4 was then determined by overlapping the partial cDNA clones of amino- and carboxyl-terminal cDNA sequences.

PCRs were carried out in 25 μL of 1× buffer B (Fisher) and 5 pmol of each primer, 5 μmol of each dNTP, 1 μL of the first-strand reaction mixture, 2 mM $MgCl_2$, and 2.5 units of Taq DNA polymerase (Fisher) for 35 cycles on a Robocycler (Stratagene). Each cycle consisted of 30 s at 94 °C, 30 s at 52 °C, and 1.5 min at 72 °C, with a final extension of 15 min. The PCR products were subjected to 1% agarose gel electrophoresis, purified, cloned into a pCR-XL-TOPO vector using the TOPO XL PCR cloning kit from Invitrogen, and transformed into competent Top10 cells for amplification, purification, and sequencing.

To obtain the 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 a gene-specific primer (antisense, 5′-GAA CGC GGT GAT TGT GAA CAT GTG AGT GTA AAA C-3′), which reversely primes the N-terminus of mcfp-4, and a GeneRacer 5′ primer from Invitrogen (sense, 5′-CGA CTG GAG CAC GAG GAC ACT GA-3′).

In Vitro Metal Binding Test with a Synthetic Histidine-Rich Peptide. The synthetic peptide, HGHVHTHRVLHKH-VHKHRVL, equivalent to two repeats of the decapeptide from mcfp-4, was synthesized (mass of 2464.89 Da; GenScript) and purified by reverse phase HPLC using a C18 column. Prior to MALDI-TOF analysis, the purified peptide (90 μM) was incubated with a 10-fold molar excess of $FeCl_3$, $NiCl_2$, $CoCl_2$, $ZnCl_2$, and $CuCl_2$ in 25 mM *N*-ethylmorpholine and 30 mM KCl buffer (pH 7.4) at room temperature for 10 min (*23*). Samples were aliquoted and examined by MALDI-TOF MS without prior purification.

Modification of the synthetic peptide by diethyl pyrocarbonate (DEPC) was performed in 50 mM Na/K phosphate buffer at pH 7.0. An ~3-fold molar excess of DEPC was incubated with peptide for 30 min at room temperature (*24*). The modified peptide was purified by C18 reversed phase HPLC with a shallow acetonitrile gradient. After lyophilization, the DEPC-modified peptide was incubated with a 10-fold molar excess of $CuCl_2$ in 25 mM *N*-ethylmorpholine and 30 mM KCl buffer (pH 7.4) at room temperature for 10 min (*23*). In some cases, prior to the DEPC treatment, the peptide was incubated with a 10-fold molar excess of $CuCl_2$ in Tris-HCl buffer at pH 7.0 for 30 min, and the samples were subjected to MALDI-TOF MS analysis without further purification.

In Vitro Calcium Binding Test with a C-Terminal Motif. A synthetic peptide, QTADDHVNDIAQTADDHVNA, which contains nearly two repeats of the undecapeptide from the C-terminal half of mcfp-4, was commercially synthesized (GenScript; mass of 2150.18 Da) and purified by C18 HPLC. To assess calcium binding, a 90 μM solution of the peptide was incubated with a 10-fold molar excess of $CaCl_2$ in 50 mM ammonium acetate buffer (pH 5.8) at room temperature for 10 min (*25*). Samples were aliquoted and examined by MALDI-TOF mass spectrometry without further purification.

RESULTS

Protein Extraction and Purification. The byssal adhesive plaques of adult *M. californianus* (shell length of 10 cm) are roughly elliptical in shape with major and minor axes of ~3 and ~2 mm, respectively, and resemble an inverted T in longitudinal section (Figure 1, right). KCl-induced byssal threads were sectioned into three pieces as shown within 15 min of secretion: (A) the footprint or plaque base, (B) the junction region, and (C) the distal thread. The pieces were extracted with 8% (v/v) acetic acid and 8 M urea, and extracts were subjected to acid urea−PAGE. Redox cycling-stained gels in Figure 1 (left panel) highlight the proteins in each of the pieces by their Dopa content. Prior studies have allowed identification of mcfp-3, -5, and -6 variants along with smaller amounts of mcfp-2 in the footprint (A) (*13*, *14*). mcfp-2 and -6 persisted or increased in the junction region (B), where slower-moving proteins such as the collagens

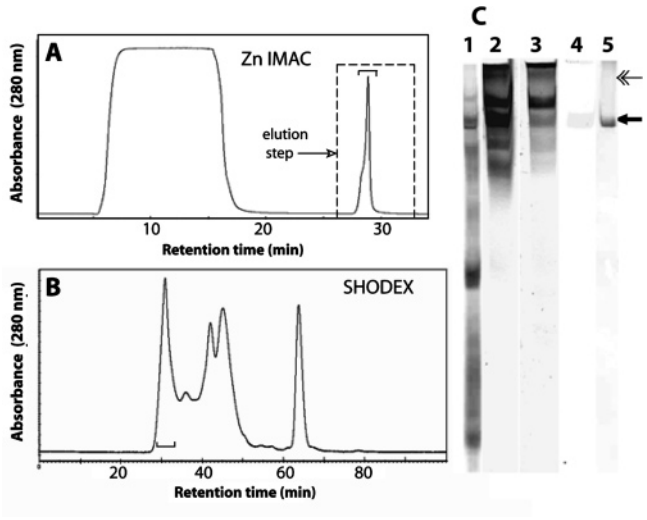

FIGURE 2: mcfp-4 from adhesive plaques. (A) Immobilized metal affinity chromatography (IMAC) of Mcfp-4 with a chelating column charged with Zn^{2+}. Protein was eluted with a stepped increase to pH 3. (B) Gel filtration chromatography on Shodex 803 of the eluted IMAC-binding fractions. His-rich fractions are denoted with a bracket. (C) Acid urea−polyacrylamide gel electrophoresis of the key steps in mcfp-4 purification: crude soluble plaque extract from plaques stained with Coomassie blue R-250 (lane 1), IMAC column binding fractions stained in parallel for protein with CBR-250 (lane 2), Dopa with NBT in glycinate (lane 3), histidine with Pauly's reagent (lane 4), and Shodex-purified mcfp-4 (arrow) stained with CBR-250 (lane 5). The double-headed arrow marks an internal calibrant, the α chain of type I collagen.

(preCOLs) and thread matrix protein variants (tmp-1) prevail (*9*−*12*). Finally, in the thread proper (C), detectable proteins are limited to the preCOLs and tmp-1. Our results suggest that at least one protein, which we shall call *M. californianus* foot protein-4 (mcfp-4), was localized to the thread−plaque junction and migrated between the preCOLs and the slow variants of mcfp-3 (Figure 1, boxed, left panel).

To better characterize this protein, a large number of plaques were collected and extracted. On the basis of superficial similarities between mcfp-4 and a partially characterized histidine-rich protein from *Mytilus edulis* (*26*), we attempted to further separate the soluble proteins on a Zn^{2+} IMAC column. Several proteins were bound by IMAC and eluted with a step to pH 3 (Figure 2A,C, lane 2). Two were strongly NBT-positive following acid urea−PAGE (suggesting Dopa) and migrated between the preCOLs and mcfp-3s. Although both may be mcfp-4 variants, the more mobile band also stained strongly for histidine (Figure 2C, lanes 3 and 4). Gel filtration on a Shodex-803 column was able to resolve the remaining proteins with the histidine-rich component being the first to elute (Figure 2B,C, lane 5). The *m*/*z* of this fraction was 92 kDa as determined by MALDI-TOF mass spectrometry.

In contrast to the plaque, extraction of dissected foot tissue (i.e., phenol glands) with 8 M urea and 5% acetic acid was inadequate in attempting to solubilize mcfp-4. Only the use of 4 M guanidine HCl and 5% acetic acid liberated significant amounts of the protein as it did in the case of mefp-5 (*19*) (Figure 3A). mcfp-4's finicky solubility, however, was used to advantage in a serially extractive strategy; that is, via successive extraction of the insoluble residue of foot tissue with 5% acetic acid, 5% acetic acid and 8 M urea, and 5% acetic acid and 4 M guanidine HCl, little but mcfp-4 and -5

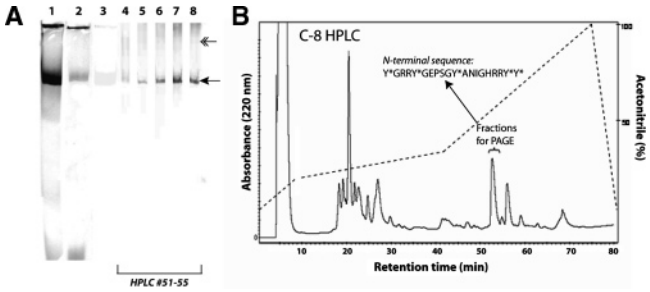

FIGURE 3: mcfp-4 from mussel feet. (A) Acid urea−polyacrylamide gel electrophoresis of key steps in the purification of foot-derived mcfp-4. Guanidine-extracted and borate-dialyzed (pH 8.0) soluble protein stained with CBR-250 (lane 1), stained for Dopa with NBT (lane 2), and stained for histidine with Pauly's reagent (lane 3). The remaining lanes (lanes 4−8) contained fractions under the bracketed peak of the reverse phase C8 HPLC column (B) stained with CRB-250. The double-headed arrow marks an internal calibrant, the α chain of type I collagen.

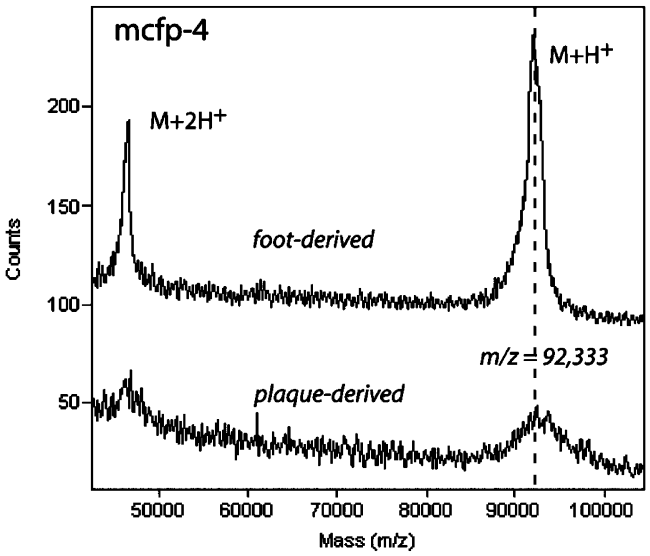

FIGURE 4: MALDI-TOF mass spectrum of purified mcfp-4 from both mussel feet and plaques. Peaks from right to left were $[M + H]^+$ and $[M + 2H]^{2+}$. Delayed extraction (200 ns) in positive ion mode with an accelerating voltage of 25 000 V, a grid voltage at 93%, and a guide wire voltage at 0.1%. The spectrum represents an average of 256 scans.

remained in the last supernatant especially after dialysis against 1% (v/v) perchlorate. mcfp-4 was finally harvested from PCA by a final dialysis against 0.1 M borate, which caused extensive but reversible precipitation as followed by acid urea−PAGE (Figure 3A, lanes 1−3). A single pass through a reverse phase C8 HPLC column (Figure 3B) was adequate to purify mcfp-4 to homogeneity (Figure 3A, lanes 4−8). Peak HPLC fractions containing pure mcfp-4 were subjected to MALDI-TOF mass spectrometry, which revealed strong singly and doubly charged ions at 92 433 and 46 217 Da, respectively (Figure 4).

Foot-derived mcfp-4 stained as a sharp band with Coomassie blue R-250, and also by redox cycling with NBT, which indicated the presence of some Dopa. The histidine-rich content of mcfp-4 was supported by intense orange staining with Pauly's reagent (Figure 3A,C).

Biochemical Characterization. The amino acid composition of mcfp-4 purified from both plaques and feet confirmed the unusual prominence of histidine detected at nearly 20 mol % (Table 1). Other abundant residues were Asx (8.7 mol %), Val (8.4 mol %), and Gly (8.3 mol %). Dopa was detected at only 1.7 mol %, which accounted for the staining with NBT being weaker than that of many of the other mcfp precursors (Figures 2C and 3A). Foot-derived mcfp-4 purified by C8 HPLC and subjected to Edman protein sequencing produced a clean, unambiguous N-terminus, Y*GRRY*-GEPSGY*ANIGHRRY*Y*, in which Y* denotes Dopa (Figure 3B). Additional sequence was obtained from a HQVLHGHVHT peptide derived by trypsinization of mcfp-4 (see the Supporting Information).

Table 1: Amino Acid Composition of mcfp-4 in Mole Percent (number of residues per 100 residues)[a]

amino acid	from foot	from plaque	predicted
Asx	8.7	11.3	11
Thr	3.0	3.3	2.8
Ser	5.1	5.2	4.4
Glx	7.4	8.8	6.8
Pro	4.7	1.3	0.7
Gly	8.3	7.3	5.2
Ala	5.1	5.6	4.5
Cys	0.4	0	0.3
Val	8.4	7.9	12.7
Met	0.5	1.8	0.7
Ile	3.2	3.9	4.3
Leu	5.7	6.1	6.8
Dopa	1.7	1.4	−
Tyr	2.2	1.5	2.3
Phe	2.2	2.2	1.3
His	19.2	18.3	23.6
Lys	6.3	6.6	5.5
Arg	7.6	7.1	7.1
total	100	100	100

[a] Values are averages of three analyses with a standard deviation that is ±15% of the mean. The predicted composition is based on mcfp-4, variant 1.

Molecular Cloning. It was not possible to obtain a complete cDNA for mcfp-4 on the basis of the limited N-terminal sequence that is available. However, additional acquisition of an internal decapeptide sequence (HQVLH-GHVH) enabled the design of a degenerate oligonucleotide that, when coupled with an abridged universal amplification primer from Invitrogen, successfully amplified the 3′ end of the transcript, including the 3′-untranslated region and poly-A tail from the RT template. The 5′ end was derived by PCR amplification of a foot-specific cDNA library using a degenerate oligonucleotide primer prepared to the partial N-terminal sequence (PSGYANIG) in combination with a downstream nondegenerate primer designed on a sequence within the partial 3′ end cDNA corresponding to the MHRVLHK sequence. With an additional 5′ RACE strategy, the complete sequence encoding the signal peptide and 5′-untranslated region was deduced. The signal peptide predicted by "SIGNALP" (EXPASY) is 21 residues long, and the cleavage site is consistent with the observed N-terminus of purified mcfp-4 (Figure 5).

The cDNA-deduced sequence revealed a very histidine-rich N-terminal moiety in which the histidine content approaches 41 mol %. The domain contains a tandemly repeated decapeptide motif, HVH*T*H*Q*VLH*G*, which is highly conserved and occurs ∼36 times in variant 1. The His-rich domain is punctuated by a linker sequence (linker A) containing the only cysteines in the protein before entering the second largest repeat sequence, the DN-rich domain, in which the levels of Asp and Asn approach 30 mol %, and

```
MKRGLFTILLITAVLVVVAES

YGRRYGEPSG YANIGHRRYY ERAISFHRHS 30

HVHGHHLLHR HVHRHSVLHG HVHMHRVSHR IMHRHRVLHG  70
HVHRHRVLHN HVHRHSVLHG HVHRHRVLHR HVHRHNVLHG 110
HVHRHRVLHK HVHNHRVLHK HLHKHQVLHG HVHRHQVLHK 150
HVHNHRVLHK HLHKHQVLHG HVHTHRVLHK HVHKHRVLHK 190
HLHKHQVLHG HIHTHRVLHK HLHKHQVLHG HVHTHRVLHK 230
HVHKHRVLHK HLHKHQVLHG HVHMHRVLHK HVHKHRVLHK 270
HVHKHHVVHK HVHSHRVLHK HVHKHRVEHQ HVHKHHVLHR 310
HVHSHHVVHS HVHKHRVVHS HVHKHNVVHS HVHRHQILHR 350
HVHRHQVVHR HVHRHLIAHR HIHSHQAAVH RHVHTHVF   387

EGNFNDDGTD VNLRIRHGII YFGGNTYRLS GGRRRFMTLW 428
QECLESYGDS DECFVQLGNQ HLFTVVQGHH STSFRSDLSN 468

DLHPDNNIEQIA NDHVNDIAQST DG-D--INDFA DTHYNDVAPIA 510
DVHVDN-IAQTA DNHVKNIAQTA HHHVNDVAQIA DDHVNDIGQTA 554
YDH-VNNIGQTA DDHVNDIAQTA DDHVNAIAQTA DDHVNAIAQTA 587
DDH-VNDIGDTA NSHIVRVQGVA KNHLYGINKAI GKHIQHLKDVS 642
NRH-IEKLNNHA                                     653

TKNLLQSALQ HKQQTIEREI QHKRHLSEKE DINLQHENAM 693

KSKVSYDGPVFNE 706
KVSVVSNQGSYNE 719
KVPVLSNGGGYNG 732
KVSALSDQGSYNE 745
GYAY          749
```

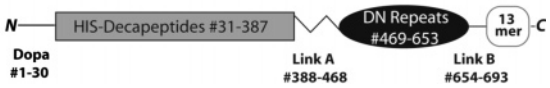

FIGURE 5: Complete deduced sequence of mcfp-4, variant 1. The schematic shows the domains of mcfp-4. The italicized portion is the signal peptide; underlined sequences denote those determined directly by Edman degradation. Variant 2 differs because it has five additional histidine-rich decapeptides.

with 16 repeats of undecapeptide *DD*HVN*DIAQT*A, in which the nonitalicized residues are substituted in less than half of the repeats (Figure 5). Only His-3 is never substituted. The DN domain is followed by another short linker sequence (linker B, Figure 5) before it ends with four repeats of a 13-amino acid sequence. The predicted amino acid composition based on the deduced mature sequence exhibited the expected overall prevalence of His and Asx at 23.6 and 11 mol %, respectively, and was consistent with the experimentally determined compositions of the purified mcfp-4s (Table 1).

Two mcfp-4 cDNA variants (variants 1 and 2) were detected and appear to represent alternatively spliced products. The variants differ only in the number of decapeptide repeats present in the His-rich domain, which is reflected in the calculated masses of 87.99 and 93.4 kDa, respectively. These are both close to the observed mass of 92 kDa, but the former would appear to require either 4 kDa of post-translational modification or the latter, 1−2 kDa of sequence trimmed from the C-terminus.

In Vitro Metal Binding. Precipitation of mcfp-4 by metals necessitated the use of a His-rich synthetic peptide for metal binding studies. The peptide is based on a sequence from the His-rich domain (residues 168−188), and binding was examined by a MALDI-based method first developed for assessing copper binding by prion peptides (*23*). Because of the low pH and probable protein denaturation during desorption and ionization, only the most robust metal binding by proteins is likely to survive MALDI. The observed *m*/*z* of the untreated monoprotonated $[M + H]^+$ peptide at 2466.1 was consistent with a predicted mass of 2464.9 Da (Figure 6). Metal adduct formation was assayed after incubation with five different metal ions. One additional peak to the parent ion at 2466.1 was observed for Zn (2530.2), Ni (2522.8), and Co (2523.1) and corresponded to the addition of 1 equiv of metal ion minus two expelled protons (for charge balance). In contrast, when incubated with Cu^{2+}, the peptide exhibited multiple *m*/*z* species, which corresponded to adducts of 1−11 copper ions (Figure 6). The data clearly suggest that the histidine-rich peptide is capable of preferential and high-capacity copper binding. Indeed, the observed maximum (11 Cu atoms/peptide) exceeds the number of histidines (eight per peptide) in the peptide.

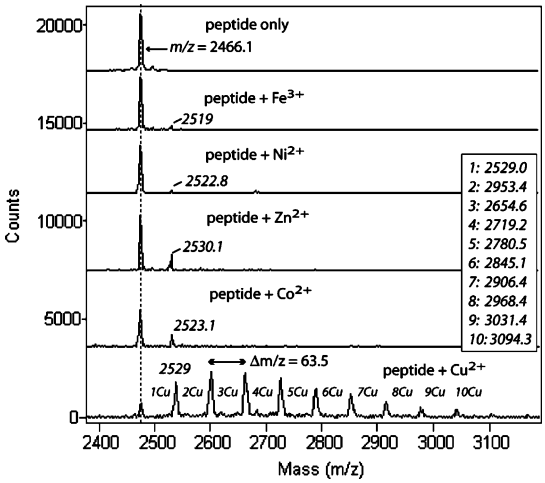

FIGURE 6: MALDI-TOF mass spectra of the synthetic His-rich decapeptide dimer and its metal adducts. The peptide (HGH-VHTHRVLHKHVHKHRVL), excerpted from the histidine-rich N-terminal domain, was incubated with a 10-fold molar excess of Fe^{3+}, Ni^{2+}, Zn^{2+}, Co^{2+}, and Cu^{2+} in 25 mM *N*-ethylmorpholine and 30 mM KCl buffer (pH 7.4).

The involvement of peptidyl-histidines in copper binding was explored by the use of diethyl pyrocarbonate (DEPC), a reagent known to preferentially modify histidine by carbethoxylation (CE) of the imidazole nitrogens (*27*). Modification of peptidyl-histidine with DEPC prior to exposure to the metal abolished copper binding (Figure 7A, top and bottom). The spectra are somewhat complicated by the presence of salt (Na + K; Δmass = 23 Da + 39 Da = 62 Da) adducts. Thus, each of the monoprotonated peptides with seven, eight, and nine carbethoxylations is accompanied by a sodiated and potassiated counterpart, e.g., $[M + 8CE + Na + K]^+$ is 62 Da above the $[M + 8CE + H]$ peak. This is a crucial control given how close the masses for Na + K and Cu are. At the concentrations that were used, DEPC was expected to monocarbethoxylate every histidine (*23*, *27*). Since the peptide mass increased by 72 Da with every monocarbethoxylation, eight histidines (*m*/*z* 3042.9) appeared to be modified along with one or two other sites (*m*/*z* 3110 and 3187.1), perhaps the N-terminal amine and an ϵ-amine of lysine (Figure 7A). Copper binding by the His-rich peptide largely prevented its modification by DEPC (Figure 7B). Apparently, not all of the sites were shielded from modification when copper was first bound by the peptide (Figure 7B), because at least two DEPC adducts (Δ = 72 Da) are clearly detectable. Presumably, these are the lower-affinity sites, but whether they are His or other residues remains to be determined.

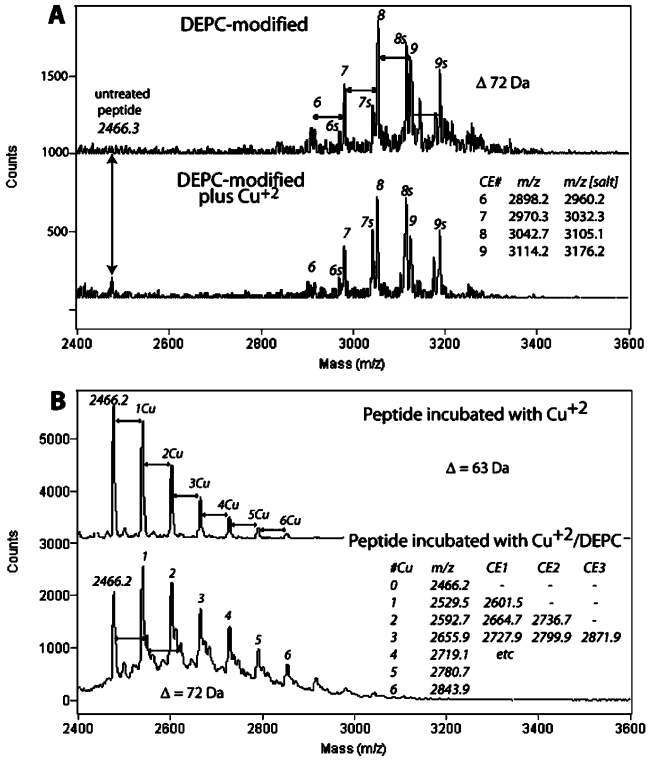

FIGURE 7: MALDI-TOF mass spectra of DEPC modification of the synthetic His-rich decapeptide before (A) and after (B) Cu^{2+} incubation. Integers over the peaks represent the calculated number of carbethoxylations. Note that all eight of the His residues and two other functionalities appear to be monocarbethoxylated DEPC (A, i), which prevents any Cu^{2+} binding that can be detected by MALDI-TOF (A, ii). Conversely, Cu^{2+} binding prior to DEPC exposure abolishes monocarbethoxylation by any peptide functionality (B, i and ii).

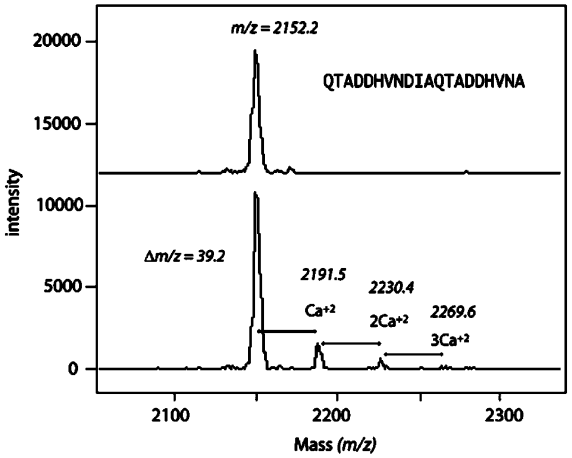

FIGURE 8: Calcium binding behavior of the synthetic peptide from the C-terminal undecapeptides of mcfp-4 assessed by MALDI-TOF MS.

The metal binding capacity of the C-terminal undecapeptide repeats was also assessed by a MALDI-TOF approach using a synthetic peptide based on the mcfp-4 sequence (residues 562−580). Whereas none of the transition metals (Cu, Ni, Zn, Co, and Fe) were detectably bound by the peptide, 1−3 equiv of bound Ca/peptide survived desorption and ionization (Figure 8 and the Supporting Information). Further work will be necessary to pinpoint the binding ligands.

Table 2: Comparison of mcfp-4 Domain Repeats with Known His- and Asp/Glu-Rich Repeats Detected in Other Proteins

protein	consensus sequence (no. of repeats)	metal implicated	ref
Mcfp-4	HQVLHGHVH (36)	Cu	this study
HMW kininogen	GH (12) R/KH (6)	Zn	*36*
HR glycoprotein	GHHPH (12)	Zn	*37, 38*
SR HCP	HRHRGH (9)	?	*39*
prion protein	PHGGGWGQ (4)	Cu	*23, 40*
nodulin	GH (10)	?	*41*
PfHRP 2	AHHAHHAAD (21)	Fe-heme	*28, 42*
Mcfp-4	DDHVN(D/N)IAQIA (16)	Ca	this study
SR HCP	EEEDEDDDEGD (9)	Ca	*38*
frustulin	CEGDCD (5)	Ca	*43*
EF hand	DKNGDGxxxxxE	Ca	*44*

DISCUSSION

mcfp-4 is localized to the apical portion of the byssal adhesive plaque where the splayed microfibrils of the thread core merge with the foam of the plaque (Figure 1). mcfp-4 is a highly repetitive and asymmetric protein with respect to composition and sequence. Of the total histidine, 84% is concentrated in the N-terminal half of the protein where it occurs four times on average in each of 36 decapeptide repeats in variant 1. In contrast, all of the Asp and 84% of the Asn are in the C-terminal half where they are prominent in a tandemly repeated peptide sequence 11 residues long with a DDHVN(D/N)IAQIA consensus sequence. Dopa occurs in the protein at less than 2 mol % and, on the basis of partial direct sequences, may be limited to the N- and C-termini as was the case for Dopa distribution in mefp-2 (*15*).

mcfp-4 joins a dozen or so known proteins with His-rich peptide repeats. Nearly all have demonstrated or proposed metal binding functions. In Table 2, the mcfp-4 consensus sequence is compared with other His-rich sequences known or proposed to bind metals. Most are involved in Zn/Cu binding, although plasmodial HRP-2 has been reported to bind iron heme (*28*). The repeat sequence of mcfp-4 is unique and unmatched in the sheer number of tandem repeats present. The preferential high-capacity Cu^{2+} binding ability of a synthetic peptide based on actual mcfp-4 sequence has been qualitatively established using MALDI-TOF mass spectrometry. Abolition of binding by prior modification of the peptides by DEPC emphasizes the role played by histidine.

The acidic consensus repeats at the C-terminus of mcfp-4 resemble repeat sequences of some known Ca binding proteins such as sarcoplasmic HRP given their high proportion of Asp/Asn and Glu/Gln (Table 2). However, in the best-characterized calcium binding proteins (EF hand and EGF domains), the binding is generally achieved by noncontiguous ligands in three-dimensional space (*29*). EGF domains are particularly widespread in proteins such as factor VI and fibrillin but also in mefp-2, which is an abundant plaque protein with 11 EGF repeats (*16*). EGF domains are endowed with Ca^{2+} binding ability (*29*) that is defined by a pentagonal bipyramidal geometry (seven ligands) involving Asp, Asn, Glu, and Gln side chains. A synthetic peptide 18 residues long based on the C-terminal repeat domain of mcfp-4 qualitatively exhibited a preference for binding one to three Ca^{2+} ions as detected by MALDI-TOF mass spectrometry.

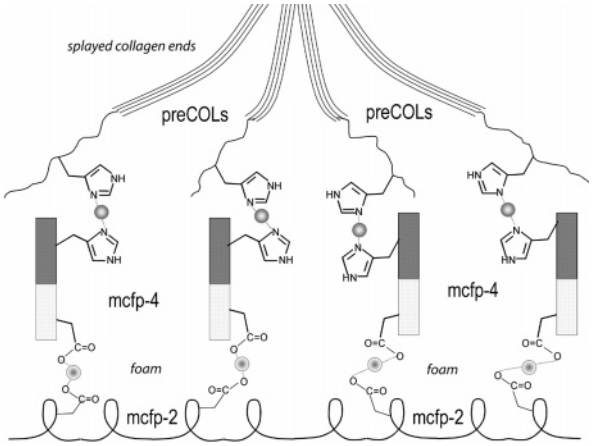

FIGURE 9: Proposed model of mcfp-4's role in joining the histidine-rich domains of the splayed preCOLs and other calcium binding proteins present in the adhesive foam. The participation of mcfp-2 is predicated on its abundance and the known calcium binding activity of its EGF domains. The involvement of the phosphorylated variants of mcfp-5 and -6 cannot be precluded.

mcfp-4 probably functions as a macromolecular bifunctional linker in the plaque−thread junction (Figure 9). On the basis of its location in the plaque, asymmetric organization, and metal ion binding behavior, mcfp-4 seems well suited to the role of coupling agent between the preCOLs in the frayed ends of the thread, on the one hand, and mcfp-2 (*16*) or the phosphorylated variants of mcfp-5 and -6 (*14*), on the other. The histidine-rich sequence in the N-terminal moiety of mcfp-4 appears to have a particularly high binding capacity for copper under MALDI conditions, e.g., up to 11 Cu ions per 2-mer. Perhaps during injection molding of a new thread, the histidine-rich repeats of mcfp-4 are presented to the preCOLs, which are also endowed with histidine-rich metal-binding domains at their termini. PreCOL-NG, for example, has a GGGHGGGHGGGRGGGH sequence at one end and an AHAHAHARAHAHA sequence at the other (*12*). Synthetic peptides of similar sequences exhibit copper binding in vitro (*30*). Transition metals such as Cu^{2+} coupled to the histidine ligands would serve as intermolecular bridges between the proteins (Figure 9). A similar relationship may apply to Ca^{2+} in coupling the C-terminal end of mcfp-4 to other calcium binding proteins in the plaque. The fate of Dopa in mcfp-4 is probably closely linked to that of the 5,5′-diDopa detected in the plaque (*31*) and thus represents an even more permanent link between plaque proteins.

The concept of connecting load-bearing proteins at junctions through coordinate or chelate interactions rather than covalent or noncovalent bonds is a not a new one. A high-resolution structure of a magnesium-mediated connection between integrin and collagen was recently reported (*32*, *33*). The high number of available binding sites in mcfp-4, however, is remarkable. The nearly 36 His-rich decapeptides and 16 Asp/Asn-rich undecapeptides should enable an excellent load distribution in mcfp-4 if all repeats participate equally in binding. mcfp-4 is an intriguing splicing element because it is so clearly multi-heterobifunctional; that is, it appears to be designed to connect to multiple domains on two different proteins. Although it remains to be shown that Cu^{2+} is the actual bridging metal in situ, it is certainly the best metal for mcfp-4 under the conditions prescribed for the MALDI assay.

Synthetic analogues of the consensus sequences of several His-rich proteins have unexpectedly provided effective metal ion binding functionalities for scaffolds used in the fabrication of nanowires and nanolithography (*34*, *35*). Given its involvement in an evolved load-bearing structure and remarkable capacity for copper binding, mcfp-4 may help to inspire a new generation of metallopolymers.

ACKNOWLEDGMENT

We thank Rachel Ngo for her assistance with byssal thread induction and analysis and Dr. Dong-Soo Hwang for helping construct the expression library of *M. californianus*.

SUPPORTING INFORMATION AVAILABLE

Isolation of mcfp-4 from foot tissue, isolation of tryptic peptides derived from mcfp-4 for sequencing, and MALDI mass spectrometric analysis of the behavior of metal ion binding by an acidic peptide. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Benjamin, M., Toumi, H., Ralphs, J. R., Bydder, G., Best, T. M., and Milz, S. (2006) Where tendons and ligaments meet bone, *J. Anat. 208*, 471−490.
2. Currey, J. D. (2002) *Bones: Structure and Mechanics*, Princeton University Press, Princeton, NJ.
3. Linke, W. A., Ivemeyer, M., Olivieri, N., Kolmerer, B., Rüegg, J. C., and Labeit, S. (1996) Towards a molecular understanding of the elasticity of titin, *J. Mol. Biol. 261*, 62−71.
4. Puxkandl, R., Zizak, I., Paris, O., Keckes, J., Tesch, W., Bernstorff, S., Purslow, P., and Fratzl, P. (2002) Viscoelastic properties of collagen: Synchrotron radiation investigations and structural model, *Philos. Trans. R. Soc. London 357*, 191−197.
5. Hasler, E. M., Herzog, W., Wu, J. Z., Müller, W., and Wyss, U. (1999) Articular cartilage biomechanics: Theoretical models material properties and biosynthetic response, *Crit. Rev. Biomed. Eng. 27*, 415−488.
6. Tamarin, A., Lewis, P., and Askey, J. (1976) Structure and formation of byssus attachment plaque in *Mytilus*, *J. Morphol. 149*, 199−221.
7. Benedict, C. V., and Waite, J. H. (1986) Ultrastructure and composition of the byssus of *Mytilus edulis* L, *J. Morphol. 189*, 261−270.
8. Waite, J. H., Holten-Andersen, N., Jewhurst, S., and Sun, C. J. (2005) Mussel adhesion: Finding the tricks worth mimicking, *J. Adhes. 81*, 297−317.
9. Sagert, J., Sun, C. J., and Waite, J. H. (2006) Chemical subtleties of mussel and polychaete holdfasts, in *Biological Adhesion* (Smith, A., and Callow, J., Eds.) Springer, New York.
10. Qin, X. X., Coyne, K. J., and Waite, J. H. (1997) Tough tendons: Mussel byssus has collagen with silk-like domains, *J. Biol. Chem. 272*, 32623−32627.
11. Waite, J. H., Qin, X. X., and Coyne, K. J. (1998) The peculiar collagens of mussel byssus, *Matrix Biol. 17*, 93−108.
12. Zhao, H., Robertson, N. B., Jewhurst, S., and Waite, J. H. (2006) Probing the adhesive footprints of *Mytilus californianus* byssus, *J. Biol. Chem. 281*, 11090−11096.
13. Zhao, H., and Waite, J. H. (2006) Linking adhesive and structural proteins in the attachment plaque of *Mytilus californianus*, *J. Biol. Chem.* (in press).
14. Rzepecki, L. M., Hansen, K. M., and Waite, J. H. (1992) Characterization of a cystine-rich polyphenolic protein family from the blue mussel *Mytilus edulis*, *Biol. Bull. 183*, 123−137.
15. Inoue, K., Takeuchi, Y., Miki, D., and Odo, S. (1995) Mussel adhesive plaque protein gene is a novel member of epidermal growth factor-like gene family, *J. Biol. Chem. 270*, 6698−6701.
16. Waite, J. H., and Qin, X.-X. (2001) Polyphosphoprotein from the adhesive pads of *Mytilus edulis*, *Biochemistry 40*, 2887−2893.
17. Waite, J. H., and Rice-Ficht, A. C. (1987) Presclerotized eggshell protein from the liver fluke *Fasciola hepatica*, *Biochemistry 26*, 7819−7825.

18. Waite, J. H., and Benedict, C. V. (1984) Assay of Dopa in invertebrate structural proteins, *Methods Enzymol. 107*, 397−413.
19. Paz, M., Flückinger, R., Boak, A., Kagan, H. M., and Gallop, P. M. (1991) Specific detection of quinoproteins by redox-cycling staining, *J. Biol. Chem. 266*, 689−692.
20. Sahal, D., Kannan, R., Sinha, A., Babbarwal, V., Gnana Prakash, B., Singh, G., and Chauhan, V. S. (2002) Specific and instantaneous one-step chemodetection of histidine-rich proteins by Pauly's stain, *Anal. Biochem. 308*, 405−408.
21. Waite, J. H. (1991) Detection of peptidyl-Dopa by amino acid analysis and microsequencing techniques, *Anal. Biochem. 192*, 429−433.
22. Qin, K.-F., Yang, Y., Mastrangelo, P., and Westaway, D. (2002) Mapping Cu(II) binding sites in prion proteins by diethyl pyrocarbonate modification and matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometric footprinting, *J. Biol. Chem. 277*, 1981−1990.
23. Follmer, C., and Carlini, C. (2005) Effect of chemical modification of histidines on the copper-induced oligomerization of jack bean urease (EC 3.5.1.5), *Arch. Biochem. Biophys. 435*, 15−20.
24. Nousiainen, M., Derrick, P. J., Kaartinen, M. T., Maepaa, P. H., Rouvinen, J., and Vainiotalo, P. (2002) A mass spectrometric study of metal binding to osteocalcin, *Chem. Biol. 9*, 195−202.
25. Weaver, J. L. (1998) Isolation, purification and partial characterization of a mussel byssal precursor protein mefp-4, Master's Thesis, University of Delaware, Newark, DE.
26. Miles, E. W. (1977) Modification of histidine residues in proteins by diethylpyrocarbonate, *Methods Enzymol. 47*, 431−442.
27. Choi, C. Y., Cerda, J. F., Chu, H. A., Babcock, G. T., and Marletta, M. A. (1999) Spectroscopic characterization of the heme binding sites in *Plasmodium falciparum* histidine-rich protein 2, *Biochemistry 38*, 16916−16924.
28. Rao, Z., Handford, P., Mayhew, M., Knott, V., Brownlee, G. G., and Stuart, D. (1995) The structure of a Ca^{2+} binding epidermal growth factor like domain: Its role in protein-protein interactions, *Cell 82*, 131−141.
29. Pappalardo, G., Impellizzeri, G., Bonomo, R. P., Campagna, T., Grasso, G., and Saita, M. G. (2002) Copper(II) and nickel binding in a histidine-containing model decapeptide, *New J. Chem. 26*, 593−600.
30. McDowell, L. M., Burzio, L. A., Waite, J. H., and Schaefer, J. (1999) REDOR detection of cross-links formed in mussel byssus under high flow stress, *J. Biol. Chem. 274*, 20293−20295.
31. Bella, J., and Berman, H. M. (2000) Integrin-collagen complex: A metal glutamate handshake, *Structure 8*, R121−R126.
32. Smith, C., Estavillo, D., Emsley, J., Bankston, L. A., Liddington, R. C., and Cruz, M. A. (2000) Mapping the collagen binding site in the domain of the glycoprotein Ia/IIa (integrin $\alpha 2\beta 1$), *J. Biol. Chem. 275*, 4205−4209.
33. Banerjee, I. A., Yu, L., and Matsui, H. (2003) Cu nanocrystal growth on peptide nanotubes by biomineralization: Size control of Cu nanocrystals by tuning peptide conformation, *Proc. Natl. Acad. Sci. U.S.A. 100*, 14678−14682.
34. Agarwal, G., Naik, R. R., and Stone, M. O. (2003) Immobilization of histidine tagged proteins on nickel by electrochemical dip pen nanolithography, *J. Am. Chem. Soc. 125*, 7408−7412.
35. Lin, Y., Pixley, R. A., and Colman, R. W. (2000) Kinetic analysis of the role of zinc in the interaction of domain 5 of HMW kininogen with heparin, *Biochemistry 39*, 5104−5110.
36. Koide, T., Foster, D., Yoshitake, S., and Davie, E. W. (1986) Amino acid sequence of human histidine rich glycoprotein derived from the nucleotide sequemce of its cDNA, *Biochemistry 25*, 2220−2225.
37. Morgan, W. T. (1985) The histidine rich glycoprotein of serum has a domain rich in histidine proline and glycine that binds heme and metals, *Biochemistry 24*, 1496−1501.
38. Hofmann, S. L., Goldstein, J. L., Orth, K., Moomaw, C. R., Slaughter, C. A., and Brown, M. S. (1989) Molecular cloning of a histidine rich Ca^{2+}-binding protein of sarcoplasmic reticulum that contains highly conserved repeat elements, *J. Biol. Chem. 264*, 18083−18090.
39. Kretzschmar, H. A., Stowring, L. E., Westaway, D., Stubblebine, W. H., Prusiner, S. B., and DeArmond, S. J. (1986) Molecular-cloning of a human prion protein cDNA, *DNA 5*, 315−324.
40. Pawlowski, K., Twigg, P., Dobritsa, S., Guan, C., and Mullin, B. C. (1997) Nodule specific gene family from *Alnus glutinosa* encodes glycine and histidine rich proteins expressed in early stages of actinorhizal nodule development, *MPMI 10*, 656−664.
41. Wellems, T. E., and Howard, R. J. (1986) Homologous genes encode 2 distinct histidine-rich proteins in a cloned isolate *Plasmodium falciparum*, *Proc. Natl. Acad. Sci. U.S.A. 83*, 6065−6069.
42. Kröger, N., Bergsdorf, C., and Sumper, M. (1994) A new calcium binding glycoprotein family constitutes a major diatom cell component, *EMBO J. 13*, 4675−4683.
43. Nakayama, S., and Kretsinger, R. H. (1994) Evolution of the EF-hand family of proteins, *Annu. Rev. Biophys. Biomol. Struct. 23*, 473−507.

BI061677N