Cement Proteins of the Tube-building Polychaete *Phragmatopoma californica**[S]

Hua Zhao[‡1], Chengjun Sun[‡1], Russell J. Stewart[§], and J. Herbert Waite[‡¶2]

From the ‡Molecular, Cell, and Developmental Biology Department, and the ¶Chemistry and Biochemistry Department, University of California, Santa Barbara, California 93106 and the §Department of Bioengineering, University of Utah, Salt Lake City, Utah 84112

The mineralized tube of the sandcastle worm *Phragmatopoma californica* is made from exogenous mineral particles (sand, shell, etc.) glued together with a cement secreted from the "building organ" on the thorax of the worm. The glue is a cross-linked mixture of three highly polar proteins. The complete sequences of Pc-1 (18 kDa) and Pc-2 (21 kDa) were deduced from cDNAs derived from previously reported peptide sequences (Waite, J. H., Jensen, R., and Morse, D. E. (1992) *Biochemistry* 31, 5733–5738). Both proteins are basic (pI ~10) and exhibit Gly-rich peptide repeats. The consensus repeats in Pc-1 and -2 are VGG$\underline{Y}$G$\underline{Y}$GGKK (15 times), and HPAVX-HKALGG$\underline{Y}$G (eight times), respectively, in which *X* denotes an intervening nonrepeated sequence and $\underline{Y}$ is modified to 3,4-dihydroxyphenyl-L-alanine (Dopa). The third protein, Pc-3, was deduced from the cement to be about 80 mol % phosphoserine/serine, and the cDNA was obtained by exploiting the presence of poly-serine repeats. Pc-3 consists of a family of at least seven variants with 60–90 mol % serine most of which is phosphorylated in the cement. Pc-1, -2, and -3 contain cysteine some of which reacts to form 5-*S*-cysteinyl-Dopa cross-links during the setting process.

The California sandcastle worm, *Phragmatopoma californica* (Fewkes), is a premier sand mason (1). In common with other sabellariid polychaetes, it exhibits an almost frenzied diligence in the collection, sorting and placement of sand grains for the construction and repair of its tubular home.[3] Although each worm builds primarily the tube in which it resides, a colony of worms can coordinate its efforts to erect massive boulder-like concretions that play a pivotal role in reef ecology (3, 4). The cement used by *Phragmatopoma* and related sabellariids to bind together grains of sand has been of interest for some time in that it adheres irreversibly to wet mineral surfaces and is used with extraordinary speed and economy. Perhaps 4–7 "spot welds," each about 100 μm in diameter, are used to hold each sand grain (diameter 500 μm) in place in the natural concrete (3, 5).

Phragmatopoma cement consists of proteins and significant levels of phosphate, calcium, and magnesium (6, 7). Two of the cement proteins, Pc-1 and Pc-2,[4] known from an earlier partial characterization (8), resemble the byssal adhesives of mussels (9) in that they are basic and contain 3,4-dihydroxyphenyl-L-alanine (Dopa) (5, 7, 8). Surprisingly, the abundant phosphate was not found to be associated with mineral but rather with serine residues in the cement (7). Indeed, the cement is dominated by phosphoserine and glycine, which together account for nearly 60 mol % of all the residues detected post-hydrolysis. Since the serine content of Pc-1 and -2 is negligible (8), the existence of a third serine-rich precursor is postulated.

The aim of the present research was to identify the serine-rich protein, to obtain full-length sequences of Pc-1 and Pc-2, and to gain some insights into the mechanism of cement solidification.

MATERIALS AND METHODS

Worm Maintenance for Tube Production—Colonies of *P. californica* were collected from the intertidal zone near Santa Barbara, CA and maintained in the laboratory with flowing filtered seawater and aeration. To collect tubes, several worms were removed from the colony with 1–2 cm of their original tubes intact and spaced out on a bed of 2-cm-thick clean sand grains in a plastic container (7). Commercial sand (grain size range 400–600 μm) (Sigma) was provided for new growth. The collected tubes were washed extensively with deionized water followed by several rinses of double deionized water, then either freeze-dried or blotted dry on tissue paper for immediate protein processing.

PCR and Preparation of cDNA—Total RNA was extracted from the cement gland in the thorax of *P. californica* with RNase plant mini kit from Qiagen (Valencia, CA) according to the supplier's protocols. First strand cDNA was synthesized from total RNA using Superscript II reverse transcriptase with Adapter Primer, 5′-GGC CAC GCG TCG ACT AGT ACT (T)$_{16}$ (Invitrogen), and used in subsequent PCR reactions. Based on specific amino acid sequences in Pc-1 and Pc-2 from a previous study (8) (*i.e.* VGGYGYGAK and WGHPAVHK, respectively), the 3′-ends of Pc-1 and Pc-2 were PCR amplified using 3′-rapid amplification of cDNA ends (3′-RACE) with degenerate oligonucleotides (sense Pc-1, 5′-GGN GGN TAY GGN TAY GGN GCN AA-3′; Pc-2, 5′-TGG GGN CAY CCN GCN GTN CAY AA-3′) and an abridged universal amplification primer (antisense 5′-GGC CAC GCG TCG ACT AGT AC-3′, Invitrogen). The PCR reaction was carried out in 25 μl of 1× Buffer B (Fisher), 5 pmol of each primer, 5 μmol of each dNTP, 1 μl of the first strand reaction mixture, and 2.5 units of *Taq* polymerase (Fisher) for 32 cycles on a Robocycler (Stratagene). Each cycle consisted of 30 s at 94 °C, 30 s at 52 °C, and 1 min at 72 °C, with a final extension time of 5 min. The PCR products were subjected to 1% agarose gel electrophoresis followed by gel purification and cloned into a PCR TA vector (TOPO TA cloning kit, Invitrogen). Plasmids were transformed

```
Pc-1: M= 18,240; pI = 9.74
MKFVIVLALVSAA YGCG
VGIGCAGGRC --GGACGG-K --GYGYGG-K LGYGAYGKGG IGGYGYGKGC
VGGYGYGGLG -------AGK LGGYGYGGSK CGGYGYGGQK LGGYGYGGKK
LGGYGYAAKK VGGYGYGAKK VGGYGYGAKK VGGYGYGAKK VGGYGYGAKK
VGGYGYGAKK VGGYGYGAKK VGGYGYGVKK VGGYGYG


Pc-2: M = 21,116; pI= 9.91
MKVLIFLATVAAVYG CGGAGG WRSGSCGG
RWGHPAV----HKALGGYG-G
YGAHPAVHAAVHKALGGYGAGAYGAG
AWGHPAV----HKALGGYGAGA
WG-HPAV----HKALGGYG-G
YGAHPAVHVAVHKALGGYGAGACGHKTGGYGG
YGAHP---VAV-KA--AY-NHGFNYGANNAIKSTKRFGG
YGAHP---V-VKKAFSRGLSHGAY-AG
SKAATGYGYGSGKAAGGYGY


Pc-3A; M = 13,979
MKLLSVFAIVVLAVYITHVEA
DSSSSSYSSSSSYSSSSSSSSSSSSSYSSSSSYSSSSSSSY
SSSSSYSSSSSYSSSSYSSSSYSSSSSYSSSSILTSTSSSDW
KRKVPARRVLRTRRFLKCVTRCTLRCILFRSAKTCARKCSR
RCLKRVF


Pc-3B; M = 30,525
MKSFTIFAAILVALCYIQISEAG
CCKRYSSSSYSSSSSSSSSSSYSSSSSSSSSYSSSSSSSSSSYSSSSSSSSSSSYSSSSSSY
SSSSSSSYSSSSSSSSSSSYSSSSSYSSSSSSSSSSYSSSSSSSSSSYSSSSSSSSSSSYS
SSSSSYSSSSSSSSYSSSSSSSSSSSYSSSSSYSSSSSSSSSSYSSSSSSSSSSYSSSSSSS
SYSSSSSYSSSSSSSSSSSYSSSSSSSSSSSYSSSSSYSSSSSSSSSSYSSSSSSSSS
YSSSSSSSSSSYSSSSSSSSSSSYSSSSSYSSSSSSSYSSSSSSSYSSSSSSSSSSSYSSSSSSSY
SSSSSSSSSSSYSSSSSSSSSSSSSSSYSSS
```

FIGURE 1. **Pc-1, -2, and -3 sequences deduced from *P. californica* cement gland cDNA and aligned to highlight internal repeats.** *Italicized* initial sequence denotes the signal peptide; *underlined* portions indicate known peptide sequence (8). Deduced protein sequences were assigned the following GenBank[TM] accession numbers AY960614 (Pc1), AY960615 (Pc2), AY960616 (Pc3A-1), AY960617 (Pc3A-2), AY960618 (Pc3A-3), AY960619 (Pc3B-1), AY960620 (Pc3B-2), AY960621 (Pc3B-3), AY960622 (Pc3B-4).

into competent Top10 cells for amplification, purification, and sequencing. The insert encoded the COOH-terminal sequence of Pc-1 and Pc-2, respectively, including the 3′-untranslated region.

To obtain the 5′-ends of the Pc-1 and Pc-2 cDNA, the GeneRacer kit (Invitrogen) was used to obtain sequence information from full-length transcripts by 5′-RACE. PCR was conducted under the same condition as above with gene specific primers (antisense Pc-1, 5′-TTT CCC CCG AAT TGG TAA TTA ATA CCC-3′; Pc-2, 5′-AGA CTT TGG TGT GAG GAA TTA ATA GCC-3′) and a GeneRacer 5′ primer.

A cDNA library was constructed from the mRNA extracted from the cement gland in the thorax of *P. californica* using the CloneMiner™ cDNA library construction kit (Invitrogen) and adapted for serine-rich protein screening. Initial screening was done by PCR using a degenerate oligonucleotide corresponding to Ser[5] (sense 5′-GAA TTC AGY AGY AGY AGY AG-3′ with engineered EcoR I site) and a vector-specific universal primer, T7 5′-AATACGACTCACTATAG-3′. PCR conditions, cloning, and sequencing involved the same strategy outlined above are shown in the supplemental material. After obtaining the 3′ sequences of Pc-3A and Pc-3B, nondegenerate gene-specific primers were designed to amplify the gene sequence with 5′ RACE strategy. To ensure integrity of each cDNA, antisense primers (Pc-3A-reverse, 5′-C TCA ATG GCC TTG AAC CTA GAA TAC-3′; Pc-3B-reverse, 5′-ACA TAT AAG TCG TGT AAA TCT ATT TCT AAC-3′) were designed within the 3′-untranslated region to amplify the full-length sequence. The PCR products were subcloned into a PCR TA vector and sequenced as described above.

In Situ Hybridization—Worms were carefully removed from their tubes, anesthetized in 33% magnesium sulfate, then fixed, dehydrated,

and embedded in methylmethacrylate (9:1 BMA with 1% benzoyl peroxide for thermal cure) as described in Warren *et al.* (10). Embedded tissue was microtomed into 2-μm-thick sections and acetone-de-embedded. Sections were then rehydrated, washed with diethylpyrocarbonate-treated Q-H$_2$O, digested with proteinase (10 μg/ml proteinase K) for 4 min at 37 °C, acetylated with 0.1 M triethanolamine, 0.5% acetic anhydride for 3 min, and blocked with blocking solution (1 × Denhardt's solution, 5% dextran sulfate, 0.2 mg/ml sheared herring sperm DNA, 4 × SSC, and 50% formamide) for 2 h at 42 °C. DIG-labeled oligonucleotide was then added to the blocking solution, and the tissue sections were hybridized overnight at 42 °C. The DIG-labeled oligonucleotide was prepared by DIG-PCR labeling kit (Roche Diagnostics) using primers sense 5′-ATGAAATCCTTCACTATTTTTGCC-3′ and antisense 5′-AGAGCTGGAACTAGAGCTGTA-3′. Negative control reactions for later *in situ* hybridization included regular dUTP instead of DIG-labeled dUTP. Hybridization product was visualized by incubating the sections with anti-DIG-AP and subsequently adding color substrates nitro blue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate as described by the supplier (Roche Diagnostics) and viewed by light microscopy. Parallel whole worms were stained for Pc-1 and -2 using the Arrow reaction (11).

Cement Analysis—For routine analysis, cement proteins were hydrolyzed in 6 M HCl and 5% phenol *in vacuo* for 24 h at 110 °C. Phosphoserine losses were corrected by extrapolation of hydrolysis to zero time (7). Hydrolysates were flash evaporated to dryness at 50 °C and subjected to amino acid analysis on a Beckman 6300 Autoanalyzer using an 85-min elution program for post-translationally modified amino acids (12). Cysteine was detected as carboxymethylcysteine following reduc-

tion by dithiothreitol and alkylation with iodoacetate as previously described (13), and phosphoserine was detected as such at 2.7 min (7). For cross-link analysis, worm cement proteins on sand grains were hydrolyzed for 1–2 h at conditions otherwise described above. Dopa and cysteinyldopa were purified from flash evaporated cement hydrolysates using phenyl boronate affinity chromatography (Affi-Gel 601 Boronate, Bio-Rad). Hydrolysate residues resuspended in 100 mM sodium phosphate buffer (pH 7.5) were applied to the boronate column (14). Bound ligands were washed with 10 column volumes of phosphate buffer, followed by 10 column volumes of 2.5 mM NH_4HCO_3 and 10 column volumes of deionized distilled water. Fractions eluted with 5% acetic acid were freeze-dried and subjected to amino acid analysis (12) and electrospray ionization mass spectrometry (Micromass QTOF2 tandem mass spectrometer) using a syringe pump to inject samples at a rate of 5 μl/min. All sample analyses were compared with authentic 2-*S*-cysteinyl-Dopa and 5-*S*-cysteinyl-Dopa.

RESULTS

Using the known peptide sequences from two cement precursors (Pc-1 and Pc-2) as the basis for gene-specific degenerate oligonucleotides, complete sequences were deduced from the cDNAs obtained by reverse transcriptase-PCR of cement gland extracts followed by 3′-and 5′-RACE. Representative and variant sequences are shown in Fig. 1 and in supporting data, respectively. Both Pc-1 and -2 are basic proteins with calculated pI values at 9.7 and 9.9, respectively, that are consistent with those measured directly by isoelectric focusing (8). Three variants of Pc-1 were sequenced, and all have a mass of about 18 kDa and consist mostly of three amino acids, glycine, lysine, and tyrosine, the last of which is extensively modified to Dopa (8). All three contain 15 repeats of a consensus decapeptide *V*GGYGYGG*K*K, in which the italicized residues are occasionally substituted (Fig. 1*A*).

With respect to Pc-2, only one variant (21 kDa) was found. This also contains degenerate copies of a consensus repeat, HPAVHKALGGYG, but with considerable variety in connecting and flanking sequences (Fig. 1*B*). Again, most of the tyrosine is converted to Dopa in the mature protein (8). While the high level of glycine was consistent with the compositional bias of the cement, Pc-1 and -2 are deficient in the other most abundant amino acid of cement, namely serine and/or phosphoserine, thus suggesting another precursor with a strong compositional bias (TABLE ONE).

The average composition of a phosphoserine/serine-rich cement precursor was deduced by subtracting the compositions of Pc-1 and Pc-2 from the composition of whole cement (TABLE ONE). Three assumptions were made in this calculation: (*i*) that the glycine in cement is contributed only by Pc-1 (>45 mol %) and Pc-2 (>29 mol %), (ii) that histidine in cement is derived from Pc-2, and (iii) that Pc-1 and Pc-2 each contribute a third to the composite whole. The accuracy of this approach is compromised mostly by the considerable amounts of tyrosine in the Pcs known to be lost first to Dopa, then to Dopaquinone, and finally to cross-links and/or oxidation products that ultimately involve collateral losses of other amino acids (15–17) as well. Given the Ser/Ser(P) content in cement at 28.5 mol %, the Ser contributions of Pc-1 and Pc-2 are 0.6/3 and 3.7/3 Ser/Ser(P), respectively, or about 1.5 parts out of 28.5. The difference (28.5–1.5 = 27 and 27 × 3 = 81 mol %) suggests that a precursor containing more than 80 mol % Ser/Ser(P) is reasonably predicted for the third cement protein.

Attempts to extract a protein with ∼80 mol % phosphoserine (Pc-3) directly from the cement gland or from cement deposited by captive worms on acid washed sand remain inconclusive. Possible reasons for this include the insolubility of cement (phosphoserine is associated with

TABLE ONE

Amino acid compositions in mol % (residues per 100 residues) determined by amino acid analysis of *P. californica* cement and deduced from cDNA for the precursor proteins, Pc-1, Pc-2, and Pc-3

Amino acid	Cement	Pc-2	Pc-1	Pc-3[a]
Asp + Asn	2.8	2.1	0	0.8
Thr	2.2	1.6	0	2.2
Ser	28.5	3.7	0.6	72.9
Ser(P)[b]	(>25)			(>64)
Glu + Gln	1.4	0	0.6	0
Pro	2.7	3.7	0	0.3
Gly	26.2	27.5	45.0	0.3
Ala	9.8	19.0	6.8	0.9
Cys/2 or Cys[c]	0.4	1.6	3.6	1.0
Val	3.4	6.4	6.0	1.2
Met	0	0	0	0
Ile	0.6	0.5	1.2	0.6
Leu	3.4	3.2	3.0	2.2
Dopa[b]	2.1	(7.3)	(9.8)	(?)
Tyr	4.0	9.0	18.7	10.0
Phe	1.1	1.6	0	0.9
His	3.5	9.0	0	0
Lys	4.4	6.9	13.9	2.2
Trp	ND[d]	2.1	0	0.3
Arg	2.9	2.1	0.6	4.1
TOTAL	99.9	100	100	100

[a] Pc-1 and Pc-3 compositions calculated from the average composition of seven variants for Pc-3 and three for Pc-1. The cement composition represents an average of three different hydrolyses. Standard deviations are approximately 10%.
[b] Parenthetical levels of Ser(P) and Dopa are from previous studies (7, 8) and represent the repeats of initial Ser and Tyr modified to Ser(P) and Dopa, respectively.
[c] Detected as cystine (Cys/2) post-hydrolysis in cement but as cysteine for the deduced compositions of the three Pc proteins.
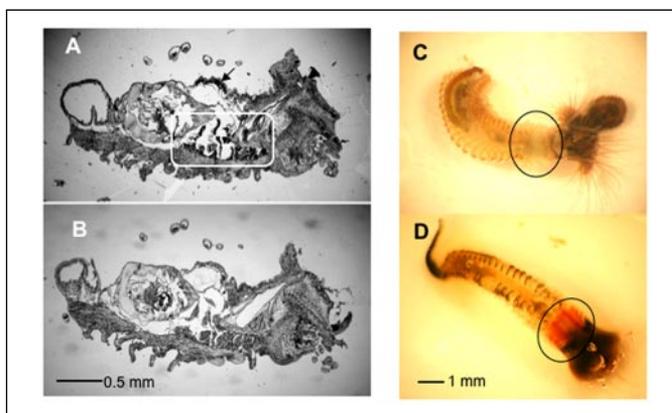[d] Not determined.

FIGURE 2. *A*, in situ reverse transcriptase-PCR of a sagittally sectioned whole *P. californica* using DIG-labeled gene-specific primers for Pc-3. *B*, negative control. Live worm was in the same approximate orientation as in *A*. The tentacular crown is to the *right*, and the cement gland is *circled* (*C*); *D*, sagittally sectioned worm with cement gland stained with the Dopa-specific Arnow reaction. The *scale bar* is as indicated.

the insoluble fractions) and also that phosphoserine-rich proteins are notoriously difficult to stain following polyacrylamide gel electrophoresis (18). In planning a molecular strategy for deducing sequence of Pc-3s, the following two points were exploited: 1) at 80 mol % serine, the probability of at least one stretch of five consecutive serines must be high, and 2) the codon preference for serine in Pc1 and -2 was AGT/C. Combining the two points, primers based on $(AGT/C)_4AG$ coupled with sequence from the 3′-untranslated region sufficed after RACE to
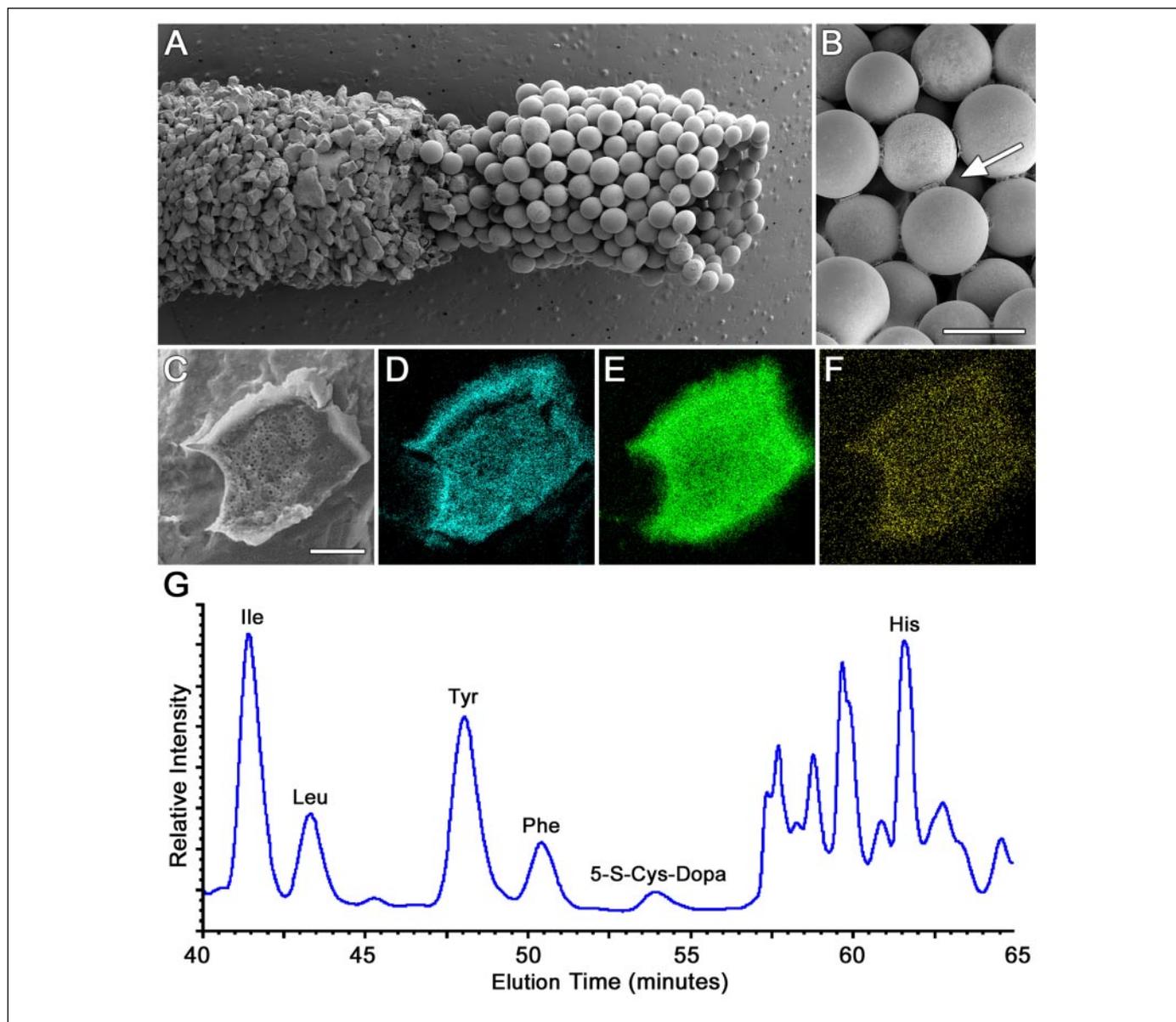
FIGURE 3. **Sulfur and the composition of the cement of *P. californica*.** *A*, S.E. of new growth in a *P. californica* tube with glass beads (0.5 mm diameter). *B*, zoom of *A* revealing cement plaques between glass beads (*arrow*); *C*, top view of a fractured cement contact between two sand grains; *D–F* have the same view as *C* with localization of sulfur (*F*) within the cement using energy dispersive x-ray spectroscopy. Added for comparison are carbon (*D*) and phosphorus (*E*). *Scale bars* are 100 μm each. *G*, amino acid analysis of hydrolyzed cement highlighting the aromatic region of the chromatogram. Peak at 54.5 min coelutes with standard 5-*S*-cysteinyl-Dopa.

discover the Pc3 sequences (Fig. 1*C*). The runs of uninterrupted serines range from Ser_4 to Ser_{13} in length. Following further screens of cement gland cDNA with primers based on the serine rich domains, two types of putative Pc-3 variants emerged: Pc-3A variants are between 50 and 60 mol % serine and contain a highly basic carboxyl terminus with six cysteines. In contrast, the Pc-3B variants lack the basic amino acids at the carboxyl terminus and have serine levels approaching 90 mol %. Calculated masses range from 10 to 30 kDa, with an expected increase of 5–25 kDa due to phosphorylation alone. The average mol % for Ser (/Ser(P)) for all seven variants was about 73 mol % (Pc-3, TABLE ONE), which compares reasonably well with the 80 mol % predicted from the cement analysis. Calculated pI values of fully phosphorylated Pc-3s would range from 0.5 to 1.5, which would place them among the most acidic proteins known.

That Pc-3 proteins are variants specifically associated with the cement gland of *P. californica* was demonstrated by *in situ* hybridization

of digoxigenin-labeled Pc-3 using a gene-specific sequence 120 bp in length in the 5′-untranslated region (Fig. 2*A*). The label was localized to the cement gland in the thorax. In a parallel study, a worm was specifically stained for Dopa, which serves as a convenient marker for Pc-1 and -2 (Fig. 2*D*). Only the thorax stained with the Dopa reagent.

Two intriguing features of the deduced sequences shown in Fig. 1 are the persistent cysteine content (range: 1–3.5%) with a cement average of a little over 2 mol % and a tyrosine content of 10 mol %. Analysis by energy dispersive x-ray spectroscopy showed the cement to be moderately endowed with sulfur (Fig. 3*F*), and acid hydrolyzed cement subjected to amino acid analysis contained cystine at 0.4 residue/100 residues (TABLE ONE). The latter, however, is less than half the expected level in terms of cysteine. To further explore the fate of cysteine, cement deposited by the worms onto acid-washed sand was collected and subjected to iodoacetate alkylation with or without prior reduction by dithiothreitol. No carboxymethylcys-
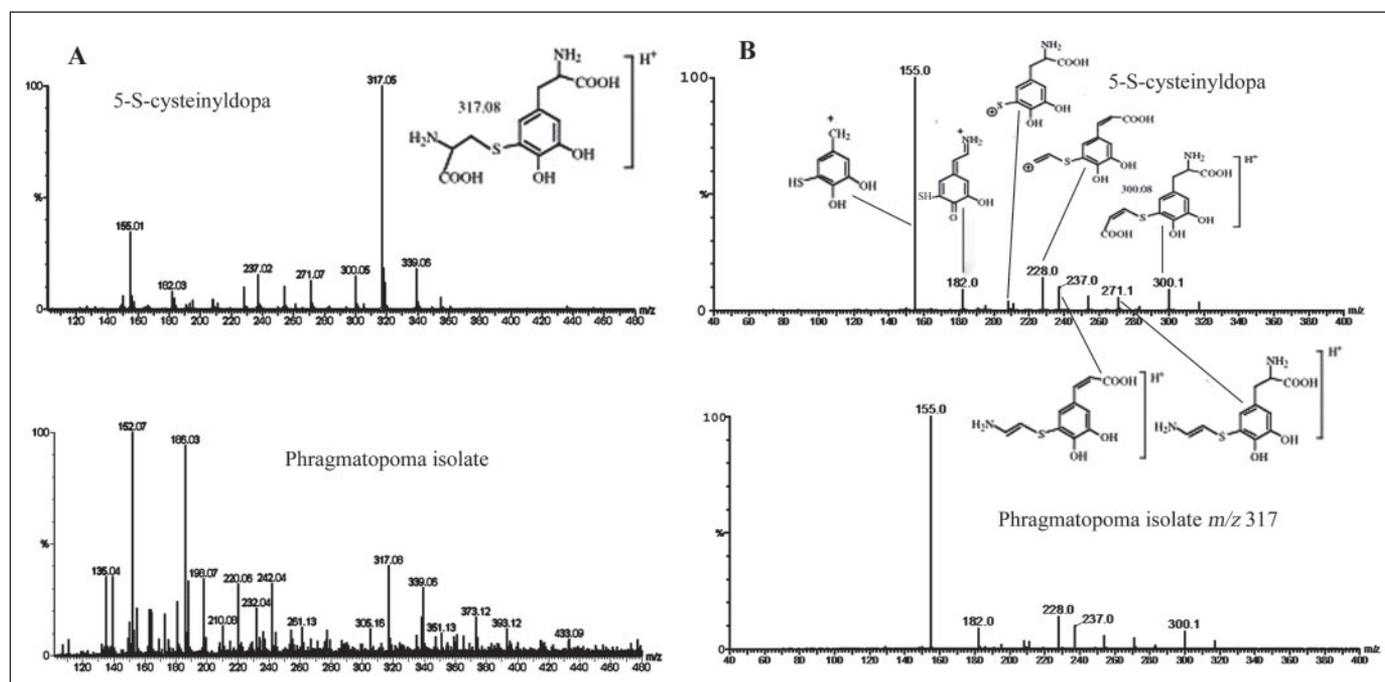
FIGURE 4. *A*, mass spectrometry by ESI TOF mass spectrometry of standard 5-*S*-cysteinyl-Dopa (*upper panels*, protonated, 317.05; sodiated, 339.06) and the peak fraction following affinity chromatography of hydrolyzed cement on phenylboronate-agarose. *B*, collision-induced decomposition and tandem mass spectrometry of the *m/z* 317 peak of standard 5-*S*-cysteinyl-Dopa (*top*) and isolate from hydrolyzed *P. californica* cement (*bottom*). Proposed fragment ion structures are as shown.

teine could be detected without prior reduction. With reduction, the detected carboxymethylcysteine (0.91 + 0.18 mol %, $n = 3$) was comparable with the cystine (0.44 + 0.16 mol %, $n = 3$) found in unreduced hydrolyzed samples.

To account for the missing half of predicted cysteine, we postulated that Pc-1, -2, and -3 might be secreted with both thiolate and disulfide bonded cysteines. The thiolates could be rapidly scavenged by Dopaquinones formed upon oxidation of Dopa residues at seawater pH (pH 8.2) and thus become the basis of an effective protein cross-linking chemistry (17, 19, 20). Using amino acid analysis, phenylboronate affinity chromatography, and electrospray ionization mass spectrometry, significant levels of 5-*S*-cysteinyl-Dopa were detected in hydrolysates of *Phragmatopoma* cement (Fig. 3, *G* and *H*). The compound isolated from hydrolyzed cement plaques binds tightly to phenylboronate, has an elution time of about 54 min by amino acid analysis, a mass of 316.1 Da by ESI mass spectrometry, and a fragmentation profile by tandem ms following collision-induced decomposition that is consistent with authentic 5-*S*-cysteinyl-Dopa (Fig. 4). All of the observed fragments have counterparts to a previously studied pentafluoropropionyl/methylated derivative of 5-*S*-cysteinyl-Dopa following fast atom bombardment ms (21) and are identified accordingly in Fig. 5. It is worth noting that although the 2- and 5-*S*-isomers are distinguished by their elution times on amino acid analysis, the 2-*S* also has distinct features on tandem ms; the intensity of the 228 and 182 peaks is consistently at 100 and 60%, respectively (results not shown).

A 5-*S*-cysteinyl-Dopa cross-link density of one per 100 amino acids was estimated from the amino acid analysis. Given the known tendency of cysteinyl-Dopa to reoxidize and undergo additional nucleophilic additions (22), this density must be considered an absolute minimum.

DISCUSSION

Phragmatopoma cement represents perhaps the simplest permanent bioadhesive investigated to date. In contrast to mussels and barnacles,

which endeavor to attach themselves directly to the substratum, *Phragmatopoma*, which remains mobile within its tube, seeks only to make a concretion of sand grains. Despite this, Pc proteins exhibit some similarities with previously characterized marine adhesive proteins. Like almost all the mussel byssal adhesive proteins, Pc-1 and -2 contain high levels of Dopa and lysine. A few mussel adhesive proteins are Gly-rich like Pc-1; fp-1 from *Aulacomya ater*, for example, has a consensus repeat of AGYGGVK (23), which is shorter but shares many of the functionalities of the VGGYGYGGK repeat. Pc-2 is more unique, and no matches for its consensus peptide could be found in Swiss-Prot. Only one other *Mytilus* adhesive protein, fp-3, has comparable tryptophan content (24).

Of the cement proteins, Pc-3 is most unlike other known adhesive precursors. Few proteins or domains exhibit the compositional serine/phosphoserine bias of Pc-3 especially Pc3B which exceeds 90 mol % Ser. Chick phosvitin has a phosphoserine rich domain that contains several runs of serine as long as 14 residues (25). Phosphophoryns, matrix proteins from tooth dentin with the consensus sequence (DpSpS)$_n$, are another example (26, 27). Both of these are involved in binding to Ca^{2+} ions and amorphous calcium phosphate and/or hydroxyapatite.

The successful sand masonry of *P. californica* is likely to be largely determined by the adhesive properties of its cement and the manner of its dispensation. Four conditions are widely considered to be prerequisites for effective practical adhesion: 1) the absence of weak boundary layers, 2) good spreading of the adhesive over the surface, 3) formation of extensive interfacial interactions, and 4) uniform setting or curing of the adhesive (28). Because these prerequisites were not formulated for underwater conditions, a further prerequisite is necessary and that is delivery of the cement/adhesive as a fluid that is nondispersible by the seawater medium. This will be considered first.

Pc-1, -2, and -3 are water-soluble polyelectrolyte solutes, thus secreting them together or sequentially would risk loss by dilution to the surrounding seawater. To overcome this problem, Stewart *et al.* (7) proposed a model based on complex coacervation. Complex coacerva-
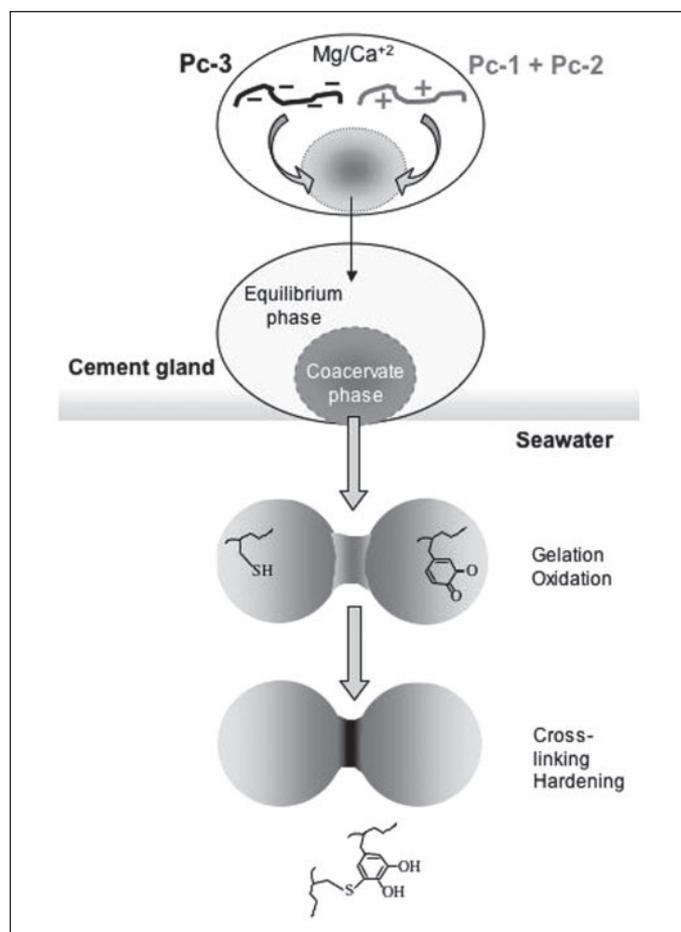
FIGURE 5. **Coacervated cement model with cross-linking.** In complex coacervation, polyions are mixed in solution in the cement gland at a pH where the net charge is zero leading to phase separation (7). Dollops of coacervate are then released from the cement gland and applied to one or more spots on each sand grain. The dollops gel shortly after each sand grain is positioned onto the tube wall and solidify due to cross-link formation. The gelling is attributed to the lower solubility of magnesium and calcium ions with phosphate groups at seawater pH (7) and the cross-linking to adduct formation between oxidized Dopa and cysteine groups in proteins.

tion refers to a liquid:liquid phase separation that occurs when polycations such as basic proteins are mixed with polyanions such as acidic proteins at a pH where there is charge equivalence (29, 30). When equivalence occurs in a symmetrical mixture of polyelectrolytes with similar charge density, molecular weight, and flexibility, a denser concentrated phase (coacervate) separates from the more dilute equilibrium phase. Coacervates have unusual properties that would seem to be able to satisfy prerequisites 1 and 2, that is, removing weak boundary layers and exhibiting good or spontaneous spreading. Coacervate proteins, by virtue of being more desolvated than noncoacervated proteins (31), may be able to absorb water from wet surfaces. In addition, given the low interfacial tension between the coacervate and equilibrium solution phases, coacervates tend to spread readily over most surfaces (29).

Pc-1, -2, and -3 ostensibly resemble commercial coacervate precursors (29) in being highly charged and flexible polymers in aqueous solution, but their influence on coacervation when mixed in roughly equal concentrations must be specifically considered. Let us assume that mixing is done at pH 5, which is thought to be the intragranular pH of a mature regulated secretory granule (32). Pc-1 and -2 are polycationic with a charge densities at pH 5 of about +0.2 and +0.35/residue and molecular masses of 18 and 21 kDa, respectively. Pc-3 charge density, assuming complete phosphorylation of serines, is higher at −0.4 to

-0.8/residue at pH 5.0, and Pc-3 masses are phosphorylation-dependent ranging from 10 to 30 kDa for unphosphorylated variants and 18 to 52 kDa for fully phosphorylated forms. Given the dominance of negative charges in Pc-3s, it is clear that charge equivalence could not be achieved with the polyelectrolytes alone. Addition of microions such as Mg^{2+} and Ca^{2+} to the mixture could help overcome the charge imbalance but high salt also suppresses coacervation (30). Cytochemical studies of cement producing cells in the cement glands of the tubeworms have established that Mg^{2+} and Ca^{2+} are colocalized with phosphate in mottled granules (6, 33), but it is not known whether Pc-1 and -2 are also present. The emerging picture is one that contrasts starkly with coacervation based on a symmetrical mixture. The necessity of microion addition for charge equivalence, high polydispersity in molecular weight and negative charges (via phosphorylation), and the effect of divalent ion binding on Pc-3 backbone flexibility all portend a multiphase coacervate morphology. Although this is consistent with the porous structure of the cement (7) and of the mottled cement precursor granules (3, 33), a fundamental understanding of the process will require much more study.

With regard to the third prerequisite of adhesion, given the high polarities of Pc proteins, extensive electrostatic and van der Waals interactions with the surface of the substratum are inevitable but their interaction strength would be diminished by the high dielectric constant of seawater. Pc proteins, however, are also capable of interfacial contacts that are independent of the dielectric constant of seawater. These are most notably the interaction of Dopa with surface oxides, which results in a coordinate bond (34), and the interaction of phosphoserines with minerals such as iron oxides and apatite to form insoluble ionic bonds (35).

Curing results from the formation of intermolecular cross-bridges (fourth prerequisite) in the adhesive. If future studies determine that the cysteinyldopa cross-links in cement are intermolecular, then this would fulfill the setting requirement. In mussel byssal adhesive plaques, di-Dopa cross-links prevail (15), but whether they also occur in *Phragmatopoma* cement is not known. A role for phosphoserines in setting is suggested by the high levels of Mg/Ca^{2+} in the cement and cement gland (6, 7, 33) and cement softening by EDTA treatment.[5] Indeed, Ca/Mg^{2+} interactions with phosphates represent an excellent pH triggered type of setting. At low pH the interaction would be largely electrostatic, whereas at pH 8, given the K_{sp} (calcium phosphate) $= 2.02 \times 10^{-33}$, it would become ionic and precipitate (36).

Fig. 5 proposes the key steps in *Phragmatopoma* cement formation. Some or all cement precursors are stockpiled together with Mg/Ca as multiphase coacervates in secretory cell granules of the cement gland. The granule coacervate contents are released onto the surface of a sand grain where they coalesce and spread. Because the interactions between Mg/Ca^{2+} and phosphate groups are less soluble at seawater pH, the cement becomes less fluid and more gel-like. Finally, Dopa residues not coordinated at the interface oxidize to Dopaquinones that react with cysteines to form irreversible cysteinyldopa cross-links.

We know of no other instance in which coacervation is exploited in adhesion. The setting by Ca^{2+}, however, shares a striking parallel with a common commercial glue (*e.g.* Elmer's White Glue^{TM}), which consists of a phosphoprotein (casein) that can be rendered water-resistant by the addition of lime (Ca(OH)$_2$) at alkaline pH (2). The adhesive advantages gained by *Phragmatopoma* cement in having an order of magnitude more phosphate groups than casein and a signif-

---

[5] C. Sun, unpublished data.

icant inclusion of polycations with Dopa groups (Pc-1 and -2) remain to be explored.

REFERENCES

1. Sisson, R. F. (1986) *Nat. Geographic* **169,** 252–255
2. Salzberg, H. K. (1977) in *Handbook of Adhesives* (Skeist, I., ed) pp. 158–171, Van Nostrand Reinhold Co., New York
3. Vovelle, J. (1965) *Arch. Zool. Exp. Gen.* **106,** 1–187
4. Chisholm, J. R. M., and Kelley, R. (2001) *Nature* **409,** 152–153
5. Jensen, R., and Morse, D. E. (1988) *J. Comp. Phys.* **158B,** 317–324
6. Gruet, Y., Vovelle, J., and Grasset, M. (1987) *Can. J. Zool.* **65,** 837–842
7. Stewart, R. J., Weaver, J. C., Morse, D. E., and Waite, J. H. (2004) *J. Exp. Biol.* **207,** 4727–4734
8. Waite, J. H., Jensen, R., and Morse, D. E. (1992) *Biochemistry* **31,** 5733–5738
9. Waite, J. H. (2002) *Integ. Comp. Biol.* **42,** 1172–1180
10. Warren, K. C., Coyne, K. J., Waite, J. H., and Cary, S. C. (1998) *J. Histochem. Cytochem.* **46,** 149–155
11. Waite, J. H. (1995) *Methods Enzymol.* **258,** 1–20
12. Waite, J. H. (1991) *Anal. Biochem.* **192,** 429–433
13. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) *Anal. Chem.* **68,** 850–858
14. Hawkins, C. J., Lavin, M. F., Parry, D. L., and Ross, I. L. (1986) *Anal. Biochem.* **159,** 187–190
15. McDowell, L. M., Burzio, L. A., Waite, J. H., and Schaefer, J. (1999) *J. Biol. Chem.* **274,** 20293–20295
16. Burzio, L. A., and Waite, J. H. (2000) *Biochem.* **39,** 11147–11153
17. Dryhurst, G., Kadish, K. M., Scheller, F., and Renneberg, R. (1982) *Biological Electrochemistry*, Vol. 1, pp. 138–139, Academic Press, New York
18. Fisher, L. W., and Termine, J. D. (1998) in *Current Advances in Skeletogenesis* (Ornoy, A. Harell, A., and Sela, J., eds) pp. 467–472, Elsevier Publishing Co., New York
19. Ito, S., Kato, T., Shinpo, K., and Fujita, K. (1984) *Biochem. J.* **222,** 407–411
20. Takasaki, S., and Kawakishi, S. (1997) *J. Agr. Food Chem.* **45,** 3472–3475
21. Agrup, G., Hansson, C., Rorsman, H., Rosengren, A.-M., and Rosengren, E. (1976) *Commun. Dept. Anat. Univ. Lund* **5,** 885–892
22. Ito, S., Inoue, S., Yamamoto, Y., and Fujita, K. (1981) *J. Med. Chem.* **24,** 673–677
23. Burzio, L. A., Saez, C., Pardo, J., Waite, J. H., and Burzio, L. O. (2000) *Biochim. Biophys. Acta* **1479,** 315–320
24. Papov, V. V., Diamond, T. V., Biemann, K., and Waite, J. H. (1995) *J. Biol. Chem.* **270,** 20183–20192
25. Byrne, B. M., van Het Schip, A. D., van de Klundert, J. A. M., Arnberg, A. C., Gruber, M., and Ab, G. (1984) *Biochemistry* **23,** 4275–4279
26. Ritchie, H. H., and Wang, L.-H. (1996) *J. Biol. Chem.* **271,** 21695–21698
27. George, A., Bannon, L., Sabsay, B., Dillon, J. W., Malone, J., Veis, A., Jenkins, N. A., Gilbert, D. J., and Copeland, N. G. (1996) *J. Biol. Chem.* **271,** 32869–32873
28. Schonhorn, H. (1981) in *Adhesion in Cellulosic and Wood-based Composites* (Oliver, J. F., ed) pp. 91–111, Plenum Publishing Corp., New York
29. Bungenberg de Jong, H. G. (1949) in *Colloid Science* (Kruyt, H. R., ed) Vol. 2, pp. 433–482, Elsevier Publishing Co., Inc., Amsterdam
30. Veis, A. (1970) in *Biological Polyelectrolytes.* (Veis, A., ed) Vol. 3, pp. 211–273, Marcel Dekker, New York
31. Ohno, H., Shibayama, M., and Tsuchida, E. (2003) *Makromol. Chem.* **184,** 1017–1024
32. Johnson, R. G. (1988) *Physiol. Rev.* **68,** 232–307
33. Vovelle, J., and Grasset, M. (1990) *Cah. Biol. Mar.* **31,** 333–348
34. Dalsin, J. L., Lin, L., Tosatti, S., Vörös, J., Textor, M., and Messersmith, P. M. (2005) *Langmuir.* **21,** 640–664
35. Boffardi, B. P. (1993) *Mater. Perform.* **32,** 50–53
36. Kuboki, Y., Fujisawa, R., Aoyama, K., and Sasaki, S. (1979) *J. Dent. Res.* **58,** 1926–1932