

# Identification, Characterization, and Expression Levels of Putative Adhesive Proteins From the Tube-Dwelling Polychaete *Sabellaria alveolata*

PIERRE T. BECKER<sup>1</sup>, AURÉLIE LAMBERT<sup>1</sup>, ANNABELLE LEJEUNE<sup>2</sup>,  
DÉBORAH LANTERBECQ<sup>3</sup>, AND PATRICK FLAMMANG<sup>1,\*</sup>

<sup>1</sup>Université de Mons—UMONS, Laboratoire de Biologie des Organismes Marins et Biomimétisme, 23 Place du Parc, 7000 Mons, Belgium; <sup>2</sup>Université de Liège, GIGA, Laboratoire de Biologie Moléculaire et de Génie Génétique, 1 Avenue de l'Hôpital, 4000 Liège, Belgium; and <sup>3</sup>Université de Mons—UMONS, Laboratoire Interfaces et Fluides Complexes, 23 Place du Parc, 7000 Mons, Belgium

**Abstract.** The shelter of the tube-dwelling polychaete *Sabellaria alveolata* is composed of mineral particles assembled with spots of a proteinaceous cement. The adhesive proteins constituting the cement were identified on the basis of their sequence similarity with proteins of a phylogenetically related species, *Phragmatopoma californica*. Two positively charged proteins, Sa-1 and Sa-2, share common features: they both have a mass of 22 kDa; are rich in glycine, tyrosine and basic residues; and show repeated peptide motifs. The consensus repeat of Sa-1 is KGAY-GAKGLGYGNKAGYGAYG (occurring 6–8 times), while Sa-2 displays the consensus heptapeptide VHKAAWG (5 times) and undecapeptide VHKAAGYGGYG (8 times). Two variants of a serine-rich protein, Sa-3A (22 kDa) and Sa-3B (21 kDa), were also identified. Their serine residues account for 75 mol% and are probably phosphorylated, meaning that Sa-3 is very acidic and negatively charged. Moreover, tyrosine residues of all adhesive proteins are presumably modified into DOPA. Although protein sequences are not well-conserved between *S. alveolata* and *P. californica*, their main characteristics (including amino acid composition, post-translational modifications, repeated patterns, isoelectric point, and mass) are shared by both species. This suggests that these features are more important for their function than the primary structure of the proteins. The mRNA abundance for each protein was estimated by quantitative real-time PCR, revealing relative expression levels

of about 5, 11, 1.5, and 1 for Sa-1, -2, -3A, and -3B, respectively. These levels could be indicative of charge neutralization phenomena or could reflect their function (interface *vs.* bulk) in the cement.

## Introduction

*Sabellaria alveolata* (family Sabelliariidae) is a gregarious tube-dwelling polychaete from Western Europe that builds its tube by assembling sand grains and mollusc shell fragments with an organic cement. Sabelliariidae are commonly called sandcastle worms or honeycomb worms because tubes constructed by a huge number of individuals are closely imbricated to form large reef-like mounds that, in *S. alveolata*, may extend over several hectares and reach heights of up to 1.8 m (Noernberg *et al.*, 2010). In *S. alveolata*, the head of the worm bears numerous thin white tentacles that collect the mineral particles in their surroundings and convey them to the building organ, a horseshoe-shaped structure situated near the mouth (Vovelle, 1965). This organ is made up of two lobes that manipulate and select suitable particles for construction. These particles are then dabbed with small spots of a proteinaceous adhesive and placed on the outer edge of the tube. The adhesive is produced by cement cells located in the thoracic area and possessing long cell processes ending near the tip of the building organ lobes, where the glue is secreted (Vovelle, 1965). Two types of cement cells are recognized on the basis of the morphology of their secretory granules, homogeneous or heterogeneous, the latter being characterized by conspicuous intra-granular inclusions (Vovelle, 1965; Gruet *et al.*, 1987).

Received 21 Dec 2011; accepted 9 August 2012.

\* To whom correspondence should be addressed. E-mail: Patrick.Flammang@umons.ac.be

In *Phragmatopoma californica*, a related sabellariid species from the Californian coast that presents a similar histological organization of the cement cells (Wang *et al.*, 2010), several adhesive proteins have been characterized (Jensen and Morse, 1988; Waite *et al.*, 1992; Zhao *et al.*, 2005; Endrizzi and Stewart, 2009; Wang and Stewart, 2012). Five major proteins, named Pc-1 to -5, were identified that showed significant expression levels (Wang and Stewart, 2012). Pc-1 (18 kDa) is a glycine and tyrosine-rich protein displaying a 10 amino acid repeated motif and a basic isoelectric point (pI) due to the abundance of lysine residues. Pc-2, Pc-4, and Pc-5 (20, 27, and 15 kDa, respectively) are all basic, histidine-rich proteins with repeated motifs. Pc-3 occurs in the form of two variants (Pc-3A and Pc-3B) both characterized by the over-expression of serine residues that account for 60–90 mol% of the protein. Most of these serine residues are phosphorylated, and Pc-3 is therefore extremely acidic with a calculated pI ranging from 0.5 to 1.5 (Stewart *et al.*, 2004; Zhao *et al.*, 2005). Noteworthy, tyrosine residues of both Pc-1 and Pc-2 are post-translationally modified into 3,4-dihydroxyphenylalanine (DOPA) by the addition of a hydroxyl group (Waite *et al.*, 1992). In the cement, the modified amino acids present in the different adhesive proteins (*i.e.*, DOPA and phosphoserine [pSer]) mediate physicochemical interactions with the substratum (adhesion) and are also involved in setting and curing reactions (cohesion) (Zhao *et al.*, 2005; Sun *et al.*, 2007; Stewart *et al.*, 2011). Curing is presumably catalyzed by a co-secreted tyrosinase (Wang and Stewart, 2012). Immunostaining and *in situ* hybridization experiments demonstrated that Pc-1 to -5 are produced by cement cells and together form the adhesive material (Wang and Stewart, 2012). These analyses also showed that the polyanionic proteins (*i.e.*, Pc-3 variants) mix poorly with the basic cement proteins (*i.e.*, Pc-1, -2, -4, and -5), invalidating the previous hypothesis of complex coacervation (separation of two oppositely charged polyelectrolytes into two immiscible aqueous phases when the charges of the polyelectrolytes are balanced) as a model for the formation of the glue (Wang and Stewart, 2012).

To date, no information is available about the composition of the cement in *S. alveolata*, but extensive anti-pSer immunolabeling was observed in the parathoracic cement cells as well as in their processes located in the building organ, indicating that these cells would secrete a polyphosphoprotein homologous to the protein Pc-3 (Flammang *et al.*, 2009). The aim of the present study was therefore to identify this polyphosphoprotein together with other putative adhesive proteins in *S. alveolata*. This was done by targeting and amplifying the corresponding cDNAs with primers designed from the protein sequences Pc-1, -2, and -3, available for *P. californica*. These sequences, which represent the first three proteins identified in the American species, were chosen because they are among the most

abundantly expressed in the cement glands (Wang and Stewart, 2012), and because they are each representative of one category of adhesive proteins (*i.e.*, respectively the GY-rich, the H-repeat, and the SY-rich categories; Endrizzi and Stewart, 2009). The expression level of each protein identified in *S. alveolata* was also estimated by quantitative real-time PCR analyses.

## Material and Methods

### Sampling

Reef fragments of *Sabellaria alveolata* Linnaeus, 1767 (Polychaeta, Sabellariidae) were collected at low tide at Champeaux, Mont-Saint-Michel Bay, France (N48°43'50, W01°33'05). They were transferred to the University of Mons (Belgium) and kept in a marine aquarium with closed circulation (14 °C, 33‰ salinity).

### RNA extraction and reverse transcription

For each extraction, three worms (about 200 mg) were removed from their tube, quickly frozen in liquid nitrogen, and directly ground in 4 ml of TRIzol Reagent (Invitrogen). Total RNA was then extracted and purified through successive incubations in bromochloropropane, isopropanol, and 75% ethanol before being dissolved in 70% ethanol with 80 mmol l<sup>-1</sup> sodium acetate. For reverse transcription, cDNA was obtained with the Transcriptor First Strand cDNA Synthesis kit (Roche) using 1 µg of total RNA and following the manufacturer's instructions.

### PCR—3' and 5' RACE

For each protein, a central fragment of the corresponding cDNA was amplified by PCR using 1 µl of the total cDNA and the Ready-To-Go PCR Beads kit (GE Healthcare); 5' and 3' extremities were obtained with the FirstChoice RLM-RACE (Ambion) with subsequent PCR amplifications. PCR conditions included an initial denaturation step of 5 min at 95 °C followed by 35 cycles with a 30-s denaturation step at 95 °C, a 30-s annealing step at 40 °C, and a 90-s elongation step at 72 °C. The last step consisted in a final extension of 10 min at 72 °C. Table 1 lists the primers used in PCR experiments. For the central fragment, the primers were designed according to the cDNA sequences coding for the adhesive proteins of *P. californica* already available in databases (Pc-1, -2, -3A, -3B; GenBank accession numbers AY960614 to AY960622; Waite *et al.*, 1992; Zhao *et al.*, 2005). Finally, as sequence variability was observed among the cDNAs from *S. alveolata*, a full-length sequence of the cDNA coding for each adhesive protein was amplified using primers matching the 5' and 3' untranslated regions (UTR).

Table 1

Primers used to amplify cDNAs coding for the adhesive proteins (f: forward primer, r: reverse primer)

Protein	Fragment	Primers sequences
Sa-1	Central	5'-GTTCTCGCTTTGGTCTCTGC-3' (f) 5'-GGCACCATATCCATATCC-3' (r)
	N-terminal end	5'-GGCTCCTCTCCCGTTCCAA-3'
	C-terminal end	5'-AGCTAAAGCAGCAGGATATG-3'
	Full length	5'-GTTTAAACTCATCCGAAA-3' (f) 5'-GCCGAAAATTGATTGTCG-3' (r)
	Central + C-terminal end	5'-GCAGTACATAAAGCATTAGGAG-3' (f) 5'-CTTTTCGCATCCGCTTT-3' (r)
Sa-2	N-terminal end	5'-GCATCCTCCTTTACTACC-3'
	Full length	5'-ATCAAACCTTGATCTATCAACC-3' (f) 5'-CTTTTCGCATCCGCTTT-3' (r)
	Central + C-terminal end	5'-AGCAGTAGCAGTAGTTACAGC-3' (f) 5'-GCATTTTCTTGCACATGTTTT-3' (r)
Sa-3A	N-terminal end	5'-GAAACGCAAGTTCATTCCAAA-3'
	Full length	5'-GCATCACTTGGAAATATA-3' (f) 5'-GCTTACACCTAGCTTGTT-3' (r)
	Central + C-terminal end	5'-AGCAGCAGCTTAGCAGTTCA-3' 5'-AGAGCTGCTGCTGTAGCT-3'
	Full length	5'-GAATTCATGCACAAACCATACTC-3' 5'-GAGTATGGTTTGTGCATGAAATTC-3'
Sa-3B	Central	5'-GTTTCTACTTCAGTCTGG-3' (f) 5'-GGTGATACTTTACTACAAA-3' (r)
	N-terminal end	5'-GAGTATGGTTTGTGCATGAAATTC-3'
	C-terminal end	5'-GTTTCTACTTCAGTCTGG-3' (f)
	Full length	5'-GTTTCTACTTCAGTCTGG-3' (f) 5'-GGTGATACTTTACTACAAA-3' (r)

### Cloning and sequencing

The PCR products were purified on agarose gel with the Quantum Prep kit (Biorad) and cloned into TOP10 chemically competent *Escherichia coli* cells using the TOPO TA cloning kit (Invitrogen). Plasmids were then isolated with the GeneJET Plasmid Miniprep kit (Fermentas), and their insert size was checked by PCR using the universal M13F and M13R primers. Clones of the expected size were sequenced with the BigDye Terminator ver. 3.1 Cycle Sequencing kit (ABI) in a 3730 DNA analyzer (ABI). The cycle sequencing reaction consisted in 25 cycles with a 10-s denaturation step at 96 °C, a 5-s annealing step at 48 °C, and a 4-min elongation step at 60 °C. For each protein, a total of 21 clones of the full-length cDNA were analyzed and translated into amino acids using the Translate tool of the ExPASy Proteomics Server. A consensus sequence, obtained with Clustal X (Thompson *et al.*, 1994), was then deduced from the alignment of these 21 sequences. Finally, signal peptides were predicted with the SignalP 3.0 Server (Center for Biological Sequence Analysis) while pI and molecular weight values were calculated using Scansite 2.0 (Obenauer *et al.*, 2003). The sequences obtained in this study have been deposited in the EMBL database under accession numbers HE599563 to HE599646.

### Histochemical and immunohistochemical analyses

Worms were first fixed overnight in non-acetic Bouin's fluid and dehydrated in a sequence of graded ethanol. They

were then embedded in paraffin wax (Paraplast, Sigma-Aldrich), and transverse sections (7 µm) through the head or the thorax were cut with a Microm HM 340E microtome. The presence of DOPA in the cement cells was investigated by using the nitro blue tetrazolium (NBT) redox cycling method (Paz *et al.*, 1991) and Arnow's staining (Arnow, 1937), while the cationic dye methyl green was used to stain polyanionic macromolecules (Gabe, 1968).

For immunohistochemistry, sections were subjected to an indirect immunoenzymatic staining method according to the following protocol. After antigen demasking by microwaving sections in citrate buffer (10 mmol l<sup>-1</sup>, pH 6.0), non-specific background staining was blocked by section pre-incubation in phosphate-buffered saline containing 0.5% bovine serum albumin (Sigma, St. Louis, MO) (PBS-BSA). Monoclonal anti-phosphoserine antibodies (clone PSR-45, mouse ascites fluid; Sigma, St. Louis, MO, ref. P 3430), diluted 1:500 in PBS-BSA, were then applied for 1 h at room temperature. After several washes in PBS, immunoreactivity was visualized by an immunoperoxidase technique using the ImmPRESS polymerized reporter enzyme staining system (Vector Laboratories, Burlingame, CA) and 3,3'-diaminobenzidine tetrahydrochloride (Sigma, St. Louis, MO) as a chromogen substrate. The sections were lightly counterstained with hemalum and luxol blue. Control reactions were performed by substituting the primary antibody with PBS-BSA and/or by using the primary antibody saturated with its specific antigen (antibodies diluted

in PBS containing 2  $\mu\text{g/ml}$  pSer-conjugated BSA [Sigma, St. Louis, MO, ref. P 3717]).

All sections were observed with a Zeiss AxioScope A1 microscope equipped with an AxioCam ICc 3 digital camera (Carl Zeiss MicroImaging, Göttingen, Germany).

### Quantitative real-time PCR

Real-time PCR was carried out on a StepOnePlus instrument (ABI) with the Power SYBR Green PCR Master Mix (ABI). Five RNA extractions followed by reverse transcriptions were performed as described above, and cDNA samples were diluted to reach a concentration of 5  $\text{ng}/\mu\text{l}$ . Only freshly collected animals were used. For each adhesive protein, the number of copies of the resulting cDNA was evaluated in the five samples by using a standard curve. Primers amplifying the full-length sequence of the corresponding cDNAs were used at a final concentration of 0.2  $\mu\text{mol l}^{-1}$ . Reactions (20  $\mu\text{l}$ ) were run in triplicate with 2  $\mu\text{l}$  of diluted total cDNA as template. Conditions were as follows: after an initial denaturation at 95  $^{\circ}\text{C}$  for 10 min, 40 cycles were performed with a denaturation step at 95  $^{\circ}\text{C}$  for 20 s, an annealing step at 47  $^{\circ}\text{C}$  for 30 s, and an elongation step at 72  $^{\circ}\text{C}$  for 75 s. For standard curves, the four cDNAs were first cloned in plasmids as described above, and a 10-fold serial dilution series of the plasmids, ranging from  $10^2$  to  $10^7$  copies/ $\mu\text{l}$ , was used. Specificity of the amplification reactions was analyzed using melt curves with a temperature range from 60  $^{\circ}\text{C}$  to 95  $^{\circ}\text{C}$ . The significance of the data was evaluated using the Wilcoxon test.

## Results

Four putative adhesive proteins were identified in *Sabelaria alveolata*, namely Sa-1, Sa-2, Sa-3A, and Sa-3B. For each protein, 21 clones were obtained and showed a significant variability among sites when aligned. This variability reached 8.1% for Sa-1 (22.4% when considering the 21-residue-long deletion and insertion of clones 10 and 20, respectively), 15.2% for Sa-2, 18.8% for Sa-3A, and 18.0% for Sa-3B. Detailed features of these proteins are described hereafter, based on their consensus sequence.

### Sa-1

To amplify a first fragment of the cDNA coding for the protein Sa-1, forward and reverse primers were designed according to peptide sequences from the protein Pc-1: respectively its signal peptide and the GYGYGA sequence derived from its repeated motif (*i.e.*, XGGYGYGAK). The consensus translated sequence of Sa-1 contains a 15-aa-long signal peptide and a 219-aa-long mature polypeptide (20.7 kDa, Fig. 1A). The mature protein is rich in glycine (39.3%), alanine (17.4%), tyrosine (16%), and lysine (13.2%), and four cysteines are present in the N-terminal part of

### A. Sa-1

```
MKVLVLLALVSAVYECGWNRGACGGGACGKACQAKAYGFGGKAGYGVY
KGAYGAKGLGYGNKPAAYGGFG KGAYGAKGLGYGNKPAAYGGFG
RGAYGAKGLGYGGKASYGAYG KGAYGAAGKFGYGNKAGYGAYG
KGAYGAKGLGYGNKVGYGAIG KGAYGAKGLGYGSKGGYGAIG
KGAYGAKGLGYGSKGGYGAIG KAKVGYGGYGAAGAAYGTRKGGYGAYGKAYGKTYG
```

### B. Sa-2

```
MKLLILVALAASISGCGGAGCGGGWGAAGKGGACGGRWGGKGGCG AHKAAWG
VHKAAWG AGHGYGG VHKAAWG AGHGYGV VHKAAG FGYG AHKAAWG
HKAAGYGAHGYYG VHKAAGYGGYG VHKAAGLGGYG VHKAAGYGGYG
VHKAAGLGGYG VHKAAGYGGYG VHKAAGYGGYG VHKAAGLGGYG AHKAAGYGGYG
AGAHKAAGYGYGAHKAAGYGYGAHKAAGYGGYG
```

### C. Sa-3A

```
MKTIVLFAFVLLALCLVHGEARAYGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
```

### D. Sa-3B

```
MKTFAVFAFVLLALCLVHGEARAYGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
```

**Figure 1.** Sa-1, Sa-2, Sa-3A, and Sa-3B consensus sequences obtained from the alignment of 21 clones. Signal peptides are highlighted in gray and internal repeats of Sa-1 and Sa-2 are underlined. For Sa-2, motifs  $X_1\text{HKAAX}_2\text{G}$  (where  $X_1$  is V or A, and  $X_2$  is W or absent) and  $X_1\text{HKAAGX}_2\text{GGYG}$  (where  $X_1$  is V or A, and  $X_2$  is Y or L) are shown with continuous and dotted lines, respectively. For Sa-3, the 12-aa-long serine-free peptides differentiating the two variants are underlined.

the sequence. The calculated pI of Sa-1 is 9.85, meaning that this protein is basic and positively charged in seawater (pH 8.2). Moreover, a 21-residue-long pattern is repeated 7 times and presents the following consensus sequence: KGAYGAKGLGYGNKAGYGYGAYG. Noteworthy, among the 21 cDNA clones sequenced, one coded for a protein variant presenting an 8th repetition, whereas another coded for a variant displaying only 6 repetitions.

### Sa-2

While performing 3'RACE experiments to obtain the sequence of the cDNA coding for Sa-1, it appeared that some clones were not related to the C-terminal end of Sa-1 but to the extremity of a histidine-rich protein, similar to the protein Pc-2 of *Phragmatopoma californica*, which was named Sa-2. A primer matching the 3' UTR of the cDNA coding for Sa-2 was then designed and used in combination with a primer based on a repeated pattern of Pc-2 (*i.e.*, AVHKALGG). This allowed the retrieval of most of the cDNA sequence of Sa-2, and this sequence was completed by 5' RACE. The translated sequence of Sa-2 comprises an 18-residue-long signal peptide and a 219-aa-long mature polypeptide that, besides histidine (9.1%), also contains numerous glycine (35.6%), alanine (23%), tyrosine (11%), and lysine (8.7%) residues (20.5 kDa, Fig. 1B). The mature protein has a basic pI of 9.76 and presents four cysteines at its N-terminus. Two repeated motifs are detected in Sa-2: (1) V\*HKAAW\*G (in which V\* and W\* can also be A and absent, respectively), present 5 times at the N-terminus of

the protein, and (2) V\*HKAAGY\*GGYG (in which V\* and Y\* can also be A and L, respectively), found 8 times in the middle and C-terminal parts of the sequence (Fig. 1).

### Sa-3

Two variants of Sa-3 were identified and named Sa-3A (19.4 kDa, Fig. 1C) and Sa-3B (18.9 kDa, Fig. 1D). The cDNA coding for the central and C-terminal parts of Sa-3A was amplified using a forward primer coding for a peptide containing 6 serines and one tyrosine (SSSSSY) and a reverse primer designed from the C-terminal end of Pc-3A (Table 1). For Sa-3B, both primers recognized a 6-serine cluster, plus another serine for the forward and one tyrosine for the reverse. Sa-3A and Sa-3B both possess a 21-aa-long signal peptide, and their mature sequences (207 and 195 aa, respectively) are characterized by the abundance of serine residues that account for the great majority of the amino acids (74.4% and 73.3%, respectively). Sa-3A mainly differs from Sa-3B by the sequence of a 12-aa-long, serine-free peptide present in the middle of the protein. Another difference is the number of tyrosine residues: 12 for Sa-3B but only 4 for Sa-3A. Calculated pI ranges from 9.46 (Sa-3B) to 10.88 (Sa-3A). However, considering a 95% phosphorylation rate, as in *P. californica* (Stewart *et al.*, 2004), the pI would drop down to extremely acidic values of 1.11 (Sa-3A) and 1.21 (Sa-3B).

### Post-translational modifications

On transverse sections of the worm thorax, the adhesive glands were clearly stained with both the NBT and the Arnou methods (Fig. 2A–C). The NBT reaction indicates the presence of redox active compounds while the Arnou staining specifically labels catechols, including DOPA. With both methods, no distinction could be made between the cement cells containing homogeneous granules and those containing heterogeneous granules (Fig. 2B).

Methyl green, a dye that stains phosphorylated polysaccharides in onuphid polychaetes (Winsnes, 1985), labeled cement cells (Fig. 2D). In this case, however, staining intensity was much more important in cement cells containing heterogeneous granules than in those containing homogeneous granules (Fig. 2D). In the former, the most strongly labeled structures are the intra-granular inclusions (Fig. 2E). Immunohistochemical analyses using anti-pSer antibodies were then performed to confirm that the methyl green staining was due to the presence of a polyphosphoprotein (see also Flammang *et al.*, 2009). An intense labeling was observed at the level of the cement cells (Fig. 2F, G). As in the methyl green staining pattern, immunoreactivity was much more important in cells containing heterogeneous granules than in cells containing homogeneous granules (Fig. 2G); and again, the most strongly labeled structures were the intra-granular inclusions (Fig. 2H). This labeling was com-

pletely abolished in control reactions without primary antibodies or with antibodies saturated with pSer-conjugated BSA (Fig. 2I).

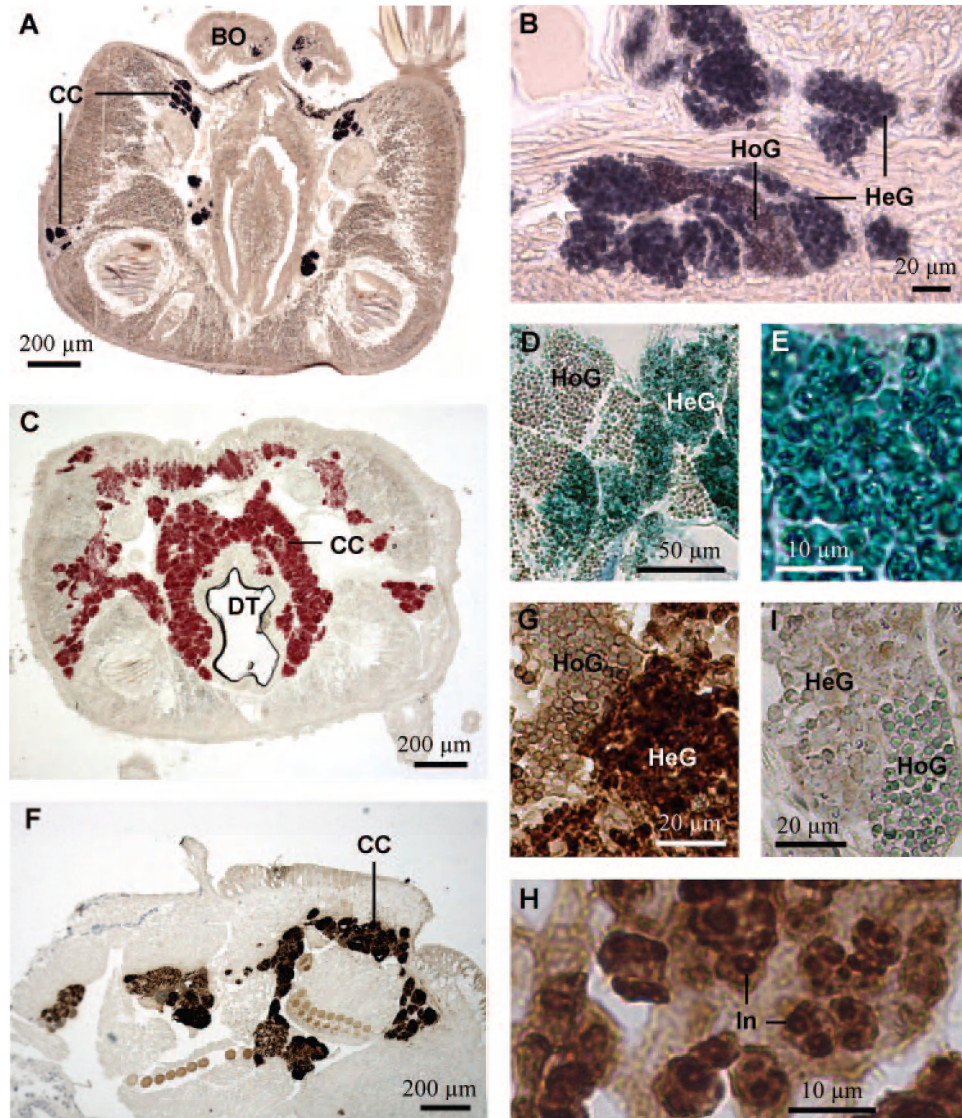
### Quantitative real-time PCR analyses

In the five samples, the expression levels of the four adhesive proteins displayed the same pattern: Sa-2 was the most abundantly expressed protein, followed in decreasing order by Sa-1, Sa-3A, and Sa-3B (Fig. 3). Considering an mRNA level of 1 for Sa-3B, the expression levels of Sa-3A, Sa-1, and Sa-2 are respectively  $1.6 \pm 0.4$ ,  $4.9 \pm 1.5$  and  $11.3 \pm 3.7$  (mean  $\pm$  standard deviation).

Sa-2 is 2.3-fold more expressed than Sa-1, although this difference is not statistically significant ( $P > 0.05$ ) because of the large standard deviation. The mRNA levels of Sa-3A and Sa-3B are also similar ( $P < 0.05$ ), the expression of Sa-3A being only 1.6-fold more important than that of Sa-3B. Sa-1 is 3.1-fold more expressed than Sa-3A, but this difference results in a borderline significance ( $P = 0.056$ ). On the other hand, its expression is 4.9-fold higher than that of Sa-3B, this difference being statistically well supported ( $P < 0.05$ ). Finally, the mRNA level of Sa-2 is significantly different from those of Sa-3A ( $P < 0.05$ ) and Sa-3B ( $P < 0.01$ ), with a respective 7.2 and 11.3-fold more important expression. Basic proteins (Sa-1 and Sa-2) are thus significantly more expressed (about 6.2-fold more,  $P < 0.01$ ) than acidic proteins (Sa-3A and Sa-3B).

### Discussion

The stonemason abilities of sabellariids have been known for centuries (Réaumur, 1711), but it is only recently that some of the secrets of their cement have been unravelled, due to its potential to inspire the development of novel biomimetic underwater adhesives (see, *e.g.*, Stewart *et al.*, 2011). Two species have attracted most of the attention in terms of tube building: *Sabellaria alveolata* and *Phragmatopoma californica*. The former was investigated mainly in terms of tube structure, building organ morphology, and cement elemental analysis (Vovelle, 1965; Gruet *et al.*, 1987; Fournier *et al.*, 2010); the latter was described in terms of building organ morphology and micro- and nano-structure as well as composition of the cement (Stewart *et al.*, 2004; Zhao *et al.*, 2005; Stevens *et al.*, 2007; Wang *et al.*, 2010; Wang and Stewart, 2012). Among sabellariids, *P. californica* is the only species in which cement proteins have been characterized. The biochemical study of the adhesive in this species was initiated through extraction of a DOPA-rich fraction from the thorax of the worms. Subsequent analyses of the DOPA-containing peptides allowed the identification of Pc-1 and Pc-2, two putative cement proteins (Jensen and Morse, 1988; Waite *et al.*, 1992). However, these two proteins were poor in serine residues even though the amino acid composition of the cement

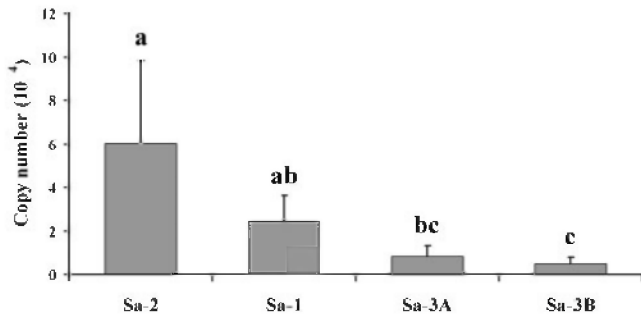


**Figure 2.** Histochemistry and immunohistochemistry of cement cells in *Sabellaria alveolata*. (A, B) NBT staining of a transverse section through a worm head (A) with detail of a group of cement cells (B). Redox active compounds appear in dark blue. (C) Arnow staining of a transverse section through a worm thoracic region. Reactive structures are labelled red. (D, E) Cement cells stained with methyl green. Staining is mostly restricted to the intra-granular inclusions of cells containing heterogeneous granules (E). (F–I) Immunolabelling (in brown) using anti-pSer antibodies. General view of a transverse section through a worm thorax (F), cement cell bodies (G), and detail of the heterogeneous granules (H). The labeling of cement cells disappears in the control reaction when antibodies are saturated with pSer-conjugated BSA (I). BO, building organ; CC, cement cell; DT, digestive tube; HeG, heterogeneous granules; HoG, homogeneous granules; In, inclusions.

revealed a high content of this amino acid (Jensen and Morse, 1988). The presence of a serine-rich protein was therefore suspected, and the cDNA sequences of two variants of Pc-3 were retrieved from PCR experiments with primers matching serine repeats (Zhao *et al.*, 2005). More recently, the transcriptome of the adhesive gland of *P. californica* was analyzed through the construction and sequencing of a cDNA library (Endrizzi and Stewart, 2009; Wang and Stewart, 2012). It appeared that Pc-1, Pc-2, and

Pc-3 were actually representatives of three protein families, respectively the GY-rich, the H-repeat, and the SY-rich protein families, each containing several different members.

The methodology used to identify the adhesive proteins of *S. alveolata* was based on the phylogenetic relationship and morphological similarity of this species with the American species, *P. californica* (Dales, 1952; Rouse and Pleijel, 2001). The primary sequences of the proteins Sa-1, -2, and -3 were obtained directly from cDNAs amplified using



**Figure 3.** Absolute quantification of the expression levels for the four adhesive proteins. Data expressed as the mean copy number ( $n = 5$ ) contained in 10 ng of cDNA. The absolute quantification is based on standard curves with  $R^2$  values of 0.984, 0.986, 0.994, and 0.990 for Sa-1, Sa-2, Sa-3A, and Sa-3B, respectively. Significant differences between the means are indicated by letters; means sharing at least one letter are not significantly different ( $P > 0.05$ ).

primers designed from their homologs Pc-1, -2, and -3, therefore allowing the extraction, purification, and characterization of the proteins to be bypassed. Proteins from the two sabellariid species share common features. Like Pc-1, Sa-1 is a small, basic protein rich in glycine, lysine, tyrosine, and alanine residues, and it presents a repeated amino acid pattern. But this repeated motif, consisting of 21 amino acids, is different from the one occurring in *P. californica*, which is only 10 residues long. Another similarity between Sa-1 and Pc-1 is the presence of 4 cysteine residues at the N-terminus. These residues could be involved in cement curing through the formation of cysteinyl-DOPA cross-links (Zhao *et al.*, 2005). Alternatively, they could also improve adhesion by participating in a thiol-mediated redox modulation mechanism similar to that described recently in the mussel adhesive (Yu *et al.*, 2011). In this mechanism, cysteine residues prevent the oxidation of DOPA, a catecholic amino acid capable of forming strong interactions with a variety of surfaces, into DOPA-quinone, its oxidized form that shows diminished adhesion (Lee *et al.*, 2006; Anderson *et al.*, 2010; Yu *et al.*, 2011). Sa-2 and Pc-2, both small and basic proteins, display numerous histidine, lysine, glycine, alanine, and tyrosine residues. Both proteins contain repeated motifs, and the consensus undecapeptide repeat of Sa-2 (VHKAAGYGGYG) is interestingly reminiscent of the dodecapeptide repeat of Pc-2 (HPAVH-KALGGYG). Like Sa-1, Sa-2 possesses 4 cysteine residues at its N-terminus, whereas Pc-2 has only 2 (Zhao *et al.*, 2005). Both Sa-3 and Pc-3 variants contain clusters of serine residues of similar size. The presence of a serine-free dodecapeptide in the middle of the sequence is, however, specific to Sa-3 and absent from Pc-3. Moreover, Sa-3A lacks the basic C-terminus displayed by Pc-3A and is therefore not polyampholitic.

Although the American and European species share related adhesive proteins with similar characteristics, align-

ment of the amino acid sequences of the proteins from both species resulted in limited percentages of identity. Indeed, a similarity of 61% was calculated between Sa-1 and Pc-1, while values of 41% were obtained for Sa-2/Pc-2, 51% for Sa-3A/Pc-3A, and 75% for Sa-3B/Pc-3B (to avoid bias due to the length of the proteins, the number of identical amino acids in the alignments was divided by the number of residues of the smaller sequence). Comparison of Sa-1 and -2 with other members of GY-rich and H-repeat protein families of *P. californica* did not result in higher percentages of identity. Therefore, although alignment with possibly yet unidentified adhesive proteins of *S. alveolata* could give higher scores, our results suggest that Sa-1, -2, and -3 are homologous to Pc-1, -2, and -3, respectively. However, low similarities indicate that natural selection has allowed significant sequence divergence without the loss of adhesive efficiency. This was also observed in other marine organisms in which the overall sequences similarity between homologous adhesive proteins, especially those involved in surface interactions (*e.g.*, proteins fp-3 from mussels and cp-19k from barnacles), is not high (Zhao *et al.*, 2006; Kamino, 2008). For some marine adhesive proteins, therefore, including those of sabellariids, shared features such as amino acid composition, post-translational modifications, presence of repeated patterns, pI, and mass appear to be more important than the primary structure. Divergent evolution of the amino acid sequences of both species, and especially their repeated motifs, could be explained by differences in the materials used to build the tube (calcareous shells *vs.* siliceous sand grains). It is indeed interesting to note that, although both types of materials are used by the American and the European species, tubes of *S. alveolata* are predominantly made up of shell fragments while those of *P. californica* are rather constituted of sand grains (Hartman, 1944; Vovelle, 1965; Simmons *et al.*, 2005).

In *P. californica*, analysis of the amino acid composition of the cement revealed that adhesive proteins are subjected to post-translational modifications, including hydroxylation of tyrosine residues into DOPA and phosphorylation of the serine residues (Waite *et al.*, 1992; Stewart *et al.*, 2004). The presence of DOPA has been confirmed in Pc-1 and Pc-2 (Waite *et al.*, 1992) and suggested in Pc-3 variants on the basis of their relatively abundant tyrosine residues (Zhao *et al.*, 2005; Stewart *et al.*, 2011), whereas pSer residues would be restricted to Pc-3 (Zhao *et al.*, 2005). In *S. alveolata*, these two modified amino acids were highlighted through histochemical and immunohistochemical detection (Vovelle, 1965; Flammang *et al.*, 2009; present study). By homology, it is likely that DOPA residues would occur in all three putative adhesive proteins of *S. alveolata*, whereas pSer residues would be present only in Sa-3. These modified amino acids are of great importance for the cohesive and adhesive strength of the cement. DOPA groups take part in surface coupling either through hydrogen bonds

or by forming complexes with metal ions and metal oxides present in mineral surfaces (Zhao *et al.*, 2006; Zeng *et al.*, 2010; Lee *et al.*, 2011). Following oxidation, DOPA groups also contribute to cement curing by forming intermolecular cross-links (*e.g.*, cysteinyl-DOPA cross-links; Zhao *et al.*, 2005). Phosphoserines, on the other hand, are thought to interact with calcareous substrates such as the mollusc shell fragments constituting the tubes, and to be involved in noncovalent cohesive interactions, possibly through  $\text{Ca}^{++}$  or  $\text{Mg}^{++}$  bridging (Waite and Qin, 2001; Sun *et al.*, 2007; Flammang *et al.*, 2009; Stewart *et al.*, 2011).

Our histochemical and immunohistochemical methods have also showed that, unlike DOPA residues, which are uniformly distributed in the secretory granules of both types of cement cells, pSer residues appear to be present mostly in the inclusions of the heterogeneous granules. This corroborates energy dispersive x-ray spectroscopic (EDS) analyses in which phosphorus was co-localized with calcium and magnesium in these inclusions in both *S. alveolata* (Gruet *et al.*, 1987) and *P. californica* (Stewart *et al.*, 2011). However, the weak anti-pSer immunoreactivity in homogeneous granules suggests a more complex distribution in *S. alveolata*. A possibility is that the two Sa-3 variants would each be specific for one type of cement cell, or that another yet unidentified polyphosphoprotein would be present in the homogeneous granules. For Sa-1 and Sa-2, as they both presumably comprise DOPA residues and as both types of secretory granules stain for DOPA, it is not possible to ascribe them to a specific type of cement cells. In the American species, it was demonstrated by immunohistochemistry and *in situ* hybridization that Pc-1, -3, and -4 are located exclusively in the heterogeneous granules, while Pc-2 and -5 are present only in the homogeneous granules (Wang and Stewart, 2012). Similar experiments should be conducted in *S. alveolata* to determine whether the distribution of Sa-1, -2, and -3 in the two types of cement glands matches that of their homologs in *P. californica*. This would also demonstrate unambiguously that Sa-1, -2, and -3 are actually cement proteins.

In *P. californica*, the co-occurrence of the positively charged and the negatively charged proteins in the cement led to the hypothesis that a phenomenon called complex coacervation could play a role in the condensation of the adhesive in the form of a dense water-immiscible fluid (Stewart *et al.*, 2004; Zhao *et al.*, 2005). Coacervation occurs when the charges of the polyelectrolytes are balanced, which seems to be the case in the cement of *S. alveolata*. Indeed, quantitative real-time PCR analyses revealed a mean expression level of positively charged proteins (Sa-1 and -2) 6.2 times greater than the one of their negatively charged counterparts (Sa-3A and -3B). Considering a phosphorylation rate of the serine residues of about 95% as in *P. californica* (Stewart *et al.*, 2004) and knowing that each phosphate group bears two negative charges at the

pH of the seawater, mature Sa-3A and -3B would respectively display 292 and 272 negative charges. On the other hand, positive charges in mature Sa-1 and -2 derive from lysine, arginine, and histidine residues and their total numbers amount to only 31 and 40, respectively. However, multiplying these values by 6.2 gives 192 and 248 positive charges, thus approaching the range of Sa-3 variants. Divalent cations (mainly  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ , Gruet *et al.* 1987) would then balance the remaining negative charges. Recently, however, Stewart *et al.* (2011) and Wang and Stewart (2012) pointed out that the rapidity of cement setting after release (*i.e.*, about 30 s; Stevens *et al.* 2007) and the poor mixing of the oppositely charged proteins within the cement are not consistent with the process of complex coacervation at the whole cement level. Nevertheless, at the granule level, electrostatic charge neutralization between oppositely charged polyions could drive the condensation of adhesive proteins (Wang and Stewart, 2012). Indeed, in *P. californica*, the negative charges of the polyanionic Pc-3 variants could be neutralized by the polybasic Pc-1 and Pc-4, by the divalent  $\text{Mg}^{2+}$ , by the basic C-terminus of Pc-3A, or by a combination of these mechanisms. In *S. alveolata*, although the fine localization of the adhesive proteins is unknown, it is reasonable to hypothesize that the relative over-expression of basic proteins would participate in the charge balance of the polyelectrolytes. Interestingly, the expression pattern of the different adhesive proteins is different in the two species. In *P. californica*, Pc-1 displays the highest level of expression (more than three times that of the other proteins; Wang and Stewart, 2012), whereas in *S. alveolata*, Sa-2 is the most expressed protein. If electrostatic interactions are at play, these different expression patterns could indicate different distributions of the proteins between the two types of cement cells. Alternatively, the relative expression levels of the adhesive proteins in *S. alveolata* could also reflect other characteristics, such as their respective function in the cement (interface *vs.* bulk).

### Acknowledgments

Authors thank Prof. P. Damman, Prof. J. Martial, Dr. C. Archambeau, and Dr. C. Van de Weerd for their help and advice. This work was supported by the "Service Public de Wallonie—Programme Winnomat 2", by the "Communauté française de Belgique—Actions de Recherche Concertées," and by COST Action TD0906 (<http://www.cost-bioadhesives.org/>). PF is Research Director of the Fund for Scientific Research of Belgium (FNRS). This study is a contribution of the Centre Interuniversitaire de Biologie Marine (CIBIM).

### Literature Cited

Anderson, T. H., J. Yu, A. Estrada, M. U. Hammer, J. H. Waite, and J. N. Israelachvili. 2010. The contribution of DOPA to substrate-



- peptide adhesion and internal cohesion of mussel-inspired synthetic peptide films. *Adv. Funct. Mater.* **20**: 4196–4205.
- Arnou, L. E. 1937.** Colorimetric determination of components of 3,4-dihydroxyphenyl-L-alanine/tyrosine mixtures. *J. Biol. Chem.* **118**: 531–537.
- Dales, R. P. 1952.** The development and structure of the anterior region of the body in the Sabellariidae, with special reference to *Phragmatopoma californica*. *Q. J. Microsc. Sci.* **93**: 435–452.
- Endrizzi, B. J., and R. J. Stewart. 2009.** Glueomics: an expression survey of the adhesive gland of the sandcastle worm. *J. Adhes.* **85**: 546–559.
- Flammang, P., A. Lambert, P. Bailly, and E. Hennebert. 2009.** Polyphosphoprotein-containing marine adhesives. *J. Adhes.* **85**: 447–464.
- Fournier, J., S. Etienne, and J.-B. Le Cam. 2010.** Inter- and intraspecific variability in the chemical composition of the mineral phase of cements from several tube-building polychaetes. *Geobios* **43**: 191–200.
- Gabe, M. 1968.** *Techniques Histologiques*. Masson, Paris.
- Gruet, Y., J. Vovelle, and M. Grasset. 1987.** Composante biominérale du ciment du tube chez *Sabellaria alveolata* (L.), annélide polychète. *Can. J. Zool.* **65**: 837–842.
- Hartman, O. 1944.** Polychaetous annelids, Part VI. Paraonidae, Magelonidae, Ctenodrilidae and Sabellariidae. Pp. 311–389 in *Allan Hancock Pacific Expeditions*, Vol. 10. University of Southern California Press, Los Angeles.
- Jensen, R. A., and D. E. Morse. 1988.** The bioadhesive of *Phragmatopoma californica* tubes: a silk-like cement containing L-DOPA. *J. Comp. Physiol. B* **158**: 317–324.
- Kamino, K. 2008.** Underwater adhesive of marine organisms as the vital link between biological science and material science. *Mar. Biotechnol.* **10**: 111–121.
- Lee, B. P., P. B. Messersmith, J. N. Israelachvili, and J. H. Waite. 2011.** Mussel-inspired adhesives and coatings. *Annu. Rev. Mater. Res.* **41**: 99–132.
- Lee, H., N. F. Scherer, and P. B. Messersmith. 2006.** Single-molecule mechanics of mussel adhesion. *Proc. Natl. Acad. Sci. USA* **103**: 12999–13003.
- Noernberg, M. A., J. Fournier, S. Dubois, and J. Populus. 2010.** Using airborne laser altimetry to estimate *Sabellaria alveolata* (Polychaeta : Sabellariidae) reefs volume in tidal flat environments. *Estuar. Coast. Shelf Sci.* **90**: 93–102.
- Obenauer, J. C., L. C. Cantley, and M. B. Yaffe. 2003.** Scansite 2.0: proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res.* **31**: 3636–3641.
- Paz, M., R. Flückinger, A. Boak, H. M. Kagan, and P. M. Gallop. 1991.** Specific detection of quinoproteins by redox-cycling staining. *J. Biol. Chem.* **266**: 689–692.
- Réaumur, R.A. Ferchault de. 1711.** Des différentes manières dont plusieurs espèces d'animaux de mer s'attachent au sable, aux pierres et les uns aux autres. *Mém. Acad. R. Sci.* 108–134.
- Rouse, G., and F. Pleijel. 2001.** *Polychaetes*. Oxford University Press, Oxford, 354 pp.
- Simmons, S. A., R. K. Zimmer, and C. A. Zimmer. 2005.** Life in the lee: local distributions and orientations of honeycomb worms along the California coast. *J. Mar. Res.* **63**: 623–643.
- Stevens, M. J., R. E. Steren, H. Vlamidir, and R. J. Stewart. 2007.** Multiscale structure of the underwater adhesive of *Phragmatopoma californica*: a nanostructured latex with a steep microporosity gradient. *Langmuir* **20**: 5045–5049.
- Stewart, R. J., J. C. Weaver, D. E. Morse, and J. H. Waite. 2004.** The tube cement of *Phragmatopoma californica*: a solid foam. *J. Exp. Biol.* **207**: 4727–4734.
- Stewart, R. J., C. S. Wang, and H. Shao. 2011.** Complex coacervates as a foundation for synthetic underwater adhesives. *Adv. Colloid Interface Sci.* **167**: 85–93.
- Sun, C., G. E. Fantner, J. Adams, P. K. Hansma, and J. H. Waite. 2007.** The role of calcium and magnesium in the concrete tubes of the sandcastle worm. *J. Exp. Biol.* **210**: 1481–1488.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994.** Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680.
- Vovelle, J. 1965.** Le tube de *Sabellaria alveolata* (L.): annélide polychète Hermellidae et son ciment. Étude écologique, expérimentale, histologique et histochimique. *Arch. Zool. Exp. Gen.* **106**: 1–187.
- Waite, J. H., and X. Qin. 2001.** Polyphosphoprotein from the adhesive pads of *Mytilus edulis*. *Biochemistry* **40**: 2887–2893.
- Waite, J. H., R. A. Jensen, and D. E. Morse. 1992.** Cement precursor proteins of the reef-building polychaete *Phragmatopoma californica* (Fewkes). *Biochemistry* **31**: 5733–5738.
- Wang, C. S., K. K. Svendsen, and R. J. Stewart. 2010.** Morphology of the adhesive system in the sandcastle worm, *Phragmatopoma californica*. Pp. 169–179 in *Biological Adhesive Systems: From Nature to Technical and Medical Application*, J. von Byern and I. Grunwald, eds. Springer, Vienna.
- Wang, S. W., and R. J. Stewart. 2012.** Localization of the bioadhesive precursors of the sandcastle worm, *Phragmatopoma californica* (Fewkes). *J. Exp. Biol.* **215**: 351–361.
- Winsnes, I. M. 1985.** The use of Methyl Green as an aid in species discrimination in Onuphidae (Annelida, Polychaeta). *Zool. Scr.* **14**: 19–23.
- Yu, J., W. Wei, E. Danner, R. K. Ashley, J. N. Israelachvili, and J. H. Waite. 2011.** Mussel protein adhesion depends on interprotein thiol-mediated redox modulation. *Nat. Chem. Biol.* **7**: 588–590.
- Zeng, H., D. S. Hwang, J. N. Israelachvili, and J. H. Waite. 2010.** Strong reversible Fe<sup>3+</sup>-mediated bridging between dopa-containing protein films in water. *Proc. Natl. Acad. Sci. USA* **107**: 12850–12853.
- Zhao, H., C. Sun, R. J. Stewart, and J. H. Waite. 2005.** Cement proteins of the tube-building polychaete *Phragmatopoma californica*. *J. Biol. Chem.* **280**: 42938–42944.
- Zhao, H., N. B. Robertson, S. A. Jewhurst, and J. H. Waite. 2006.** Probing the adhesive footprints of *Mytilus californianus* byssus. *J. Biol. Chem.* **281**: 11090–11096.