Integrative Biology, 2019, 11(5), 235–247

doi: 10.1093/intbio/zyz020
Advance Access publication June 28, 2019
Original Article

ORIGINAL ARTICLE

Pressure cycling technology for challenging proteomic sample processing: application to barnacle adhesive

Janna N. Schultzhaus[1,2,*], Scott N. Dean[1,2], Dagmar H. Leary[2], W. Judson Hervey[2], Kenan P. Fears[3], Kathryn J. Wahl[3], and Christopher M. Spillmann[2]

[1] National Research Council Research Associateship Programs Fellow, Washington, D.C., USA, [2] Center for Bio/Molecular Science & Engineering, Naval Research Laboratory, Washington, D.C., USA, and [3]Chemistry Division, Naval Research Laboratory, Washington, D.C., USA

*Corresponding author. E-mail: Janna.Schultzhaus.ctr@nrl.navy.mil

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019

Abstract

Successful proteomic characterization of biological material depends on the development of robust sample processing methods. The acorn barnacle *Amphibalanus amphitrite* is a biofouling model for adhesive processes, but the identification of causative proteins involved has been hindered by their insoluble nature. Although effective, existing sample processing methods are labor and time intensive, slowing progress in this field. Here, a more efficient sample processing method is described which exploits pressure cycling technology (PCT) in combination with protein solvents. PCT aids in protein extraction and digestion for proteomics analysis. Barnacle adhesive proteins can be extracted and digested in the same tube using PCT, minimizing sample loss, increasing throughput to 16 concurrently processed samples, and decreasing sample processing time to under 8 hours. PCT methods produced similar proteomes in comparison to previous methods. Two solvents which were ineffective at extracting proteins from the adhesive at ambient pressure (urea and methanol) produced more protein identifications under pressure than highly polar hexafluoroisopropanol, leading to the identification and description of >40 novel proteins at the interface. Some of these have homology to proteins with elastomeric properties or domains involved with protein-protein interactions, while many have no sequence similarity to proteins in publicly available databases, highlighting the unique adherent processes evolved by barnacles. The methods described here can not only be used to further characterize barnacle adhesive to combat fouling, but may also be applied to other recalcitrant biological samples, including aggregative or fibrillar protein matrices produced during disease, where a lack of efficient sample processing methods has impeded advancement. Data are available via ProteomeXchange with identifier PXD012730.

Received March 8, 2019; **revised** May 8, 2019; **editorial decision** May 22, 2019; **accepted** May 29, 2019

© The Author(s) 2019. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Insight

Optimization of proteolysis and minimization of sample loss are important considerations during proteomics sample preparation, yet the recalcitrant nature of some insoluble biological samples (i.e. aggregative proteins) necessitates procedures that work against these goals. Here, we demonstrate how pressure cycling technology can be used with strong denaturants to characterize a notoriously intractable biological sample, barnacle adhesive, and aid in sample dissolution as well as increase proteolysis. The methods resulted in the identification of many previously seen adhesive proteins and allowed for the characterization of novel proteins, offering insight into the function and evolution of one of nature's most robust adhesives.

Background

The production of reliable bottom-up proteomics data depends on efficient proteolysis, a sample processing step that cannot occur without protein release and solubilization. This requires sufficient homogenization and lysis when working with tissue or cell culture samples, as intact cells or cellular compartments will block protein and protease interactions. Denaturants also improve proteolysis by revealing proteolytic sites that are inaccessible in folded proteins. Lysis and denaturation protocols have been developed and standardized for many biological samples, yet some samples require additional methods. These samples include difficult to lyse cells (i.e., fungi with rigid cell walls) and cell fractions containing insoluble proteins, such as transmembrane proteins, which are highly hydrophobic and resist proteolysis during standard procedures. Modified protocols employing mechanical or chemical disruption and strong chaotropic or solvating agents for protein denaturation have been developed to address these issues [1, 2]. Protein aggregates are a different type of biological sample that also present sample processing challenges due to protein insolubility, but have not received as much focus for proteomics method development. Protein aggregates can develop during disease states or aging, when misfolding results in the formation of amyloid fibrils [3, 4], or during normal growth to allow for biological adhesion, including in the extracellular matrix of bacterial biofilms [5] or in the formation of the proteinaceous adhesive produced by barnacles [6–9].

Barnacles are a diverse infraclass of marine crustaceans that employ a variety of strategies to survive a sessile existence, from parasitic Rhizocephala to stalked (Pedunculata) and acorn (Sessilia) barnacles. Stalked and acorn barnacles adhere tenaciously to a wide variety of substrates via a secreted proteinaceous adhesive, relying on the permanency of the bond for survival. This process is remarkable from both a biofouling and novel materials perspective. Biomolecules released to the surface during growth cure to form an adhesive, contributing to the status of barnacles as a costly pest species for commercial and military fleets [10]. Further characterization of the components of this adhesive could lead to the development of coatings that inhibit attachment without the use of biocides or environmentally harmful chemicals [11], as well as next-generation adhesives [12]. Many marine organisms rely on a proteinaceous adhesive in direct contact with the interface for attachment, yet, as arthropods that are distantly related to other sessile marine organisms, barnacles likely possess a novel method for adherence [13]. During barnacle growth and expansion, proteins are deposited at the interface [14] where a poorly understood curing process produces the insoluble cement critical for permanent adhesion. While some interface proteins are water soluble (pheromones [15, 16], peroxidases, and protease inhibitors [17]), the matrix of proteins that provide structural support and adhesion are not. Until recently, only a small number of proteins in the adhesive layer of acorn barnacles had been identified [18–20] largely due to the insolubility of the cement to commonly used denaturants such as guanidine hydrochloride, urea, and formic acid, as well as reliance on early protein sequencing techniques like Edman degradation. Early reports on partial adhesive solubilization used a combination of formic acid and a variety of chemicals to break disulfide bonds, resulting in the identification of several proteins named according to species and average molecular weight [19, 20]. Homologous proteins in other species were later identified (*Amphibalanus amphitrite* [21], *Tetraclita japonica formosana* [22], *Pollicipes pollicipes* [23], *Lepas antifera* [24]) though the number of novel proteins found in the barnacle cement layer remained unchanged for several decades.

Recently, around 50 proteins were identified using a combination of a transcriptome of the sub-mantle tissue of *A. amphitrite* [21] and treatment of the adhesive layer collected from multiple surfaces with the polar solvent hexafluoroisopropanol (HFIP) [6, 25]. Several of the proteins possess conserved low-complexity domains with homology to spider silks. Other proteins, including enzymes with oxidase activity [17], were also identified, supporting earlier observations of the presence of enzymes at the interface [25–28]. While these methods have been effective in uncovering a host of previously unidentified proteins at the adhesive interface [17], the gel-based methods used to prepare samples for subsequent mass spectrometry (MS) analysis were time consuming and required relatively large amounts of starting material for use with proteomics approaches. There is a need to develop alternative sample processing techniques to maximize the amount of solubilized material for downstream analysis, which would facilitate experiments examining how barnacle cement proteins enable adhesion.

The current work examines whether pressure cycling technology (PCT) in combination with a variety of solvents can overcome the difficulties associated with barnacle adhesive to prepare samples in less time and with smaller amounts of starting material per sample for proteomics analysis. PCT involves pressure application to enhance protein extraction [29] and enzymatic digestion [30]. Cyclic bursts of high pressure (up to 45 kpsi) disrupt molecular interactions, leading to homogenization of solid samples and increased solubility of proteins and improved protease accessibility. Trypsin kinetics are also increased under pressure [30], reducing sample processing time. PCT could therefore facilitate maximal protein extraction from barnacle cement for subsequent MS analysis. As many of the barnacle adhesive proteins are water-insoluble, sample processing concerns include precipitation and aggregation during digestion. These are similar problems encountered during proteomic analysis of transmembrane proteins or misfolded amyloid fibrils formed

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019

during disease [3]. Various solvents have been used in these fields to efficiently solubilize and process these difficult samples, including urea, methanol, and HFIP. Using relatively small amounts (milligrams) of starting material, barnacle adhesive was subjected to PCT for homogenization and promotion of protein digestion under different solvent conditions.

Materials and Methods

Barnacle husbandry and adhesive collection

Young adult *Amphibalanus* (= *Balanus* [31, 32]) *amphitrite* attached to silicone coated panels were obtained from Duke University Marine Laboratory where settlement and husbandry were performed as previously described [33]. At the Naval Research Laboratory (Washington, DC), *A. amphitrite* were maintained as described previously [6]; briefly, barnacles were kept in an incubator at 23°C with 12 hour day/night cycles in 32 ppt artificial seawater (Instant Ocean, Blacksburg, VA) for three months and fed *Artemia* spp. *nauplii* (Brine Shrimp Direct, Ogden, UT) three times a week.

Individuals that had formed thickened, white, opaque adhesive [33–36], were gently dislodged from silicone coated panels the day of glue collection. The adhesive was peeled off the baseplate of the barnacles using a metal probe and divided into 1–2 mg (wet weight) amounts in 150 µl Barocycler tubes (Fig. 1**B, step 1**) and stored at 4°C overnight. Every 1 mg of wet weight adhesive is estimated to contain 0.875 mg water and 0.125 mg of protein [36].

Electron microscopy

In preparation for electron microscopy (EM), an adult barnacle attached to a gold-coated (100 nm) microscope slide was fixed with 3% glutaraldehyde in artificial seawater for 30 min, triple rinsed with deionized water, and then the main body and side plates were removed. The remaining basal region of the barnacle was dehydrated by soaking the slide for 30 min in each of the following ethanol solutions in order, 10%, 20%, 30%, 40%, 50%, 70%, 90%, and 100%, dried in a critical point dryer, and sputter coated with 5.0 nm of gold to minimize charging. The sample was imaged on a Leo field emission secondary electron microscope (Carl Zeiss Microscopy) at a voltage of 3.0 kV and a working distance of 2.1 mm using the InLens detector.

Protein extraction from barnacle adhesive and digestion

Proteins were extracted from adhesive with Pressure Cycling Technology (PCT) by incubation with 30 µl of different solvents in a Barocycler (Pressure BioSciences, South Easton, MA) with 30 µl micropestle caps for 60 cycles of 45 kpsi for 50 s and ambient pressure for 10 s at 30°C (Fig. 1**B, steps 2–3**). **Table 1** contains descriptions of tested solvents (HFIP (100% or 30%), 8 M urea, 60% methanol in 100 mM ammonium bicarbonate, and combinations thereof). Each condition was performed n = 6.

Cysteines were then reduced and modified by adding 10 mM tris(2-carboxyethyl)phosphine (TCEP) and 40 mM iodoacetamide (IAA) and incubating for 30 minutes in the dark (Fig. 1**B, step 4**) [37]. Sequential digestion by LysC and trypsin were performed by first adding 5 µl *n*-propanol and 4 µg LysC/Trypsin mix (Promega Corporation, Madison, WI) to a final volume of 55 µl, using the 50 µl caps. At this concentration of urea (~4.5 M), trypsin is inhibited and only LysC is active. LysC digestion occurred in the Barocycler for 45 cycles of 20 kpsi for 50 s and ambient pressure for 10 s at 37°C (Fig. 1**B, step 5**). Next, 10 µl of *n*-propanol and 85 µl of 100 mM ammonium bicarbonate was added to bring the final volume to 150 µl, and the 150 µl caps were used. Trypsin digestion occurred in the Barocycler for 90 cycles of 20 kpsi for 50 s and ambient pressure for 10 s at 37°C (Fig. 1**B, step 6**).

After digestion, two samples of each condition were pooled together, resulting in a total of three samples. These samples were desalted with Strata-X Polymeric SPE 30 mg/3 ml columns (Phenomenex, Torrence, California) and eluted with 250 µl of 70:30 acetonitrile:1.0% formic acid in water. The eluted samples were dried via speed vacuum and analyzed via tandem mass spectrometry analysis.

Tandem mass spectrometry and bioinformatics

Mass spectrometry was completed as described in [38]. Briefly, digested proteins were resuspended in 100 µl of 0.1% FA in water and analyzed by liquid chromatography mass spectrometry/mass spectrometry (LC-MS/MS) with a Tempo-MDLC coupled to a TripleTOF 5600+ mass spectrometer (AB Sciex, Foster City, CA). 7 µl of the samples were loaded onto and eluted from dual 3 µm 120 Å ChromXP C_{18}CL RP Columns with a gradient from 80:5 to 5:80 0.1% FA in H_2O:acetonitrile over 140 minutes. Tandem mass spectra were extracted by Sciex script and searched against the BarnALL database using Mascot as described in So *et al.* [6]. The BarnALL database was generated from translated cDNA sequences produced from RNA-seq experiments of the sub-mantle tissue [21] and also contains the 52 proteins identified in the shell [39]. The database also contains a concatenated list of common contaminants. Samples were analyzed assuming trypsin digestion with a precursor and MS/MS tolerance of ±0.6 Da. Deamidation and oxidation were listed as variable modifications in the Mascot search parameters. Scaffold was then used to assess and verify peptide and protein

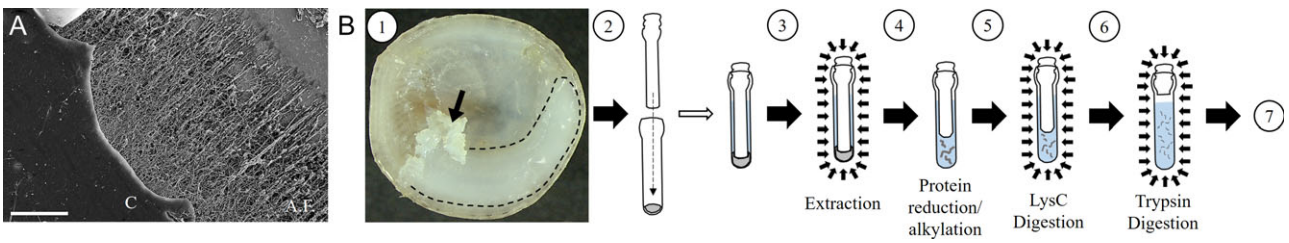

Figure 1. PCT methods for barnacle cement proteomics. (A) Electron microscopy image of the cuticle (C) and underlying adhesive fibrils (A.F.) that anchor *A. amphitrite* to the substrate. Scale bar = 10 µm. (B:1) Collected adhesive (arrow; undisturbed thickened adhesive is outlined) is placed [2] in a MicroTube with solvent and the MicroPestle cap. [3] Proteins are then extracted using pressure cycling technology. [4] After cysteine modification, LysC/Trypsin mixture and the 50 µl cap are added. [5, 6] Sequential LysC/trypsin digestion occurs under pressure. Trypsin digestion happens after further sample dilution in 150 µl volume. [7] After cleanup, samples can be submitted for analysis via mass spectrometry.

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019

Table 1. Solvent concentrations during protein extraction and protease digestion.

Condition	HFIP	Urea	Methanol
Full strength HFIP	100% (20%)	–	–
Dilute HFIP	30% (6%)	–	–
Dilute HFIP/Urea	30% (6%)	8 M (1.4 M)	–
Urea	–	8 M (1.4 M)	–
Dilute HFIP/MeOH	30% (6%)	–	60% (12%)
MeOH	–	–	60% (12%)

Numbers listed first describe concentrations during protein extraction and numbers in parentheses describe concentrations during digestion with trypsin.

assignments (peptide threshold > 95%, minimum number of peptides ≥ 3, protein threshold > 99.0%; estimated false discovery (FDR) for peptides = 0.0% and for proteins = 0.0% at these settings).

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [40] partner repository with the dataset identifier PXD012730.

Protein sequence analysis

Peptides and their amino acid composition were examined using R [41] and the package Peptides [42]. Heat maps were generated using R and the function heatmap.2 with the package gplots [43]. ExPASy ProtParam tool was used to analyze the amino acid composition and the aliphatic and grand average of hydropathy (GRAVY) indices of protein sequences [44].

Protein sequences and conserved domains were analyzed using NCBI's BLASTP with default parameters [45] and the Conserved Domain Database [46]. Homology to annotated sequenced and conserved domains was deemed significant at E values ≥ 1e-4.

Results

Method development

Amphibalanus amphitrite produces a fibrillar network of proteinaceous adhesive (Fig. 1**A**) that is highly resistant to solubilzation. PCT sample preparation methods developed by Guo *et al.* [37] were modified to process *A. amphitrite* adhesive (Fig. 1**B**) and assess protein identification under different solvent conditions. This method allows for protein extraction and digestion to occur in the same tube, minimizing sample loss and increasing throughput. Six conditions using three solvents (HFIP [full strength or dilute], urea, and methanol and combinations thereof) were tested (**Table 1**). This methodology increases throughput, both in terms of sample number (16 samples can be digested simultaneously), and time as PCT also facilitates protease digestion, decreasing total digestion time to 135 minutes with LysC/trypsin compared to overnight trypsin digestion at ambient pressure.

Solvent effects on peptide and protein characteristics

Since the *A. amphitrite* genome has not yet been fully sequenced, protein identification is limited to a database created from transcripts of submantle tissue and other sequences in the public domain [21]. This database may be lacking sequences of proteins which are present in the adhesive as it is only a subset of protein complements expressed by *A. amphitrite*. For this reason, the number of acquired spectra for each sample as well as the number of protein identifications were examined to make comparisons among the extraction solvents.

The various extraction solvents resulted in different spectral and proteomic profiles of *A. amphitrite* adhesive (**Fig. 2 & Files S1-2**). Overall, the total number of spectra in each sample was relatively consistent except for the sample prepared with urea, which had significantly more (30 200 ± 1100 [mean ± SE]) (Fig. 2**A**). The fewest spectra were produced by the dilute HFIP extraction treatment (4800 ± 555). All other conditions (full strength HFIP, dilute HFIP + urea, dilute HFIP + methanol, and methanol) yielded approximately 10 000 spectra per sample. The variation for the dilute HFIP + urea treatment was higher than all others due to one of the three replicates being an outlier with a higher number of spectra acquired. This replicate aligned more closely with the results seen in the pure urea condition (i.e., high number of spectra), whereas the other two aligned more closely with the dilute HFIP condition (i.e., low number of spectra). For all other parameters examined, this particular replicate is consistently higher than the other two, causing large variation for the dilute HFIP condition.

Of the spectra acquired, the percentage assigned to peptide sequences was affected by the extraction conditions (Fig. 2**B**). The largest percentage of identified spectra were found in the pure urea (11.2% ± 1.1), methanol (10.1% ± 2.0), and full strength HFIP (9.0% ± 0.8) samples. The spectrum assignment in dilute HFIP + methanol samples was slightly lower at 7.2% ± 0.2. The dilute HFIP + urea condition had 4.8% ± 4.6 spectra identified as peptides, and dilute HFIP alone had the lowest assignment with only 0.6% ± 0.3 spectra matched to peptide sequences. Overall, the number of assigned spectra is low, and many spectra remain unassigned, an issue likely due to the limitations of the current database.

The number of missed cleavages in identified peptides was quantified for each extraction condition to estimate solvent influence on protease activity (Fig. 2**C**). The full strength HFIP condition and the HFIP/urea mixture were markedly different from all other conditions with a greater percentage of peptides (approximately 65%) containing at least one missed cleavage. The other five conditions had relatively similar missed cleavage profiles with 25–35% peptides displaying a remaining tryptic site. These results indicate that HFIP could either have a deleterious effect on trypsin efficiency or extract modified proteins that resist proteolysis more effectively than other conditions.

The total number of proteins identified by 3 or more peptides ranged from 74 in the pure urea condition to 5 in the dilute HFIP treatment (Fig. 2**D**). Most proteins were consistently identified in the pure urea and methanol extractions, therefore throughout the remainder of this manuscript comparisons will be made only between these two conditions and to the full strength HFIP condition which has been used previously to solubilize *A. amphitrite* adhesive [6] (**Table S1**).

Next, the primary sequences of identified peptides and their polarity among the extraction conditions were characterized (**Fig. 3**). The amino acid properties of peptides identified in HFIP samples differed significantly from both the peptides in the urea and methanol conditions for six (tiny, small, aromatic, non-polar, polar, and basic amino acids [Kruskal-Wallis test, χ^2>7, df = 2, p > 0.02 for all]) of the nine types examined (aliphatic, charged, and acidic amino acids were not significantly different) (Fig. 3**A** & Tables S2-3 for results of Kruskal-Wallis and Dunn post-hoc test; see also **Fig. 3** legend for definition of amino acid types). The HFIP condition extracted more peptides with a higher portion of tiny, small, and polar amino acids. There were

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019

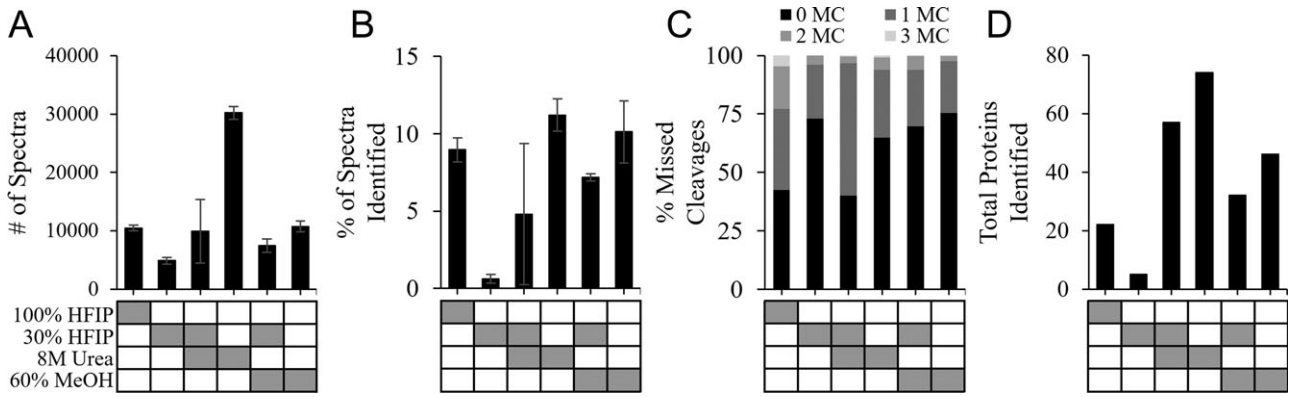

Figure 2. Comparison of solvent effect on mass spectrometry results. (A) The total spectra, (B) percentage of spectra identified at the peptide level, (C) the percentage of missed cleavages, and (D) the total number of identified proteins for each solvent tested. Solvent composition is indicated by the gray boxes beneath the graphs. N = 3, error bars = mean ± SE.

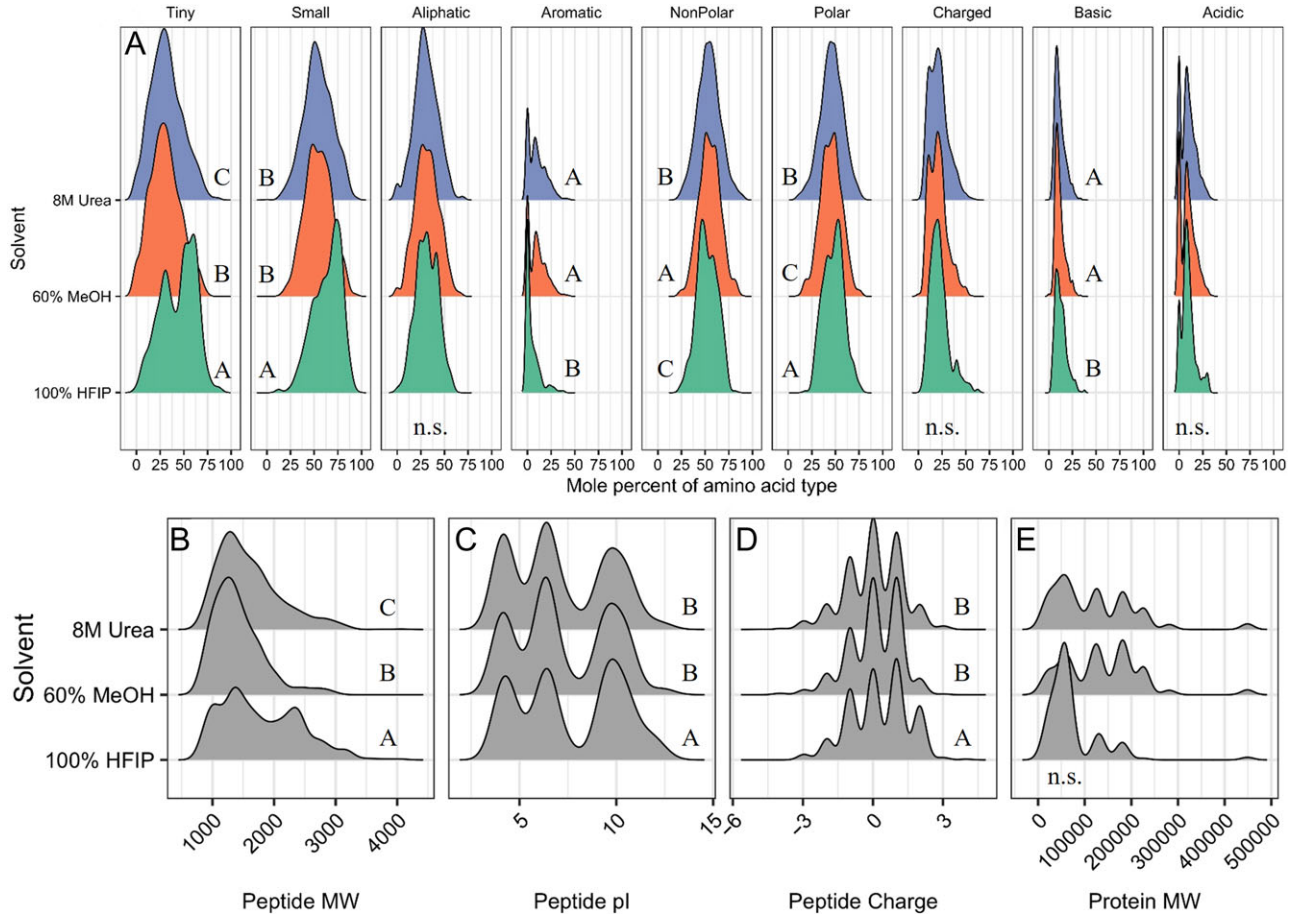

Figure 3. Peptide and protein properties. (A) The amino acid composition, (B) molecular weight, (C) isoelectric point, and (D) charge of the peptides and (E) molecular weight of the proteins identified in the urea, methanol, and HFIP treatments are shown. Amino acid type: Tiny (A, C, G, S, and T); Small (A, B, C, D, G, N, P, S, T, and V); Aliphatic (A, I, L, and V); Aromatic (F, H, W, and Y); Non-polar (A, C, F, G, I, L, M, P, V, W, and Y); Polar (D, E, H, K, N, Q, R, S, T, and Z); Charged (B, D, E, H, K, R, and Z); Basic (H, K, and R); Acidic (B, D, E, and Z); B = asparagine/aspartic acid; Z = glutamine/glutamic acid. Y-axis units are density. Letters next to plots indicate significant differences at $\alpha = 0.05$. n.s. = not significant.

also more peptides with a lower portion of aromatic and basic amino acids compared to the urea and methanol conditions (**Table S3**; $p < 0.05$). The only significant differences in amino acid composition between the urea and methanol extracted peptides were seen in the categories of tiny ($p = 0.0400$) and nonpolar/polar ($p = 0.0346$) amino acids, where peptides consisting of more tiny and fewer polar amino acids were found in urea samples.

Each solvent condition led to peptides of different MW being identified (Kruskal-Wallis test, $\chi^2 = 139.641$, df = 2, $p = 0$). HFIP

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019

extracted longer peptides, (methanol: p = 5.49e-35; urea: p = 9.90E-12), and methanol extraction conditions resulted in shorter peptides than the urea condition (p = 1.11e-20, Fig. 3**B**). The isoelectric point also showed significant differences (Kruskal-Wallis test, χ^2= 13.067, df = 2, p = 0), with fewer HFIP peptides with a pI around 6 and more with a pI higher than 10 (methanol: p = 9.44e-3; urea: p = 9.02e-4; Fig. 3**C**). Solvents also affected peptide charge (Kruskal-Wallis test, χ^2= 21.748, df = 2, p = 0), with HFIP producing fewer neutral peptides and more +2 charged peptides (methanol: p = 2.59e-4; urea: p = 1.01e-5; Fig. 3**D**). The Boman Index, which computes the probability of peptide interaction based on residue solubility [47], was also examined, though no significant differences were identified. Finally, the effect of solvents on the molecular weight of the identified proteins (as opposed to peptide MW in Fig. 3**B**) was examined. No significant effect on protein molecular weight among the extraction conditions was observed, although full strength HFIP plots show MW to be shifting toward smaller values (Fig. 3**E**).

Comparison of the number of proteins identified per condition

The overlap of the proteins identified from each condition was examined next (**Fig. 4**). When looking at the overlap between all conditions (Fig. 4**A**), it is apparent that each condition produced a subset of both unique and overlapping proteins, though the variation was large. Minus one protein, those identified in the dilute HFIP treatment are contained completely within the full strength HFIP treatment, and the conditions where HFIP was mixed with urea and methanol cluster closer to the full strength HFIP treatment than the pure urea or methanol treatments. Mixtures of HFIP and methanol or urea also cluster more closely with their pure counterpart, *i.e.*, dilute HFIP + urea is close to the pure urea condition while dilute HFIP + methanol is closer to pure methanol.

Comparing the number of proteins identified in HFIP, urea, and methanol, 17 proteins were common to all three conditions (Fig. 4**B**). Samples processed in urea had 37 uniquely identified proteins with very few unique proteins identified in either the HFIP (2) or methanol conditions (3). In terms of overlap, 31 proteins were found in both the urea and the methanol conditions, only 3 proteins found in both the HFIP and urea conditions, and none were shared between the methanol and HFIP conditions.

Novel interface proteins identified via PCT

Figure 5 shows the total spectrum counts per protein identified in this work for the full strength HFIP and pure urea and methanol conditions. In addition to proteins found at the interface that have previously been described [6, 17], 39 other proteins or protein clusters were identified via PCT (bolded protein names in **Fig. 5**, **File S3**). Some of these proteins had been detected via proteomics analysis in previous work but not described (**Table S1**) [6, 17]. To further sort and predict the function of this set of proteins, their homology to known proteins was ascertained via BlastP (**File S4**). During this assessment, a naming convention was adopted to facilitate continued analysis of *A. amphitrite* proteins and was retroactively applied to previously identified and named proteins. Proteins are named AaXx$_{(n)}$-m, Aa for *Amphibalanus amphitrite*, and Xx$_{(n)}$ for proteins with a contracted name (for example, AaMuc for mucin), or AaXXx for proteins with multiple names (for example, AaLOx for lysyl oxidase), followed by m (1 through m) to denote individual proteins in a group. For proteins named after a molecular weight (whether or not this molecular weight applies to the protein in question was not addressed here, historical naming conventions were followed), the scheme AaCP##-n has been adopted, where CP stands for 'cement protein,' followed by ## = MW and –n to denote individual proteins within a group.

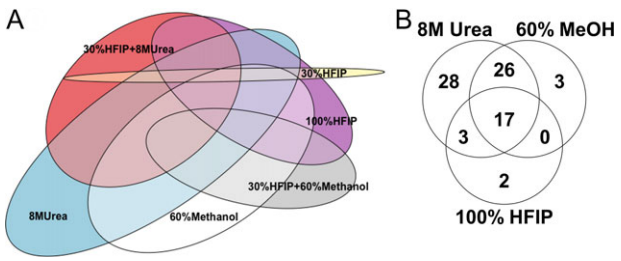

Figure 4. Overlap of proteins identified from each condition. (A) Total overlap between all tested conditions and (B) overlap between 8 M Urea, 60% MeOH, and 100% HFIP extraction conditions at 3 peptide protein identification threshold.

Information from homology and conserved domains was used to classify proteins as bulk proteins, enzymes, or pheromones (**File S5**). Several proteins either had conserved domains or homology to proteins with annotations that indicated either a structural function as a bulk protein, including hemocytin- (AaHem), glutenin- (AaGlut), titin- (AaTitin), and fasciclin domain containing- (AaFas) like proteins; an enzymatic function, including peroxidase- (AaPx), serine protease- (AaSP), and protease inhibitor- (AaPI) like proteins; or a role in communication as pheromones. Known barnacle pheromones include MULTIFUNCin (AaMulti), settlement inducing protein complex (AaSIPC), and waterborne settlement pheromone (AaWSP). All of these pheromones were identified in this study, and each has homology to pheromones found in other barnacle species. Two additional variations of AaMulti (-4 and -5), 3 variations of AaWSP (-1 through -3), and one version of AaSIPC were identified in the adhesive via PCT. AaMulti-3 was not identified in the adhesive here, but had been previously [6]. AaMulti-5 did not have significant homology to known MULTIFUNCin proteins in other barnacle species, but did contain the conserved domains found in the other AaMulti proteins, including A2M and ISOPREN_C2_like domains which function as proteinase inhibitors [16], and displayed homology to AaMulti-1 through -4.

A number of proteins had no significant homology to any publically available protein sequences. The sequence of these proteins were compared to the bulk structural proteins and to each other, and this analysis resulted in the identification of three proteins with homology (E value ≥ 1.0e-4) to AaCP105-1 (AaCP105-3, -4, and -5), two proteins with homology to AaMuc-1 (AaMuc-2 and -3), and a group of 5 proteins with homology to each other (AaCP34-1 through -5; **File S6**). Previously, interface proteins were characterized by enriched amino acid content [6]. Like AaCP105-1 and -2, AaCP105-3, -4, and -5 have high levels of polar amino acids (Gly/Ala/Ser/Thr; Glycine/Serine rich Cement Proteins [GSrCPs]). The AaCP34-like proteins all have high leucine content (Leucine rich Cement Proteins [LrCPs]) and a high aliphatic index (**Table 2**), more closely aligning with the AaCP52- and AaCP100-like proteins which are also LrCPs with a high aliphatic index.

Four proteins have homology that indicates they could possess multiple functions, one of which is in immunity. These proteins included two variations of vitellogenin (AaVit), β-1,3-glucan-binding protein precursor (AaβGBPP), and a hemocytin-like protein (AaHem). Evolution from ancestral immune to adhesive roles has been suggested [27], leading us to place these proteins within the list of bulk proteins.

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019

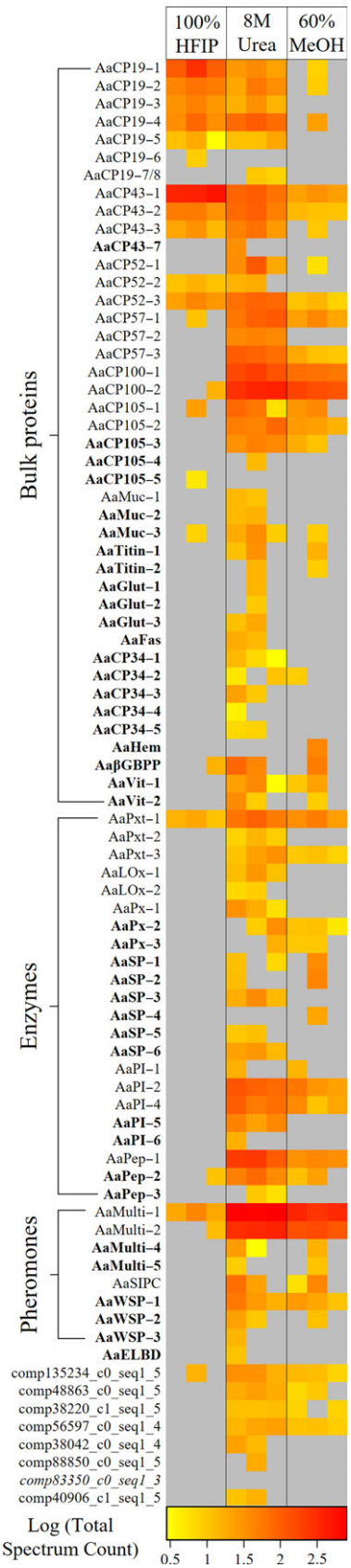

Figure 5. Heat map of the total spectrum counts of proteins identified in 100% HFIP, 8 M Urea, and 60% MeOH extraction conditions. Three replicates per condition are shown. Identified proteins are displayed in three predicted functional groups (structural proteins, enzymes, and pheromones). Bolded names indicate proteins identified or described for the first time in this study. One protein (comp83350_c0_seq1_3) was not identified in the three conditions shown here which is indicated by italicizing the name. Total spectrum counts for each replicate were log10 transformed. Gray cells indicate where proteins were not identified at the three peptide threshold.

The conserved domains found in these sequences were also examined (**File S5**). Few conserved domains were found in the bulk proteins (only 7 out of 39). Several proteins had low homology to domains with functions that would not be expected to play a role in adhesion (AaCP43-1 contained the Flagellin_C domain, E value = 1.3e-3; AaCP43-2 contained DNA_pol3_delta2, E value = 2.2e-5; AaCP57-3 contained the Nucleoporin_FG2 domain, E value = 6.0e-4; and AaGlut-2 contained the dnaA domain, E value = 1.2e-6). Three bulk proteins (AaCP105-1 and -3 and AaHem) contained chitin binding domains. AaHem contained a number of other conserved domains, including mucin, FA58C, TIL, and superoxide dismutase domains. The final structural protein containing conserved domains is AaFas, which has Fasciclin domains (E value = 9.9e-14). Many of the domains found in AaCP105-1, AaCP105-3, AaHem, and AaFas function in cell or protein adhesion.

All proteins identified as enzymes (except for AaPep-1, which has no conserved domains) have at least one conserved domain that indicates their function as either an oxidase, protease, protease inhibitor, or endopeptidase. AaLOx-1 and -2 both contain SRCR domains, and AaLOx-1 has a Lysyl_oxidase domain (E value = 3.3e-87). The three peroxinectin- and three peroxidase-like proteins all contained An_peroxidase type domains with high homology (E value < 1e-26). AaSP-1 through 6 all contained domains with predicted serine protease functions at moderate to high homology (E value < 1e-20). All AaPI proteins contained SERPIN domains (SERine Protease INhibitors, E value > 1e-29). AaPep-2 and -3 contain predicted cysteine-rich secretory protein domains (SCP and CAP, E value < 2e-8) and have homology with predicted venom allergens (AaSP-7 E value = 5e-42; AaSP-8 E value = 6e-45). SCP and CAP domain containing proteins may act as peptidases, among a number of other potential functions [48].

One final protein contained a predicted extracellular ligand binding domain (Neur_chan_LBD superfamily, E value > 9.9e-32; AaELBD), which has been shown to bind acetylcholine and other neurotransmitters [49]. How this domain could contribute to adhesion is unclear. Ultimately, a final 8 proteins had no homology to any known proteins in NCBI or to other *A. amphitrite* interface proteins, and 7 had no conserved domains. One protein (comp38220_c1_seq1_5) had homology to juvenile hormone binding domain (E value = 3.1e-6). These 8 undefined proteins are listed by their accession number at the bottom of **Fig. 5**.

Solvent effects on the interface proteome

Finally, the effect of solvents on protein identification in the three predicted functional groups was examined. The pure urea treatment led to the most protein identifications overall, with no bias in what type of protein was identified. All of the 8 pheromones, all but one of the enzymes (AaSP-4), and all except for three of the structural proteins (AaCP19-6, AaCP105-5, and AaHem) were seen. The methanol treatment led to fewer identifications, and fewer structural proteins (23/42) and enzymes (12/22) were identified, but almost all (7/8) pheromones were seen. In contrast, the full strength HFIP treatment led to the

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019

Table 2. Amino acid composition, aliphatic index, and GRAVY scores for AaCP34- and AaCP105-like proteins.

Protein Name	Accession	Enriched for:	AA Composition	Aliphatic Index	GRAVY
Leucine rich Cement Proteins [LrCPs]					
AaCP34-1	comp54786_c0_seq1_4	Leucine	16.5%	107.96	0.072
AaCP34-2	comp51936_c0_seq1_5	Arginine	11.7%	78.53	–0.4233
		Leucine	9.5%		
AaCP34-3	comp55944_c0_seq1_4	Leucine	11.0%	93.62	0.073
AaCP34-4	comp55167_c0_seq1_4	Leucine	11.0%	97.13	–0.052
AaCP34-5	comp80764_c0_seq1_4	Leucine	13.5%	92.85	–0.255
Glycine/Serine rich Cement Proteins [GSrCPs]					
AaCP105-3	comp42253_c0_seq1_4	Alanine	11.8%	71.8	–0.342
		Glycine	10.0%		
AaCP105-4	comp45569_c1_seq2_6	Serine	11.3%	68.45	–0.445
AaCP105-5	comp57780_c0_seq1_6	Asparagine	10.5%	69.92	–0.897
		Threonine	12.1%		

identification of few pheromones (2) and enzymes (2), but a modest number of structural proteins (17) were found.

To gain a better understanding of what proteins are most likely to be identified in *A. amphitrite* adhesive, core proteins, i.e. those found in each of the 3 replicates of each condition, were identified and compared between solvents (**Table S4**). While HFIP only led to the identification of 22 proteins overall, 12 of those were found in each replicate. The structural proteins AaCP19-1 through -5, AaCP43-1 through -3, and AaCP52-2 and -3 can be expected to be consistently identified with HFIP, and although this condition identified few enzymes and pheromones, AaPxt-1 and AaMulti-1 were identified in all 3 HFIP replicates. Nineteen core proteins were identified by the methanol condition, and while these core proteins included structural proteins, enzymes, and pheromones, none of these protein categories had deep coverage. The urea condition, alternatively, provided 40 core proteins. In fact, only 2 of the total core proteins (AaCP52-2 and AaPx-2) were not found to be core proteins for urea.

Discussion

Acorn barnacles produce limited amounts of water-insoluble adhesive in an obstructed location, making this material notoriously difficult to analyze. The body of the barnacle is confined within calcified parietal (side shell) and opercular plates which obscure the interface in acorn barnacles, where radial expansion of the adhesive interface takes place in a manner correlated to their molt cycle [50, 51]. Further, this interfacial region is itself only a few microns thick, and, in some species, lies beneath a calcified base plate. While the adhesive itself is primarily composed of protein [52], the presence of carbohydrates and lipid at this critical region have also been noted [14, 53–55]. New details related to the formation and composition of barnacle adhesive have recently come to light based on confocal microscopy of live barnacles [8, 14, 55] and novel approaches to break down the adhesive interface [6].

Direct observation of the expanding adhesive interface in live adult acorn barnacles (specifically *A. amphitrite*) has revealed a complex spatiotemporal process [8, 14, 26, 55]. Lipid, protein, and carbohydrate development and deposition occur at various time points throughout the molt cycle. An epidermal layer, located at the leading radial edge (and tens of microns removed from the interface) and extending toward the center of the barnacle, has a critical role in providing the building blocks for the multi-layered interface [14], including protein delivery. Further, the delivery of lipids and reactive oxygen species to the leading edge contributes to surface cleaning, an important priming step for adhesion [14]. A well-described capillary network with distributed ducts terminating close to the leading edge have traditionally been assigned the role of directly providing cement [25, 26] or molting fluid [26] to the leading edge interface. Burden *et al.* (2012) showed that the molt process, including development of the capillary networks and mineralization, coincided with increased adhesion but also later showed that a layer of protein is present and delivered to the interface before the capillary network is completed [55]. The experiments of Saroyan *et al.* suggested the role of the capillary network is related to sclerotization [25], though direct biophysical evidence confirming the composition of chemistries in the capillary network is lacking for acorn barnacles. While the process of acorn barnacle adhesive formation is not fully understood, it is critical to understand the *composition* of the adhesive itself and do so in an efficient manner with a minimal amount of starting material – a primary goal of the current investigation.

Recent advances have been made in dissolution of the adhesive with HFIP and resulted in the identification of novel proteins [6, 17], yet enhanced methods will enable processing samples from single organisms and allow investigations of barnacle adhesive relative to biophysical or environmental conditions. In this work, PCT methods were developed to process the hydrated, thick adhesive often produced by barnacles when grown on silicone substrates [36]. High pressure aided in barnacle interface protein solubilization, leading to the identification and characterization of additional proteins at the interface. The PCT method offers substantial methodological improvements by subjecting the adhesive to high levels of pressure in the presence of a concentrated solvent, the most effective being urea, which led to the most reliable identification (i.e. consistent over multiple replicates) of the greatest number of proteins. Elevated pressure has been shown to increase proteolysis efficiency and decrease digestion time compared to ambient pressure conditions, meaning that PCT methods can be completed in under 8 hours - a massive reduction in the amount of time necessary to prepare barnacle adhesive using in-gel methods for proteomics (typically spanning at least 2-3 days [57]). There is also an increase in throughput as 16 samples can be processed simultaneously and with significantly less starting material (~1–2 mg of adhesive) for analysis. In addition to reducing the volume and time of the workload during processing, the PCT method led to the identification of a majority of the proteins already known to form the adhesive and to the discovery of novel proteins at the interface.

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019

Around 80% of the *A. amphitrite* adhesive proteins previously identified with HFIP and gel-based methods [6] were also found in this study using PCT. A majority of the proteins were identified with a high number of minimum peptides (5), though ~30% were identified only when the threshold was dropped below 4 minimum peptides. The proteins found at the interface can be divided into three main predicted functional groups: bulk structural proteins, enzymes, and pheromones. Twenty-six of the previously identified proteins predicted to be bulk proteins were further grouped into 6 subgroups (19-, 43-, 52-, 57-, 100-, and 105-like proteins). Of these, the majority (23) were identified using PCT. Those proteins not identified were a single instance of AaCP19-like (AaCP19-9) and 2 instances of AaCP43-like (AaCP43-5 and -6) proteins. AaCP19-9 and AaCP43-5 were only identified from transcriptomics and have not been identified in the adhesive using proteomics. Using the clustering algorithm in Scaffold, AaCP19-7 and -8 were clustered together, which highlights the similar sequence between the two proteins. Of the seventeen enzymes that have been identified at the surface interface [6, 17, 27], peroxinectin- (AaPxt), lysyl oxidase- (AaLOx), serine protease inhibitor- (AaPI), peroxidase- (AaPx), and peptidase- (AaPep) like proteins were identified in the present study, but whey acid proteins (AaWAP) were not. Three variations of MULTIFUNCin (AaMulti), a settlement aggregation pheromone [16], have also been found in the adhesive [6, 17], and two of these (AaMulti-1 and -2) were identified here.

The solvent conditions examined in conjunction with PCT proved to have a large effect on protein identification. Methanol and HFIP are both organic solvents, while urea is a chaotropic agent. Organic solvents and chaotropes aid in protein solubilization in distinct manners: organic solvents stabilize hydrophobic regions that interact in aqueous conditions while chaotropes disrupt protein interactions and stabilize the resulting unfolded protein via hydrogen bonding and electrostatic interactions [58]. Cement solubilization via PCT in 8 M urea resulted in the greatest number of spectra per sample, indicating that this treatment provided the greatest extent of adhesive solubilization and subsequently the highest identification of both spectra and proteins (74). The 60% methanol condition had the second largest group of identified proteins (46) with only 3 unique proteins identified compared to the 8 M urea condition. In general, these two conditions had relatively similar amino acid, peptide, and protein profiles. Of the 22 proteins identified in the 100% HFIP condition, only two were unique to this condition. HFIP also resulted in differences in the type of amino acids as well as the peptide and protein molecular weight profiles observed, indicating that proteolysis was affected. In particular, the average peptide molecular weight was significantly larger using 100% HFIP compared to 60% MeOH or 8 M urea, yet the average protein molecular weight was much smaller. Among all conditions examined, these data indicate 8 M urea was the optimal cement solubilization technique with PCT, while the dilute HFIP performed the worst and addition of HFIP to urea or methanol was detrimental to protein identification. Although the 100% HFIP condition resulted in the fewest peptide/protein identifications, this condition did result in high solubilization of AaCP19- and AaCP43-like proteins, which are difficult to solubilize and are thought to be integral to barnacle adhesive function [6]. Differences in protein identification from HFIP extracts in this work with PCT to previous gel-based methods [6] can be attributed to the decrease in protease function and to the types of samples analyzed. Here, only the thickened adhesive was examined, whereas previously multiple sample types (thickened adhesive, adhesive deposited on glass beads, and 'plaques' – composed of the basal cuticle and the underlying layer of hard cement) were subjected to HFIP. From the perspective of building a more complete protein profile of barnacle adhesive, PCT with 8 M urea provides the most information, though it is also clear that different solvents are able to solubilize different proteins and groups of proteins. These results suggest that the adhesive contains proteins spanning a wide array of physical and chemical properties. As such, a 'one-size-fits-all' sample preparation protocol may not yield a representative profile of the adhesive, potentially explaining why identifications presented herein are the 'deepest' to date.

The application of PCT to the barnacle adhesive allowed for the identification of several new proteins at the *A. amphitrite* interface. Only one of these novel proteins was identified in the 100% HFIP condition (AaCP105-5), an unsurprising result as HFIP has been used to dissolve and characterize *A. amphitrite* adhesive proteins previously. The majority of the novel proteins were identified using urea as a solvent, although two (AaHem and AaSP-4) were identified only in the methanol condition.

To characterize these novel proteins, sequences or domains homologous to publically available data were identified. Characterization of barnacle adhesive proteins has often relied on identifying enriched amino acids to infer function [6, 23], yet the growth of publically available data have made a sequence similarity based analysis feasible here. Of the 81 proteins identified in this study, 8 were left without any assigned putative function or identity. These proteins share no sequence similarity with any known proteins or any of the other interface proteins. Sequence similarity and previous analysis [6, 17] led us to divide the remaining 73 interface proteins into three putative functional categories – structural proteins, enzymes, or pheromones. One protein, AaCP43-7, was added to the structural protein AaCP43-like family because it had significant homology to the partial sequence of AaCP43-2, which had been previously deposited to NCBI by our research group (GenBank: AQA26374.1). Five proteins were also added to the structural protein group because they had sequence similarity to the AaCP105-like (AaCP105-3 through -5) and AaMuc-like (AaMuc-1 and -2) families. A novel protein family was identified, AaCP34, where the 5 members had similar sequences to one another. The AaCP34-like proteins had no homologous sequences or domains to other known proteins.

Another 7 proteins were added to the structural protein category because their putative identities indicate they could serve some sort of adhesive function. Two proteins (AaTitin) showed weak to moderate homology to a predicted Titin protein (NCBI Accession Number: XP_015810555.1). Titin is a giant protein that imparts elasticity to muscle fibers [59]. The AaTitin proteins have no conserved domains and are small proteins (predicted MW = 33 & 20 kDa). Their sequence similarity to a predicted titin-like protein (XP_015810555.1) appears to arise from repetitive valine residues. Another group of three proteins (AaGlut) show partial homology to a protein annotated as a predicted glutenin from *Plutella xylostella*, the diamondback moth (XP_011568564.1), despite glutenin being a plant protein that imparts elasticity to wheat dough [60]. The AaGlut-like proteins contain no homology to any conserved domains, but the *P. xylostella* glutenin-like protein has a putative Glutenin_hmw super family domain, suggesting convergent evolution [61] may have occurred to impart this protein with elastomeric properties. Another protein (AaFas) contains two repeating predicted fasciclin domains and has homology to multiple transforming growth factor-β-induced protein ig-h3-like proteins, which often contain four repeating fasciclin domains [62]. Ig-h3-like proteins

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019

are collagen-like and can be secreted into fiber rich tissues, and their facsiclin domains may function in integrin binding [63]. AaCP19- and AaCP43-like proteins have been shown to have repetitive complex and simple domains and homology to silk proteins, indicating that nanofibrils formed by these proteins may be important for barnacle adhesion [6, 64].

Four of the interface proteins have homology to and share domains with proteins involved in invertebrate immunity (AaHem: hemocytin; AaVit-1 and -2: vitellogenin; AaβGBPP: β-1,3-glucan-binding protein precursor). Hemocytin functions as an invertebrate lectin that binds to carbohydrates and promotes hemagglutination in response to pathogens or foreign structures [65, 66]. Vitellogenin is a multifunctional protein that, aside from its primary role in providing nutrients to developing embryos, binds to carbohydrates and promotes clotting via a conserved von Willebrand factor domain (vWD) [67, 68]. The barnacle vitellogenins contain vWD and many of the other conserved domains [38] found in all vitellogenin sequences [69]. β-1,3-glucan-binding protein initiates crustacean immune responses after binding to β-1,3-glucan [70–72], a carbohydrate found in fungal and algal cell walls. The protein aggregative and cross-linking processes that take place during immunological responses were previously proposed to have been evolutionarily repurposed in barnacle adhesive polymerization [27]. The ability of these proteins to bind to carbohydrates is also of interest as the proteinaceous adhesive is in direct contact with a carbohydrate and chitin-rich cuticle layer [55], yet no chitin-binding proteins have been previously found via mass spectrometry at the surface. Four transcripts containing chitin binding GO terms were identified from the prosoma region of *T. japonica formosana* [22], but whether those proteins are located in the adhesive region has not been examined. Here, AaCP105-1 and -3 and AaHem were found to contain putative chitin-binding domains (ChtBD2, of the CBM_14 superfamily). These immune-related proteins along with the AaCP105-family proteins may therefore function at the interface by binding to the cuticle layer and promoting protein aggregation.

Over a quarter of the identified proteins have homology that suggests potential enzymatic function. These enzymes include peroxinectins, lysyl oxidases, peroxidases, a peptidase, and serine proteases and inhibitors. The presence of these enzymes at the interface has previously been noted, as has the activity of peroxidases and lysyl oxidases [6, 17, 25, 28]. Oxidation occurs during cuticular sclerotization in arthropods [73, 74], and may occur during barnacle adhesive curing by oxidation of AaCP43-1 [17]. Eight potential serine proteases and six protease inhibitors were also identified. Serine proteases and inhibitors regulate the phenoloxidase cascade that controls invertebrate immunity by protease activation of pro-tyrosinases involved in melanization [72, yet tyrosinases have not been identified in this work or in previous proteomics analysis of the interfacial material [6, 17]. Proteases are also involved in the degradation of the cuticle during molting in insects and crustacea [75–78], either by direct digestion of the cuticular protein matrix or by activating other necessary enzymes. Acorn barnacles form a new layer of folded cuticle at the base which stretches as radial growth occurs. When the cuticle cannot stretch any more, molting occurs and the old cuticle tears to make room for the expansion of a new layer of folded cuticle [14, 55, 79, 80]. The proteases and protease inhibitors found at the surface interface may therefore play a dual role in molting and adhesion, as has been previously suggested [27].

Pheromones were also identified in the proteins at the surface interface, as previously reported [6, 17]. The MULTIFUNCin glycoproteins localize to the barnacle itself, and are intra- and interspecies chemical communication cues, attracting both settling barnacle larvae as well as predators [16]. AaMulti-5 had homology to α_2-macroglobulin proteins and contained multiple proteinase inhibitor (A2M and ISOPREN_C2) and receptor (A2M_recep) domains, features shared with all other AaMulti proteins as well as settlement inducing protein complex (SIPC). α_2-macroglobulin proteins are active in the innate immune system [81], leading Ferrier *et al.* (2016) to hypothesize that the communication function of the orthologous proteins MULTIFUNCin and SIPC arose from an ancestral immune function.

Three of the identified proteins had homology to a deposited sequence identified as waterborne settlement pheromone (WSP; GenBank: BAM34601.1); this is likely the same ~30 kDa protein described by Endo *et al.* [82], which, unlike MULTIFUNCin and SIPC, is released into the water column [83]. Proteolytically cleaved peptides of SIPC are also likely released into the water column [84], but the amino acid sequence of the WSP proteins is distinct from the α_2-macroglobulin-like pheromones, indicating that WSP and SIPC are distinct. The only domain found in WSP is Cupin_5 (cl01418), although the level of homology is low (E value < 1.e-5). Cupin domain containing proteins have a diverse array of functions and are found in prokaryotes and eukaryotes [85], making a prediction of protein function based on the presence of this domain difficult.

It is important to note the transcriptome generated from sub-mantle tissue [21] was used to construct a protein database for *A. amphitrite* and has made untargeted proteomics possible [6, 17, 38], yet progress is limited by a lack of genomic information. Many of the peptides detected by the mass spectrometer are ambiguously assigned to multiple proteins with similar predicted sequences, resulting in many protein clusters. Likewise, many of these proteins with similar sequences have high homology when comparing their sequences with BLAST. In the absence of a complete genome sequence, whether these proteins are isoforms or are produced by unique genes is unknown. The protein database is also limited because it was made from transcripts in only one tissue, and does not contain all of the genetic information of *A. amphitrite*. Future efforts to produce a genome would aid in all proteomic analyses of *A. amphitrite*.

Conclusions

The insoluble nature of barnacle adhesive has simultaneously fascinated scientists while confounding characterization efforts. Effective solubilization of proteins in *A. amphitrite* adhesive has been demonstrated using PCT and the results validate an effective and relatively fast method of cement solubilization with minimal starting material. The variety of solvent conditions tested reveal the majority of identified proteins were found in the 8 M urea, though different conditions do present different protein profiles and may be more appropriate depending on the experimental questions being asked. These methods would likely be applicable to a number of other insoluble biological samples that are medically and environmentally relevant.

The expansion of the number of identified proteins within the cement layer of *A. amphitrite* extends the general knowledge of its composition, which historically has been a severe bottleneck to understanding the mechanism of cement production. How the interface proteins work in a concerted manner to form the permanent adhesive is unknown, yet the identified proteins at the interface have a variety of potential functions which indicate that a complex balance of protein-protein and protein-carbohydrate interactions along with oxidative and protease

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019

processes underlie adhesive formation. Many of the identified proteins have no sequence similarity to publically available proteins, suggesting that unique processes evolved by barnacles may contribute to adhesion and have yet to be characterized.

Supplementary data

Supplementary data is available at *INTBIO* online.

Acknowledgments

The authors thank Mr. Shrey Patel for help processing the samples for proteomics analysis during his summer Naval Research Enterprise Internship Program at the Naval Research Laboratory, and Drs. Chenyue Wang and Christopher So for discussion.

Funding

This work was supported by the internal Basic Research Program at the United States Naval Research Laboratory through the Office of Naval Research and the National Research Council Post-Doctoral Associateship Program [J.N.S and S.N.D].

References

1. Tan S, Tan HT, Chung MC. Membrane proteins and membrane proteomics. *Proteomics* 2008;**8**:3924–32.
2. Vit O, Petrak J. Integral membrane proteins in proteomics. How to break open the black box? *J Proteomics* 2017;**153**:8–20.
3. Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. *Annual Rev Biochem* 2006;**75**:333–66.
4. Walther DM, Kasturi P, Zheng M et al. Widespread proteome remodeling and aggregation in aging C. *elegans*. *Cell* 2015; **161**:919–32.
5. Bleem A, Christiansen G, Madsen DJ et al. Protein engineering reveals mechanisms of functional amyloid formation in *Pseudomonas aeruginosa* biofilms. *J Mol Biol* 2018;**430**:3751–63.
6. So CR, Fears KP, Leary DH et al. Sequence basis of barnacle cement nanostructure is defined by proteins with silk homology. *Sci Rep* 2016;**6**:36219.
7. Barlow DE, Dickinson GH, Orihuela B et al. Characterization of the adhesive plaque of the barnacle *Balanus amphitrite*: amyloid-like nanofibrils are a major component. *Langmuir* 2010;**26**:6549–56.
8. Burden DK, Barlow DE, Spillmann CM et al. Barnacle balanus amphitrite adheres by a stepwise cementing process. *Langmuir* 2012;**28**:13364–72.
9. Nakano M, Kamino K. Amyloid-like conformation and interaction for the self-assembly in barnacle underwater cement. *Biochemistry* 2015;**54**:826–35.
10. Schultz M, Bendick J, Holm E et al. Economic impact of biofouling on a naval surface ship. *Biofouling* 2011;**27**:87–98.
11. Nir S, Reches M. Bio-inspired antifouling approaches: the quest towards non-toxic and non-biocidal materials. *Curr Opin Biotechnol* 2016;**39**:48–55.
12. Stewart RJ, Ransom TC, Hlady V. Natural underwater adhesives. *J Polym Sci B Polym Phys* 2011;**49**:757–71.
13. Kamino K. Molecular design of barnacle cement in comparison with those of mussel and tubeworm. *J Adhes* 2010;**86**: 96–110.
14. Fears KP, Orihuela B, Rittschof D et al. Acorn barnacles secrete phase-separating fluid to clear surfaces ahead of cement deposition. *Adv Sci* 2018;**5**:1700762.
15. Matsumura K, Nagano M, Fusetani N. Purification of a larval settlement-inducing protein complex (SIPC) of the barnacle, *Balanus amphitrite*. *J Exp Zool A Ecol Genet Physiol* 1998;**281**: 12–20.
16. Ferrier GA, Kim SJ, Kaddis CS et al. MULTIFUNCin: a multifunctional protein cue induces habitat selection by, and predation on, barnacles. *Integr Comp Biol* 2016;**56**:901–13.
17. So CR, Scancella JM, Fears KP et al. Oxidase activity of the barnacle adhesive interface involves peroxide-dependent catechol oxidase and lysyl oxidase enzymes. *ACS Appl Mater Interfaces* 2017;**9**:11493–505.
18. Naldrett M, Kaplan D. Characterization of barnacle (*Balanus eburneus* and *B. cenatus*) adhesive proteins. *Mar Biol* 1997;**127**: 629–35.
19. Kamino K, Inoue K, Maruyama T et al. Barnacle cement proteins importance of disulfide bonds in their insolubility. *J Biol Chem* 2000;**275**:27360–5.
20. Kamino K, Odo S, Maruyama T. Cement proteins of the acorn-barnacle, *Megabalanus rosa*. *Biol Bull* 1996;**190**:403–9.
21. Wang Z, Leary DH, Liu JN et al. Molt-dependent transcriptomic analysis of cement proteins in the barnacle *Amphibalanus amphitrite*. *BMC Genomics* 2015;**16**:859.
22. Lin H-C, Wong YH, Tsang LM et al. First study on gene expression of cement proteins and potential adhesion-related genes of a membranous-based barnacle as revealed from Next-Generation Sequencing technology. *Biofouling* 2014;**30**:169–81.
23. Rocha M, Antas P, Castro LFC et al. Comparative Analysis of the Adhesive Proteins of the Adult Stalked Goose Barnacle *Pollicipes pollicipes* (Cirripedia: Pedunculata). *Mar Biotechnol* 2019;**21**:38–51.
24. Jonker J-L, Abram F, Pires E et al. Adhesive proteins of stalked and acorn barnacles display homology with low sequence similarities. *PLoS One* 2014;**9**:e108902.
25. Saroyan J, Lindner E, Dooley C. Repair and reattachment in the balanidae as related to their cementing mechanism. *Biol Bull* 1970;**139**:333–50.
26. Golden JP, Burden DK, Fears KP et al. Imaging active surface processes in barnacle adhesive interfaces. *Langmuir* 2016;**32**: 541–50.
27. Dickinson GH, Vega IE, Wahl KJ et al. Barnacle cement: a polymerization model based on evolutionary concepts. *J Exp Biol* 2009;**212**:3499–510.
28. Rittschof D, Orihuela B, Harder T et al. Compounds from silicones alter enzyme activity in curing barnacle glue and model enzymes. *PLoS One* 2011;**6**:e16487.
29. Tao F, Li C, Smejkal G et al. Pressure cycling technology (PCT) applications in extraction of biomolecules from challenging biological samples. *High Press Biosci Biotechnol* 2007;**1**: 166–73.
30. López-Ferrer D, Petritis K, Hixson KK et al. Application of pressurized solvents for ultrafast trypsin hydrolysis in proteomics: proteomics on the fly. *J Proteome Res* 2008;**7**:3276–81.
31. Clare AS, Høeg JT. *Balanus amphitrite* or *Amphibalanus amphitrite*? A note on barnacle nomenclature. *Biofouling* 2008;**24**: 55–7.
32. Clare AS, Rittschof D. What's in a name? *Nature* 1989;**338**:627.
33. Holm ER, Orihuela B, Kavanagh CJ et al. Variation among families for characteristics of the adhesive plaque in the barnacle *Balanus amphitrite*. *Biofouling* 2005;**21**:121–6.
34. Berglin M, Gatenholm P. The barnacle adhesive plaque: morphological and chemical differences as a response to substrate properties. *Colloids Surf B Biointerfaces* 2003;**28**:107–17.

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019

35. Wiegemann M, Watermann B. Peculiarities of barnacle adhesive cured on non-stick surfaces. *J Adhes Sci Technol* 2003;**17**:1957–77.
36. Ramsay DB, Dickinson GH, Orihuela B et al. Base plate mechanics of the barnacle *Balanus amphitrite* (= *Amphibalanus amphitrite*). *Biofouling* 2008;**24**:109–18.
37. Guo T, Kouvonen P, Koh CC et al. Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps. *Nat Med* 2015;**21**:407–13.
38. Wang C, Schultzhaus JN, Taitt CR et al. Characterization of longitudinal canal tissue in the acorn barnacle *Amphibalanus amphitrite*. *PLoS One* 2018;**13**:e0208352.
39. Zhang G, He L-s, Wong Y-H et al. Chemical component and proteomic study of the *Amphibalanus* (= *Balanus*) amphitrite shell. *PLoS One* 2015;**10**:e0133866.
40. Vizcaíno JA, Csordas A, Del-Toro N et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 2016; **44**:D447–D56.
41. Team RC. R: A language and environment for statistical computing. 2015.
42. Osorio D, Rondón-Villarrea P, Torres R. Peptides: a package for data mining of antimicrobial peptides. *R Journal* 2015;**7**: 4–14.
43. Warnes GR, Bolker B, Bonebakker L et al. gplots: Various R programming tools for plotting data. *R package version* 2009;**2**:1.
44. Gasteiger E, Gattiker A, Hoogland C et al. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 2003;**31**:3784–8.
45. Altschul SF, Gish W, Miller W et al. Basic local alignment search tool. *J Mol Biol* 1990;**215**:403–10.
46. Marchler-Bauer A, Bo Y, Han L et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 2017;**45**:D200–D3.
47. Boman H. Antibacterial peptides: basic facts and emerging concepts. *J Intern Med* 2003;**254**:197–215.
48. Gibbs GM, Roelants K, O'bryan MK. The CAP Superfamily: Cysteine-Rich Secretory Proteins, Antigen 5, and Pathogenesis-Related 1 Proteins—Roles in Reproduction, Cancer, and Immune Defense. *Endocr Rev* 2008;**29**:865–97.
49. Brejc K, van Dijk WJ, Klaassen RV et al. Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* 2001;**411**:269.
50. Power AM, Klepal W, Zheden V et al. Mechanisms of adhesion in adult barnacles. In: von Byern J, Grunwald I (eds). *Biol Adhes Syst*. Vienna: Springer, 2010, 153–68.
51. Fyhn UE, Costlow JD. A histochemical study of cement secretion during the intermolt cycle in barnacles. *Biol Bull* 1976;**150**:47–56.
52. Walker G. The biochemical composition of the cement of two barnacle species, *Balanus hameri* and *Balanus crenatus*. *J Mar Biol Assoc U. K* 1972;**52**:429–35.
53. Essock-Burns T, Gohad NV, Orihuela B et al. Barnacle biology before, during and after settlement and metamorphosis: a study of the interface. *J Exp Biol* 2017;**220**:194–207.
54. Gohad NV, Aldred N, Hartshorn CM et al. Synergistic roles for lipids and proteins in the permanent adhesive of barnacle larvae. *Nat Commun* 2014;**5**:4414.
55. Burden DK, Spillmann CM, Everett RK et al. Growth and development of the barnacle Amphibalanus amphitrite: time and spatially resolved structure and chemistry of the base plate. *Biofouling* 2014;**30**:799–812.
56. Darwin C. A monograph on the sub-class Cirripedia, with figures of all the species. Volume 2. In: *The Balanidæ (or sessile cirripedes); the Verrucidæ, etc*. London: Ray Society, 1854, 25, 684 pp.
57. Shevchenko A, Tomas H, Havli J et al. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 2006;**1**:2856.
58. Dormeyer W, van Hoof D, Mummery CL et al. A practical guide for the identification of membrane and plasma membrane proteins in human embryonic stem cells and human embryonal carcinoma cells. *Proteomics* 2008;**8**:4036–53.
59. Trinick J. Cytoskeleton: Titin as a scaffold and spring. *Curr Biol* 1996;**6**:258–60.
60. Shewry P, Halford N, Tatham A. High molecular weight subunits of wheat glutenin. *J Cereal Sci* 1992;**15**:105–20.
61. Storz JF. Causes of molecular convergence and parallelism in protein evolution. *Nat Rev Genet* 2016;**17**:239.
62. Zinn K, McAllister L, Goodman CS. Sequence analysis and neuronal expression of fasciclin I in grasshopper and Drosophila. *Cell* 1988;**53**:577–87.
63. Kawamoto T, Noshiro M, Shen M et al. Structural and phylogenetic analyses of RGD-CAP/βig-h3, a fasciclin-like adhesion protein expressed in chick chondrocytes. *Biochim Biophys Acta* 1998;**1395**:288–92.
64. So CR, Yates EA, Estrella LA et al. Molecular recognition of structures is key in the polymerization of patterned barnacle adhesive sequences. *ACS Nano* 2019;**135**:5172–83.
65. Kotani E, Yamakawa M, Iwamoto S-i et al. Cloning and expression of the gene of hemocytin, an insect humoral lectin which is homologous with the mammalian von Willebrand factor. *Biochim Biophys Acta* 1995;**1260**:245–58.
66. Lesch C, Goto A, Lindgren M et al. A role for Hemolectin in coagulation and immunity in *Drosophila melanogaster*. *Dev Comp Immunol* 2007;**31**:1255–63.
67. Zhang S, Wang S, Li H, Li L. Vitellogenin, a multivalent sensor and an antimicrobial effector. *Int J Biochem Cell Biol* 2011; **43**:303–5.
68. Baker M. Invertebrate vitellogenin is homologous to human von Willebrand factor. *Biochem J* 1988;**256**:1059.
69. Hayward A, Takahashi T, Bendena WG et al. Comparative genomic and phylogenetic analysis of vitellogenin and other large lipid transfer proteins in metazoans. *FEBS Lett* 2010; **584**:1273–8.
70. Duvic B, Söderhäll K. Purification and characterization of a beta-1, 3-glucan binding protein from plasma of the crayfish Pacifastacus leniusculus. *J Biol Chem* 1990;**265**:9327–32.
71. Cerenius L, Liang Z, Duvic B et al. Structure and biological activity of a 1,3-beta-D-glucan-binding protein in crustacean blood. *J Biol Chem* 1994;**269**:29462–7.
72. Cerenius L, Söderhäll K. The prophenoloxidase-activating system in invertebrates. *Immunol Rev* 2004;**198**:116–26.
73. Andersen SO. Cuticular sclerotization and tanning. In: *Insect Molecular Biology and Biochemistry*. Elsevier, 2012, 167–92.
74. Andersen SO. Insect cuticular sclerotization: a review. *Insect Biochem Mol Biol* 2010;**40**:166–78.
75. Wei Z, Yin Y, Zhang B et al. Cloning of a novel protease required for the molting of *Locusta migratoria manilensis*. *Dev Growth Differ* 2007;**49**:611–21.
76. Warner A, Matheson C. Release of proteases from larvae of the brine shrimp *Artemia franciscana* and their potential role during the molting process. *Comp Biochem Physiol B Biochem Mol Biol* 1998;**119**:255–63.
77. Gimenez AF, Garcıa-Carreno F, Del Toro MN et al. Digestive proteinases of *Artemesia longinaris* (Decapoda, Penaeidae) and relationship with molting. *Comp Biochem Physiol B Biochem Mol Biol* 2002;**132**:593–8.

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019

78. Gimenez AF, Garcıa-Carreno F, Del Toro MN et al. Digestive proteinases of red shrimp *Pleoticus muelleri* (Decapoda, Penaeoidea): partial characterization and relationship with molting. *Comp Biochem Physiol B Biochem Mol Biol* 2001;**130**:331–8.
79. Bourget E. *Barnacle shells: composition, structure and growth*. In: Southward AJ (ed.). *Barnacle Biology*. Rotterdam: AA Balkema, 1987, 267–85.
80. Anderson DT. *Barnacles: Structure, Function, Development and Evolution*. London: Chapman & Hall, 1994.
81. Armstrong PB, Quigley JP. α2-macroglobulin: an evolutionarily conserved arm of the innate immune system. *Dev Comp Immun* 1999;**23**:375–90.
82. Endo N, Nogata Y, Yoshimura E et al. Purification and partial amino acid sequence analysis of the larval settlement-inducing pheromone from adult extracts of the barnacle, *Balanus amphitrite* (= *Amphibalanus amphitrite*). *Biofouling* 2009;**25**:429–34.
83. Clare AS, Matsumura K. Nature and perception of barnacle settlement pheromones. *Biofouling* 2000;**15**:57–71.
84. Rittschof D, Cohen JH. Crustacean peptide and peptide-like pheromones and kairomones. *Peptides* 2004;**25**: 1503–16.
85. Dunwell JM, Purvis A, Khuri S. Cupins: the most functionally diverse protein superfamily? *Phytochemistry* 2004;**65**: 7–17.

Downloaded from https://academic.oup.com/ib/article-abstract/11/5/235/5524728 by guest on 28 July 2019