The complex barnacle perfume: identification of waterborne pheromone homologues in *Balanus improvisus* and their differential expression during settlement

Anna Abramova, Ulrika Lind, Anders Blomberg and Magnus Alm Rosenblad

Department of Chemistry and Molecular Biology, University of Gothenburg, Gothenburg, Sweden

**ABSTRACT**

A key question in barnacle biology is the nature of cues that induce gregarious settlement. One of the characterised cues is the waterborne settlement pheromone (WSP). This study aimed to identify WSP homologues in *Balanus improvisus* and to investigate their expression during settlement. Six WSP homologues were identified, all containing an N-terminal signal peptide, a conserved core region, and a variable C-terminus comprising several -GR- and -HDDH- motifs. The *B. improvisus* WSP homologues were expressed in all settlement stages but showed different expression patterns. The homologue most similar to the *B. amphitrite* WSP was the most abundant and was constantly expressed during settlement. In contrast, several of the other WSP homologues showed the greatest expression in the juvenile stage. The presence of several WSP homologues suggests the existence of a pheromone mix, where con-specificity might be determined by a combination of sequence characteristics and the concentration of the individual components.

**ARTICLE HISTORY**

Received 12 December 2018
Accepted 24 March 2019

**KEYWORDS**

Barnacles; settlement cues; waterborne settlement pheromone; *Balanus improvisus*; gregarious settlement

**Introduction**

Barnacles are among the most important fouling organisms in the marine environment. During settlement, the barnacle cyprid larvae respond to chemical and physical cues in the surrounding environment while swimming in the water column as well as when exploring various surfaces. The surface exploration comprises wide search, close search, and inspection behaviours that precede the settling decision leading to permanent adhesion through the release of cement (Crisp 1976; Aldred et al. 2018). Because adult barnacles are permanently attached, cyprid substratum selection is the main factor determining the local abundance and distribution of barnacles in natural populations as well as on marine constructions (eg ships’ hulls), and is therefore a key question in barnacle biology and ecology (Thiyagarajan 2010). Understanding the mechanisms underlying cyprid attraction and substratum selection will facilitate the tailoring of anti fouling technologies resulting in the development of surfaces that are actively rejected by cyprids during the initial stages of the surface exploration, thus preventing attachment.

Gregariousness in barnacles, ie the attraction of conspecific individuals and the subsequent settling and aggregation in dense communities, is crucial for reproduction and is achieved through cyprid settlement in response to chemical cues produced by individuals of the same species (Clare and Matsumura 2000). Since Knight-Jones and Stevenson (1950) described gregarious barnacle settlement behaviour for the first time almost 70 years ago, considerable advances have been made in the understanding of the mechanisms of this behaviour. Early research in gregarious settlement led to the discovery that extracts of whole adults contain a ‘settling factor’ that induces conspecific settlement of cyprids (Crisp and Meadows 1962). Later studies resulted in the identification of the contact pheromone in *Balanus amphitrite* (=*Amphibalanus amphitrite*), called settlement-inducing protein complex (SIPC) (Matsumura et al. 1998), the cDNA sequence of which was cloned and sequenced (Dreanno et al. 2006). It was shown that SIPC is a 171 kDa cuticular glycoprotein with similarity to the α2-macroglobulin protein family. SIPC is active when bound to a surface (Matsumura et al.
proteins in the available databases. The full sequence did not show any resemblance to SIPC or to any other peptide (Clare and Yamazaki 2000). Later, a protein corresponding to 32 kDa was reproduced by another study (Clare and Matsumura 2000). In addition, several synthetic di- and tripeptides were tested for settlement-inducing activity. Altogether, this indicated that more than one WSP gene is present in B. improvisus (Lin et al. 2014) and none are deposited in public databases. Moreover, the study by Endo et al. (2009) mentioned preliminary unpublished data showing that several proteins with a molecular mass of around 32 kDa detected in barnacle-conditioned seawater have settlement-inducing activity. In particular, the sea slug Aplysia releases a pheromone blend comprising more than three different types of waterborne pheromones to attract mates (Cummins et al. 2005). The indications of the presence of several WSP homologues in some barnacle species led the present authors to explore the presence of WSP homologues in B. improvisus using transcriptome data from cyprids and adults. Several WSP candidates in B. improvisus were identified, cloned, and compared to homologues present in B. amphitrite. In addition, gene expression analysis of the WSP homologues in B. improvisus during the settling process was performed showing that the genes are clearly differentially expressed. This study adds yet another dimension to the complexity of barnacle pheromones forming a basis for a better understanding of the gregarious behaviour as well as providing essential information for developing highly barnacle-selective antifouling strategies.

Materials and methods
Identification and cloning of WSP homologues in B. improvisus
Blast searches of several different RNA datasets (unpublished data) from cyprids and adults of
B. improvisus were performed using B. amphitrite WSP (BAM34601) as a query. Six sequences that constituted full or partial fragments of WSP-like sequences were found, five of which were cloned and confirmed by Sanger sequencing. The sixth sequence was identified from a single molecule PacBio RNA dataset obtained from an adult of B. improvisus (Lind et al. 2017). The sequences identified were named according to the level of identity to the published B. amphitrite WSP, e.g., the Bi_WSP clone was 74% identical and 91% similar at the protein level to the B. amphitrite WSP (Table 1), while the WSP-like 1–5 sequences had a gradually lower (in the range of 39%–70% identity and 71%–88% similarity) level of identity/similarity that followed the gene numbering.

Cloning of the cDNA containing the complete open reading frame (ORF) for Bi_WSP and Bi_WSP-like 2–5 was done by first performing RACE using PCR primers based on sequences found in a Sanger-sequenced cDNA library prepared from B. improvisus cyprids (unpublished data). RNA preparations of cyprids and cDNA synthesis were carried out as described in Lind et al. (2010). 3′-RACE was performed for all five pheromones (GeneRacer kit; Invitrogen), and 5′-RACE was performed for Bi_WSP-like 3. As a first step, a touch-down PCR was performed using the following PCR program: an initial denaturing step of 98 °C for 2 min followed by 5 cycles of 98 °C for 30 s, 72 °C for 1 min, 5 cycles of 98 °C for 30 s, 70 °C for 1 min, and 25 cycles of 98 °C for 30 s, 65 °C for 30 s, and 72 °C for 1 min. A final elongation step of 72 °C for 7 min was added. A nested PCR was thereafter performed using 1 μl of the touch-down reaction applying the following PCR program: an initial denaturing step of 98 °C for 2 min, followed by 35 cycles of 98 °C for 30 s, 65 °C for 30 s, and 72 °C for 1 min and a final elongation step of 72 °C for 7 min. The polymerase PfuUltra (Stratagene) was used in all PCR reactions. Sequences obtained from RACE were used to design primers to clone the complete ORF for the genes using the same PCR program as for the nested PCR, but with an annealing temperature of 60 °C. For the primers used for the cloning, see Supplemental material Table S1. Bi_WSP-like 1 was not cloned. Instead, a full-length sequence corresponding to Bi_WSP-like 1 was identified from an adult single-molecule PacBio RNA dataset (Lind et al. 2017), where the corresponding consensus contig covering the ORF was supported by more than 30 PacBio reads. The sequences have been deposited in GenBank with accession numbers MK275628–MK275633. To find the sequences that are most divergent from the B. improvisus WSP homologues but still having the B. amphitrite WSP as the best match in NCBI nr, the identified B. improvisus WSP and WSP-like 1–5 protein sequences were used as queries for searching in the authors’ different datasets. The search resulted in the identification of one additional sequence showing 30% identity or less to any of the queries, and this was named Bi_pb29993 according to the PacBio contig.

**Searching for WSP homologues in other species**

To determine whether barnacles other than B. improvisus also have several WSP homologues, the available larval and adult B. amphitrite NCBI/SRA datasets (SRX120025/SRR426833, SRX1035030/31, SRX1035107/27) were downloaded and transcriptome assemblies were created using Trinity 2.0.6 (Haas et al. 2013). These B. amphitrite assemblies were searched using the six identified B. improvisus WSP homologues. To avoid using wrongly assembled sequences, only B. amphitrite WSP-like candidates that were found in at least two assemblies were used for further sequence analysis. All candidate B. amphitrite ORFs were checked vs NCBI nr to make sure they were most similar to the published B. amphitrite WSP and not more similar to other types of proteins. B. amphitrite sequences that had the best match to the B. amphitrite WSP in NCBI nr searches were regarded as WSP-like candidates and named according to their contig numbers.

| Bi_WSP | Bi_WSP-like 1 | Bi_WSP-like 2 | Bi_WSP-like 3 | Bi_WSP-like 4 | Bi_WSP-like 5 |
|--------|---------------|---------------|---------------|---------------|---------------|
| Ba_WSP | 100%(100%)    | 74%(91%)      | 70%(88%)      | 66%(86%)      | 51%(78%)      |
| Bi_WSP | 74%(91%)      | 100%(100%)    | 80%(93%)      | 64%(84%)      | 55%(81%)      |
| Bi_WSP-like 1 | 70%(88%) | 80%(93%) | 100%(100%) | 59%(82%) | 56%(80%) |
| Bi_WSP-like 2 | 66%(86%) | 64%(84%) | 59%(82%) | 100%(100%) | 41%(70%) |
| Bi_WSP-like 3 | 51%(78%) | 55%(81%) | 56%(80%) | 41%(70%) | 100%(100%) |
| Bi_WSP-like 4 | 51%(78%) | 48%(79%) | 50%(78%) | 54%(82%) | 42%(74%) |
| Bi_WSP-like 5 | 39%(71%) | 38%(71%) | 42%(72%) | 38%(72%) | 36%(69%) |

**Table 1.** Pairwise identity and similarity (in parenthesis) of the B. amphitrite WSP (BAM34601) and the identified B. improvisus WSP homologues.
SignalP and TargetP available at http://www.cbs.dtu.dk/services/. All six *B. improvisus* WSP homologues as well as *Bi_pb29993* had weak matches (see Supplemental material Table S2 for the details) to the Pfam cupin_5 domain (PF06172), referred to as the RmlC-like cupin domain in InterPro. Sequences with a better match to the cupin_5 protein domain were also identified in both *B. improvisus* and *B. amphitrite* and were named *Bi_cupin* and *Ba_cupin*, respectively.

To further investigate whether crustaceans other than barnacles have proteins similar to the WSP homologues, an alignment of the six *B. improvisus* WSP homologues with the signal peptide and C-terminus removed was used to create an HMM profile with the HMMER 3.0 package (Potter et al. 2018). A set of all available 867,743 crustacean proteins was downloaded from the NCBI protein database (2018-05-02) and searched using the obtained WSP-specific HMM profile. The HMM search identified 27 sequences including the *B. amphitrite* WSP and 26 sequences from either *Daphnia* or *Eurytemora*. Among these 26 sequences, none had the *B. amphitrite* WSP as the best Blastp match in NCBI nr nor were they more similar to the *B. improvisus* WSP homologues than to the barnacle cupin_5 representatives in the present authors’ dataset. To identify any other domains than cupin_5 that were present in the identified sequences, the protein sequences were searched against the collection of profiles in Pfam with hmmscan HMMER 3.0 (Potter et al. 2018). The search showed that there were no other domains than cupin_5 in the identified sequences, with scores well above the trusted Pfam cut-off (37.2–98.3).

**Protein alignments**

The *B. improvisus* and *B. amphitrite* WSP homologue protein sequences were aligned using Clustalw 2.0.10 and visualised with Jalview 2.10.1. The alignment revealed the presence of a well-aligned central region of ~200 aa with a few gaps (Figure 1b) that is referred to as the ‘core region’ throughout the text. The fasta36 package (Pearson 2016) was used to create the identity table for the core region of the identified *B. improvisus* WSP homologues and the *B. amphitrite* WSP (Table 1). The Zappo sequence colour scheme was used to show the physico-chemical properties of amino acids in the C-terminal region (Figure 2 and Figure S1B in the Supplemental material).

**Phylogenetic trees**

The *B. improvisus* and *B. amphitrite* WSP homologues identified were used for phylogenetic analysis. The *Daphnia* cupin_5, the closest homologue to the *B. improvisus* WSPs found among crustaceans, was used as an outgroup together with *B. improvisus* and *B. amphitrite* cupin_5 representatives.

The two methods used for the phylogenetic analysis were MrBayes 3.2.6 and PhyML 3.0 at phylogeny.fr. The ClustalW alignment of the identified *B. improvisus* and *B. amphitrite* WSP homologues and cupin-5 proteins (Figure S1A) was used as input into MrBayes 3.2.6 with one cold and three hot chains running for 100,000 generations at which point the average standard deviation was < 0.01. Alignments of full-length sequences or core regions only were used as input to PhyML.

**Expression analysis of the *B. improvisus* waterborne pheromones identified**

Transcriptomic data on settlement stages were used to investigate the expression of the *B. improvisus* WSP homologues identified, and the details of the experimental set up and the sampling of the different settling stages are described in Abramova et al., 2018 (manuscript). Briefly, cyprid larvae of *B. improvisus* were reared in a laboratory culture as described in Jonsson et al. (2018). For the expression analysis, 1–2-day-old cyprids were settled on Petri dishes (Nunc No 150340, Ø 48 mm) for 4–5 days. Four different settling stages were collected: free-swimming, close-search, attached cyprid, and juvenile. Each stage was represented by three biological replicates to ensure the generality of the findings (three independent batches of cyprids each coming from different sets of *B. improvisus* parents) with 20 individuals in each. Extracted RNA was used to make sequencing libraries with the Illumina TruSeq Stranded mRNA protocol following poly-A selection, and these were multiplexed and sequenced on three Illumina lanes. Library preparation and Illumina sequencing were performed at the national genomics facility at the Science for Life Laboratory in Stockholm (https://www.scilifelab.se/platforms/ngi/). Obtained RNA-seq data were assembled with the Trinity 2.0.6 software (Haas et al. 2013), and the previously identified *B. improvisus* WSP homologues were used to identify the corresponding contigs in the transcriptome assembly. The sequences of *Bi_WSP* and *Bi_WSP-like 2* were complete in the assembly, whereas the other WSP homologues were either found as partial contigs or were split between different Trinity contigs. To make a statistical analysis possible, all the fragmented WSP-like contigs were removed and substituted with the
corresponding WSP-like sequences with complete ORFs that were previously identified. This edited assembly was used as the template for read mapping and to calculate the transcript abundances using the pseudoalignment-based quantification method kallisto (Bray et al. 2016) through the use of a Trinity 2.0.6 wrapper. The resulting abundances were used to generate a matrix of counts and a TMM (trimmed mean
of M-values)-normalised matrix. Lowly-expressed genes were filtered out by keeping only genes with at least 2 CPM (corresponding to 60 counts on average) in at least two samples across the entire experiment. The preliminary data quality check indicated a pronounced batch effect reflecting the different parental backgrounds of the cyprid batches. To minimise the batch effect, the script removeBatchEffect from the edgeR 3.4.2 was applied to the count matrix (Robinson et al. 2010). The differential gene expression analysis was performed using the edgeR package. Genes with false discovery rate (FDR) 0.05 and a logFC > 2 were considered to be significantly differentially expressed.

Results

Identification of WSP-like sequences in B. improvisus

To identify sequences in B. improvisus that are similar to the published WSP in B. amphitrite, tBlastn searches of RNA-seq datasets obtained from both cyprids and adults of B. improvisus were performed using the published B. amphitrite WSP as a query. Six unique B. improvisus RNA-contigs were found, of which five encoded partial WSP proteins. These sequences were used to obtain full-length variants of all WSP homologues in B. improvisus. The protein sequence identity between the six obtained B. improvisus homologues and the B. amphitrite WSP was 39%–74% (Table 1), with Bi_WSP being the most similar. The numbering of the WSP-like homologues in B. improvisus is related to their sequence identity to the Bi_WSP, with Bi_WSP-like 1 being the highest (80%). All B. improvisus protein sequences contained a predicted N-terminal signal peptide, a rather well-conserved central core region, and a highly variable C-terminal region (Figure 1). Four of the six identified B. improvisus WSP homologues were of almost the same length (255–266 aa) while Bi_WSP-like 3 and Bi_WSP-like 5 (217 and 222 aa, respectively) had a roughly 40 aa shorter C-terminus (Figure 1 and Table 2).

Except for the relatively high sequence identity to WSP from B. amphitrite, the WSP homologues in B. improvisus did not have any reliable matches to other proteins in the databases. However, all B. improvisus WSP homologues contained a region with weak resemblance to the cupin_5 domain (PF06172) (Table S2). Cupin_5 is a domain of unknown function, and proteins containing this domain belong to the cupin superfamily (http://www.ebi.ac.uk/interpro/entry/IPR009327). In the Pfam database, for metazoans there are currently only 17 sequences from 10 species indicated to have a cupin_5 domain. Among these sequences, the scores for the predicted cupin_5 domain varied from 41.3 to 139.2 with the only two crustacean representatives having scores of 70.5 or 78.7, both coming from Daphnia (see Table S3 in Supplemental material). The Pfam cupin_5 domain scores for the B. improvisus WSP homologues varied between 19.8 and 36.3, with some of them thus being below the Pfam cupin_5 profile threshold score of 25.0 (Table S2). However, proteins with higher cupin_5 Pfam scores in both
B. improvisus and B. amphitrite (with scores 45.4 and 46.0) were found, and these were named Bi_cupin and Ba_cupin, respectively. The Bi_cupin sequence was only 15% identical to Bi_WSP. Even if some of the WSP homologues in B. improvisus have a domain with weak similarity (slightly above the threshold score) to cupin_5, for the ease of the following discussion only proteins with high domain scores, ie Bi_cupin, Ba_cupin and Daphnia cupin_5, will be referred to as ‘cupin_5 proteins’ further in the text.

A search for more WSP-like sequences in B. improvisus, using the six initially identified B. improvisus WSP homologues as queries, resulted in the identification of one additional sequence showing a roughly 30% identity to the other WSPs and 15% identity to the Bi_cupin. This sequence was named Bi_pb29993 according to the name of the assembly contig. An alignment with the WSP-like sequences (Figure S1A) showed that the Bi_pb29993 protein was rather different from the six B. improvisus WSP homologues, but it still had the B. amphitrite WSP as the best match in the NCBI protein database.

Analysing the sequences of the B. improvisus WSP homologues showed that the long C-terminus found in most homologues was clearly distinct from the conserved core region by containing many charged residues (D/E, R/H/K) as well as glycines, with several -GR- and -HDDH- motifs or close variants of these (Figure 2). It should be noted that the frequent -HDDH- and -GR- motifs in the C-terminal regions of the identified WSP homologues vary greatly in their number, position, and composition, making it difficult to align this particular region of the sequences.

### Additional WSP-like sequences in B. amphitrite

By using the identified WSP homologues from B. improvisus as queries for searching the published B. amphitrite RNA datasets, three more B. amphitrite WSP homologues in addition to the published Ba_WSP (Endo et al. 2009) were identified, of which two were encoding complete proteins (Figure 1b). Furthermore, two B. amphitrite sequences with the best match either to Bi_cupin or Bi_pb29993 were identified.

The alignment of the identified WSP homologues (Figure 1b) clearly shows that the B. amphitrite WSP-like sequences are structurally similar to the B. improvisus sequences, with an N-terminal signal peptide, a conserved core region, and the greatest difference found in the C-terminal region. The C-terminus of three of the B. amphitrite sequences, including the published Ba_WSP, had a prevalence of charged amino acids similar to what was found for the B. improvisus homologues Bi_WSP and Bi_WSP-like 2 and 4 (Figure 2 and Figure S1B in Supplemental material), while the fourth B. amphitrite sequence had a shorter C-terminus resembling B. improvisus WSP-like 3 and 5.

### Phylogenetic analysis

Phylogenetic analyses showed that all B. improvisus WSP homologues were found in a well-supported core clade together with the B. amphitrite WSP, except for the Bi_WSP-like 5, which formed a sister group (Figure 3). The Bi_WSP-like 3 seems to be more closely related to Bi_WSP than to Bi_WSP-like 5, even though both Bi_WSP-like 3 and 5 lack the long, charged C-terminus (proteins with a long C-terminus are marked with an asterisk in Figure 3). The WSP homologues from B. improvisus and B. amphitrite were not divided into separate clades, thus suggesting that several WSP homologues were already present before these two species diverged. While the phylogenetic method used (either Bayesian or Maximum Likelihood) did not influence the overall topology of the phylogenetic tree, it did change the exact branching within the core clade indicating the difficulties in resolving the true topology in this part of the tree (Figure S2).

Overall, the phylogenetic analysis provided strong evidence that the B. improvisus WSP homologues identified are clearly more closely related to the B. amphitrite WSP than to the cupin_5 proteins (Daphnia cupin_5, Bi_cupin, and Ba_cupin).

### Expression of B. improvisus WSP candidates during settlement

The WSP homologues identified from B. improvisus showed different expression patterns during settlement (Figure 4). The expression of Bi_WSP did not change significantly during settlement progression and was by far the highest, with roughly a hundred-fold higher expression compared to the other WSPs in the cyprid stages (Figure 4g). In contrast, Bi_WSP-like 1 was overall the least-expressed candidate in any of the settling stages (note the different scales on the y-axes). Bi_WSP-like 2, Bi_WSP-like 3, and Bi_WSP-like 5 showed similar expression patterns, being relatively lowly expressed in the cyprids and in the attachment stage and substantially upregulated when
cyprids metamorphosed into juveniles; for example, Bi_WSP-like 5 underwent about a 7-fold upregulation in the juvenile stage compared to the attached stage (Figure 4f). In summary, the metamorphosed juvenile stage of B. improvisus displayed substantial expression of all WSP homologues, with Bi_WSP having the highest expression.

Discussion

Since the initial publication of the discovery of WSP in 2009 (Endo et al. 2009), there have been several studies indicating that more than one WSP gene is present in some barnacle species (Lin et al. 2014; So et al. 2017; Wang et al. 2018). Originally, Endo et al. (2009) reported the results of a preliminary study that detected several proteins with weights around 32 kDa in B. amphitrite-conditioned seawater that exhibited settlement-inducing activity. In addition, several WSP sequences were found in the water-soluble fraction of proteins rinsed from the cement (So et al. 2017) and among the proteins from longitudinal canal and sub-mantle tissues of B. amphitrite (Wang et al. 2018). In adults of T. j. formosana, Lin et al. (2014) reported the presence of two different WSP sequences displaying different expression in prosoma and basis, suggesting the existence of more than one WSP also in this species. However, only one of the sequences from these studies was published. The indications of the presence of several WSP homologues in some barnacle species led the present authors to explore the presence of WSP homologues in B. improvisus using transcriptome data from cyprids and adults. Six sequences that are homologous to the published B. amphitrite WSP were identified from B. improvisus. The similar size and isoelectric point observed for several of the WSP homologues identified here (data not shown) suggest that in the previous studies based
on protein purifications (Endo et al. 2009; So et al. 2017) a single band on protein gels corresponding to 32 kDa could theoretically represent co-purification of several different proteins.

Phylogenetic analysis clearly showed that five of the six identified WSP sequences from *B. improvisus* proteins belong to the same clade as WSP homologues from *B. amphitrite*, suggesting that several WSPs were already present before the two species diverged. Despite some of the sequences appearing to form orthologous groups in the phylogenetic analysis (Figure 3), it is not presently possible to infer the true orthologous relationships between the WSP-like candidates from *B. improvisus* and *B. amphitrite* due to the lack of complete genome-based proteomes for any of these barnacle species. Furthermore, it should be pointed out that it is not yet possible to establish the true number of WSP family homologues in *B. improvisus* and *B. amphitrite* using only transcriptome data because some of the WSP genes might be expressed at extremely low levels at the life stages and tissues examined and thus not be detected. The phylogenetic analysis presented here should therefore be seen as a first step to aid in the classification of WSP-like sequences.

Although an evolutionary relationship is not evidence of similar function, it is clear that the identification of several WSP-like sequences in these species suggests that WSP is a substantially larger protein family than previously thought and that multiple WSP paralogues also exist in other barnacle species.

There is currently no evidence for the presence of WSP homologues in any crustaceans other than

---

**Figure 4.** Gene expression of WSP homologues and Bi_pb29993 during settlement. (a–f) The relative expression of the WSP homologues; (g) The relative expression of the WSP homologues and Bi_pb29993. The order of the bars in the graph corresponds to the order of the names in the legend. FS: free-swimming, CS: close-search, AT: attached cyprid, JUV: juvenile.
barnacles, suggesting that the WSPs are a relatively late evolutionary invention and thus a protein family that might be unique to barnacles. However, part of the core region of the WSP homologues has some similarity to the cupin_5 domain found in both prokaryotes and eukaryotes. Nevertheless, the sequence and phylogenetic analyses indicated that the WSPs are clearly distinct from other cupin_5 domain proteins found in B. improvisus, B. amphitrite, and Daphnia. Moreover, identification of the cupin_5 domain does not provide any information about function because the cupin_5 family is poorly characterised (Zhou et al. 2005; Gaowa and Zhang 2009; Du et al. 2010).

Analysis of the sequence-features of the identified WSP homologues revealed that all have quite conserved N-terminal signal peptides and core regions, with the greatest variability residing in the C-terminal regions. The importance of the C-terminus for water-soluble pheromones was noticed many years ago because digestion with carboxypeptidases that cleave carboxyl-terminal lysine and arginine residues eliminated pheromone activity (Tegtmeyer and Rittschof 1988). Furthermore, the C-terminus has a striking number of charged amino acids compared to the core region (Figure 2 and Figure S1B). A charged, basic C-terminus has previously been suggested to be present in a small barnacle waterborne cue of 3-5 kDa (≈ 40 aa) (Tegtmeyer and Rittschof 1988). Interestingly, the sizes of the peptides that showed potent induction of settlement in the study by Tegtmeyer and Rittschof (1988) correspond to the average size of the C-termini in the WSP homologues reported here. This opens up the possibility that the short peptides reported (Rittschof 1985; Tegtmeyer and Rittschof 1988; Clare and Matsumura 2000) with settlement-inducing activity are derived from cleavage of the WSPs, releasing biologically active peptides from the C-termini as earlier proposed (Rittschof 1993). A similar mechanism was found in lobsters and hermit crabs, where olfactory serine proteases have been suggested to cleave after arginine and lysine residues in inactive odorant peptide molecules to generate active chemical cues (Rittschof 1990; Levine et al. 2001). Notably, the charged amino acids in the C-terminus of some of the identified WSP homologues are organised in repeated motifs. The most commonly observed were -HDDH- and -GR- motifs (Figure 2), but some of the WSP homologues also contained variations of these motifs, eg -HDEH-, -HDH-, and -GGR-. Interestingly, the latter motif was previously reported to be important for settlement-inducing activity in both barnacle and oyster larvae (Zimme-Faust and Tamburri 1994; Browne and Zimmer 2001).

In comparison to SIPC, WSP is largely under-investigated. An important piece of missing evidence is whether WSP is actually released into the water by adult organisms in nature and functions as a waterborne cue, as earlier suggested (Endo et al. 2009). Further experimental studies are required to establish whether the multiple WSP homologues identified here in both B. amphitrite and B. improvisus act as pheromones. However, assuming that the WSP homologues exhibit pheromone activity, the current study brings a new level of complexity to the mechanisms of barnacle chemical communication and barnacle settlement ecology. The presence of several WSP-like candidates in barnacles is consistent with previously reported data on pheromones in some insects, nematodes, and marine invertebrates that consist of a mix of several components. It has been suggested that the specificity of these pheromones is determined by the relative concentrations of the different components in the mixture (Linn et al. 1984; Roelofs 1995; Cummins et al. 2004). Moreover, evidence exists that the concentration of the different pheromone components depends on age, sex, and habitat, thus providing additional information to the individuals receiving the pheromone signal (Coppée et al. 2011; Choe et al. 2012). One particularly well-studied example of pheromone mixtures among marine invertebrates is the pheromone blend in the sea slug Aplysia (Cummins et al. 2005). Mate attraction in these animals involves several waterborne pheromones such as attractin, seductin, and temptin, that in different combinations act to maintain mating aggregations (Cummins et al. 2005). In barnacles, according to the current hypothesis, cyprids detect the waterborne cue released by adults while swimming and respond to it by transitioning from the water column to a substratum where the surface-bound SIPC pheromone in turn induces permanent attachment (Elbourne and Clare 2010).

Details on how the waterborne cue is released, distributed, and sensed by cyprids remain unknown. However, studies on host location by cyprids of the parasitic barnacle Heterosaccus dollfusi (Pasternak et al. 2004) showed that larvae use chemoreception to initiate the motion and rheoreception (sensing direction of the flow) to follow the host’s odour plume upstream until locating the organism releasing the odour. This study also suggested that larvae do not use the concentration gradient of the attractant to find the direction but rather its presence or absence in the water flow. In the context of a pheromone mixture, it is therefore
probable that cyprids’ first encounter the most abundant WSP that travels the longest distance from where it was released. Following the odour plume, cyprids will sense the other more lowly expressed WSP homologues that would serve as a confirmation of the direction towards the odour source as well as possibly trigger downward movement thus bringing cyprids closer to the surface where the cue is released by already established adults.

The relative abundance of the WSPs in combination with the sequence differences between species might be the basis for the species specificity of the cue. However, no obvious and consistent amino acid differences between the WSPs from B. amphitrite and B. improvisus were observed in the core region. Instead, the C-terminus with its intriguing repetitions of acidic and basic residues might contain species-specific combinations of motifs. Although there is some evidence that glycosylation patterns might be involved in the species specificity of the SIPC protein (Yorisue et al. 2012), nothing currently indicates that this type of post-translational modification plays a role in the activity of WSP. On the contrary, Endo et al. (2009) found that lentil lectin treatment (which binds to and inhibits sugars on proteins) did not influence the pheromone activity of their WSP fraction. On the other hand, the waterborne cue might not necessarily be species specific as has been shown for Aplysia attractins, which are the only peptide pheromone family in invertebrates known not to be species specific (Painter et al. 2003). Behavioural studies on Aplysia also showed that fewer individuals were attracted by a single component than by a pheromone blend, suggesting that pheromones act in concert (Painter et al. 2003). The present data show that different barnacle life stages express different combinations of WSPs. In particular, the cyprids expressed mostly Bi_WSP, whereas juveniles expressed all six identified WSP homologues. It should be pointed out that newly metamorphosed juvenile barnacles were found to be particularly attractive to cyprids (Crisp and Meadows 1962). Therefore, the blend of several WSPs produced by juveniles might signal to cyprids that it is a suitable place to successfully settle and metamorphose.

Except for the reported pheromone function of the published B. amphitrite WSP (Endo et al. 2009), neither pheromone activity nor any other function for the WSP homologues identified here from B. amphitrite or B. improvisus have been experimentally examined. In addition to inducing cyprid settlement, the WSP homologues might have completely different or dual functions. For instance, the identification of WSP among the proteins rinsed from cement collected from adult B. amphitrite (So et al. 2017) suggests that WSPs might be a part of the permanent adhesive from where they leak into the water and induce cyprid settlement. A patent has been issued for an adhesive 35 kDa protein in B. improvisus, B. crenatus, and S. balanoides extracted from the hemolymph or base plate (Kaplan et al. 2003) that contains an N-terminal fragment of about 20 amino acids that exactly matches the Bi_WSP downstream of the signal peptide region (starting at aa number 16). Furthermore, the Semibalanus sequence (File S1), which is a protein of 245 aa and found both in the hemolymph and the base plate, is clearly a WSP homologue with 67%, 65%, and 62% identity to Bi_WSP-like 2, Ba_WSP, and Bi_WSP, respectively. Additionally, the triad -HDDH- motif in the C-terminus of this 35 kDa Semibalanus sequence (mentioned in the same patent) is found in Bi_WSP-like 4 with a G separating the three -HDDH- motifs, whereas Bi_WSP-like 2 contains three repetitions of the alternative -HDH- motif. It should be noted that a repeated -HDH- motif in the C-terminus is also found in two of the B. amphitrite WSP sequences (Figure 2). It is interesting that these motifs consistently occur in all the barnacle species mentioned, albeit sometimes with small differences, especially because the C-terminus is the least conserved part of these WSP homologues. The presence and particular combinations of motifs might indicate differences in function for the different WSP homologues, whether acting as pheromones or as a part of the cement.

In the current study, several WSP homologues in B. improvisus were identified and characterised. Further experimental studies are needed to determine the function of all of the WSP homologues, especially what features of WSPs are behind the settling-inducing effect and whether there is a functional connection to the properties of the adhesives. Although several WSP-like sequences exist in both B. improvisus and B. amphitrite, a complete set of genes encoding WSP-like proteins in any barnacle species can only be obtained using a high-quality genome as a reference. Despite that, the identification of a set of WSP homologues is an important discovery that will first of all promote further research into the great complexity of barnacle chemical communication. Furthermore, it is clear from the present study that the WSP homologues seem to be barnacle-specific and thus might be used as targets for non-toxic antifouling agents analogous to pheromone-based applications in insect management.
strategies (Witzgall et al. 2010). Finally, the protein sequences of the identified WSP homologues are now available in the NCBI database, and this will aid in the annotation of WSPs in future studies.

Acknowledgements

The authors would like to acknowledge support from the SciLifeLab National Genomics Infrastructure (NGI) and Uppmax for providing assistance in massive parallel sequencing and computational infrastructure, respectively. The authors would also like to thank Martin Oegmark for maintaining the B. improvisus culture and providing cyprid larvae, Johan Lagmo for technical assistance during the initial phase of the project in cloning some of the WSP-like genes from B. improvisus, and Per Jonsson for valuable comments on the manuscript.

Declaration of interest statement

No potential conflicts of interest were reported by the authors.

Funding

Funding for this project was provided to AB by the Swedish Research Council (VR, #2017-04559), the EU Commission (SEAFRONT—Synergistic Fouling Control Technologies; #614034) and Swedish Research Council for Sustainable Development (Formas; #210-2011-1733), as well as to MAR from Carl Tryggers Foundation (CTS11:14). This work was supported by the Centre for Marine Evolutionary Biology at the University of Gothenburg (http://www.cemeb.science.gu.se/). MAR was partially funded by National Bioinformatics Infrastructure Sweden (NBIS).

ORCID

Anna Abramova http://orcid.org/0000-0002-0493-7808
Ulrika Lind http://orcid.org/0000-0003-1746-9043
Anders Blomberg http://orcid.org/0000-0002-1260-3920
Magnus Alm Rosenblad http://orcid.org/0000-0002-4077-7821

References

Aldred N, Alsaab A, Clare AS. 2018. Quantitative analysis of the complete larval settlement process confirms Crisp’s model of surface selectivity by barnacles. Proc R Soc B. 285:20171957. doi:10.1098/rspb.2017.1957
Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 34:525–527. doi:10.1038/nbt.3519
Browne KA, Zimmer RK. 2001. Controlled field release of a waterborne chemical signal stimulates planktonic larvae to settle. Biol Bull. 200:87–91. doi:10.2307/1543088
Choe A, Chuman T, von Reuss SH, Dossey AT, Yim JJ, Ajredini R, Kolawa AA, Kaplan F, Alborn HT, Teal PEA et al. 2012. Sex-specific mating pheromones in the nematode Panagrellus redivivus. P Natl Acad Sci USA. 109:20949–20954. doi:10.1073/pnas.1218302109
Clare AS, Matsumura K. 2000. Nature and perception of barnacle settlement pheromones. Biofouling. 15:57–71. doi:10.1080/08927010009386298
Clare AS, Yamazaki M. 2000. Inactivity of glycy-glycyl-arginine and two putative (Q SAR) peptide analogues of barnacle waterborne settlement pheromone. J Mar Biol Assoc UK. 80:945–946. doi:10.1017/S0025315400002952
Coppee A, Mathy T, Cammaerts M-C, Verheggen FJ, Terzo M, Iserbyt S, Valterović I, Rasmont P. 2011. Age-dependent attractiveness of males’ sexual pheromones in Bombus terrestris (L.) [Hymenoptera, Apidae]. Chemoecology. 21:75–82. doi:10.1007/s00049-011-0070-x
Crisp D. 1976. Settlement responses in marine organisms. Butterworths, London.
Crisp DJ, Meadows P. 1962. The chemical basis of gregariousness in cirripedes. Proc R Soc Lond B Biol Sci. 156:500–520.
Cummins SF, Nichols AE, Amare A, Hummon AB, Swedler JV, Nagle GT. 2004. Characterization of Aplysia enticin and temptin, two novel water-borne protein pheromones that act in concert with attractin to stimulate mate attraction. J Biol Chem. 279:25614–25622. doi:10.1074/jbc.M313585200
Cummins SF, Nichols AE, Warso CJ, Nagle GT. 2005. Aplysia seducin is a water-borne protein pheromone that acts in concert with attractin to stimulate mate attraction. Peptides. 26:351–359. doi:10.1016/j.peptides.2004.10.024
Dreanno C, Kirby RR, Clare AS. 2007. Involvement of the barnacle settlement-inducing protein complex (SIPC) in species recognition at settlement. J Exp Mar Biol Ecol. 351:276–282. doi:10.1016/j.jembe.2007.07.003
Dreanno C, Matsumura K, Dohmae N, Takio K, Hirota H, Kirby RR, Clare AS. 2006. An alpha2-macroglobulin-like protein is the cue to gregarious settlement of the barnacle Balanus amphitrite. Proc Natl Acad Sci U S A. 103:14396–14401. doi:10.1073/pnas.0602763103
Du Y, He YX, Gaowa S, Zhang X, Chen Y, Zhang SC, Zhou CZ. 2010. Crystal structures of the apo and GDP-bound forms of a cupin-like protein BbDUF985 from Branchiostoma belcheri tsingtauense. Proteins. 78:2714–2719.
Elbourne PD, Clare AS. 2010. Ecological relevance of a conspecific, waterborne settlement cue in Balanus amphitrite (Cirripedia). J Exp Mar Biol Ecol. 392:99–106.
Endo N, Nogata Y, Yoshimura E, Matsumura K. 2009. Purification and partial amino acid sequence analysis of the larval settlement-inducing pheromone from adult extracts of the barnacle Balanus amphitrite (＝Amphibalanus amphitrite). Biofouling. 25:429–434. doi:10.1080/08927010902875113
Ferrier GA, Kim SJ, Kaddis CS, Loo JA, Ann Zimmer C, Zimmer RK. 2016. MULTIFUNCin: a multifunctional protein cue induces habitat selection by, and predation on, barnacles. Integr Comp Biochem. 56:901–91. doi:10.1016/j.icb.2016.01.010
Gaowa S, Zhang S. 2009. Identification, expression, function and localization of a DUF985 domain-containing hypothetical gene from amphioxus Branchiostoma belcheri. Gen. Comp Biochem Physiol B Biochem Mol Biol. 152:28–37. doi:10.1016/j.gcbpb.2008.09.085
Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J,Couger MB, Eccles D, Li B, Lieber M
et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 8:1494–1512. doi: 10.1038/nprot.2013.084

Jonsson PR, Wrange AL, Lind U, Abramova A, Ogemark M, Blomberg A. 2018. The barnacle Balanus improvisus as a marine model - culturing and gene expression. J Vis Exp.

Kaplan D, Gatenholm P, Berglin K, Platko J, Pepper L, Ngangan A, inventors; Kaplan, D, assignee. 2003 patent No. 10,134,507.

Kelly DR. 1996. When is a butterfly like an elephant? Chem Biol. 3:595–602. doi:10.1016/S1074-5521(96)90125-8

Knight-Jones E, Stevenson J. 1950. Gregariousness during settlement in the barnacle Elminius modestus Darwin. J Mar Biol Assoc UK. 29:281–297. doi:10.1017/S0025315400055375

Kotsiri M, Protopapa M, Mouratidis S, Zachariadis M, Vassilakos D, Kleidas I, Samiotaki M, Dedos SG. 2018. Should I stay or should I go? The settlement-inducing protein complex guides barnacle settlement decisions. J Exp Biol. 221:jeb185348. doi:10.1242/jeb.185348

Levine MZ, Harrison PJ, Walthall WW, Tai PC, Derby CD. 2001. A CUB-serine protease in the olfactory organ of the spiny lobster Panulirus argus. J Neurobiol. 49:277–302. doi:10.1002/neu.10010

Lin HC, Wong YH, Tsang LM, Chu KH, Qian PY, Chan BKK. 2014. First study on gene expression of cement proteins and potential adhesion-related genes of a membrane-based barnacle as revealed from Next-Generation Sequencing technology. Biofouling. 30:169–181. doi:10.1080/08927014.2013.853051

Lind U, Jarva M, Rosenblad MA, Pingitore P, Karlsson E, Wrangle AL, Kamdal E, Sundell K, Andre C, Jonsson PR et al. 2017. Analysis of aquaporins from the euryhaline barnacle Balanus improvisus reveals differential expression in response to changes in salinity. Plos One. 12.

Lind U, Rosenblad MA, Frank LH, Falkbringer S, Brive L, Laurila JM, Pohjanoksa K, Vuorenpää A, Kukkonen JP, Gunnarsson L. 2010. Octopamine receptors from the barnacle Balanus improvisus are activated by the α2-adrenoceptor agonist medetomidine. Mol Pharmacol. 78:80–90. doi:10.1124/mol.110.008788

Linn C, Bjostad L, Du J, Roelofs W. 1984. Redundancy in a chemical signal: Behavioral responses of male Aplysia to a 6-component sex pheromone blend. J Chem Ecol. 10:1635–1640. doi:10.1080/08927014.2012.695776

Lorang WR. 2016. Finding protein and nucleotide similarities with FASTA. Curr Protoc Bioinformatics. 53:3–9.

Matsuura K, Nagano M, Fusetani N. 1998. Purification of a larval settlement-inducing protein complex (SIPC) of the barnacle, Balanus amphitrite. J Exp Zool. 281:12–20. doi:10.1002/(SICI)1097-010X(19980501)281:1<12::AID-JEZ3>3.0.CO;2-F

Painter SD, Clough B, Black S, Nagle GT. 2003. Behavioral characterization of attractin, a water-borne peptide pheromone in the genus Aplysia. Biol Bull. 205:16–25. doi:10.2307/1543441

Pasternak Z, Blasius B, Abelson A. 2004. Host location by larvae of a parasitic barnacle: larval chemotaxis and plume tracking in flow. J Plankton Res. 26:487–493. doi:10.1093/plankt/fbh040

Pearson WR. 2016. Finding protein and nucleotide similarities with FASTA. Curr Protoc Bioinformatics. 53:3–9.

Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. 2018. HMMER web server: 2018 update. Nucleic Acids Res. 46:200–204.

Rittschof D. 1985. Oyster drills and the frontiers of chemical ecology: unsettling ideas. Bull Am Mar Union. 1:111–116.

Rittschof D. 1990. Peptide-mediated behaviors in marine organisms Evidence for a common theme. J Chem Ecol. 16:261–272. doi:10.1007/BF01021283

Rittschof D. 1993. Body odors and neutral-basic peptide mimics: a review of responses by marine organisms. Am Zool. 33:487–493.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 26:139–140. doi:10.1093/bioinformatics/btp616

Roelofs WL. 1995. Chemistry of sex attraction. Proc Natl Acad Sci U S A. 92:44–49. doi:10.1073/pnas.92.1.44

So CR, Scancella JM, Fears KP, Essock-Burns T, Haynes SE, Leary DH, Diana Z, Wang CY, North S, Oh CS et al. 2017. Oxidase activity of the barnacle adhesive interface involves peroxide-dependent catechol oxidase and l-lysyl oxidase enzymes. ACS Appl Mater Inter. 9:11493–11505. doi:10.1021/acsami.7b01185

Tegtmeier K, Rittschof D. 1988. Synthetic peptide analogs to barnacle settlement pheromone. Peptides. 9:1403–1406. doi:10.1016/0169-4758(88)90209-4

Thiyagarajan V. 2010. A review on the role of chemical cues in habitat selection by barnacles: new insights from larval proteomics. J Exp Mar Biol Ecol. 392:22–36. doi:10.1016/j.jembe.2010.04.030

Wang C, Schultzhaus JN, Taitt CR, Leary DH, Shriver-Lake LC, Snellings D, Sturiale S, North SH, Orihuela B, Rittschof D. 2018. Characterization of longitudinal canal tissue in the acorn barnacle Amphibalanus amphitrite. Plos One. 13:e0208352. doi:10.1371/journal.pone.0208352

Witzgall P, Kirsch P, Cork A. 2010. Sex pheromones and their impact on pest management. J Chem Ecol. 36:493. doi:10.1007/s10886-009-9737-y

Yorisue T, Matsumura K, Hirota H, Dohmae N, Kojima S. 2012. Possible molecular mechanisms of species recognition by barnacle larvae inferred from multi-specific sequencing analysis of proteinaceous settlement-inducing pheromone. Biofouling. 28:605–611. doi:10.1080/08927014.2012.695776

Zimmer RK, Tambrurri MN. 1994. Chemical identity and ecological implications of a waterborne, larval settlement cue. Limnol Oceanogr. 39:1075–1087. doi:10.4319/lo.1994.39.5.1075

Zimmer RK, Ferrier GA, Kim SJ, Kaddis CS, Zimmer CA, Loo JA. 2016. A multifunctional chemical cue drives opposing demographic processes and structures ecological communities. Ecology. 97:2232–2239. doi:10.1002/ecy.1455