© 2018. Published by The Company of Biologists Ltd | Journal of Experimental Biology (2018) 221, jeb185348. doi:10.1242/jeb.185348

RESEARCH ARTICLE

Should I stay or should I go? The settlement-inducing protein complex guides barnacle settlement decisions

Manto Kotsiri[1], Maria Protopapa[1], Sofoklis Mouratidis[1], Michael Zachariadis[1,2], Demetrios Vassilakos[1], Ioannis Kleidas[1], Martina Samiotaki[3] and Skarlatos G. Dedos[1,*]

ABSTRACT

Reproduction in barnacles relies on chemical cues that guide their gregarious settlement. These cues have been pinned down to several sources of settlement pheromones, one of which is a protein termed settlement-inducing protein complex (SIPC), a large glycoprotein acting as a pheromone to induce larval settlement and as an adhesive in surface exploration by the cyprids. Settlement assays in laboratory conditions with *Amphibalanus* (=*Balanus*) *amphitrite* cyprids in the presence of SIPC showed that cyprids exhibit settlement preference behaviour at lower concentrations of SIPC [half maximal effective concentration (EC_{50})=3.73 nmol l^{-1}] and settlement avoidance behaviour at higher concentrations (EC_{50}=101 nmol l^{-1}). By using truncated fragments of SIPC in settlement assays, we identify that domains at the N-terminus of SIPC transduce settlement preference cues that mask the settlement avoidance cues transduced by domains at its C-terminus. Removing the N-terminal 600 amino acids from SIPC resulted in truncated fragments that transduced only settlement avoidance cues to the cyprids. From the sexual reproduction point of view, this bimodal response of barnacles to SIPC suggests that barnacles will settle gregariously when conspecific cues are sparse but will not settle if conspecific cues inform of overcrowding that will increase reproductive competition and diminish their reproductive chances.

KEY WORDS: Barnacles, *Amphibalanus amphitrite*, Settlement-inducing protein complex, Gregarious settlement, Predation, Biofouling

INTRODUCTION

The reproductive behaviour of barnacles has always fascinated scientists, among them Charles Darwin (Darwin, 1854), partly because their sessile trait makes them adopt reproductive strategies that rely on proximity to conspecifics (Barnes and Barnes, 1977; Crisp and Meadows, 1962; Crisp, 1961), but also because these animals are model organisms for marine biofouling (Aldred and Clare, 2008; Rittschof and Cohen, 2004; Rittschof et al., 1984, 1992). The life cycle of these animals consists of the nauplius (larval) stage, the cyprid stage, in which the barnacle actively seeks a preferred substrate for settlement, after which it metamorphoses into a juvenile barnacle, and then the adult stage, in which it is capable of sexual reproduction (Høeg and Møller, 2006). Early observations on the surface exploration behaviour of barnacles (Yule and Walker, 1985; Yule and Crisp, 1983; Crisp and Meadows, 1962; Crisp, 1961) led to the discovery of the proteinaceous nature of the signals that these animals use to inform their conspecifics of their presence and thus increase their reproductive chances (Matsumura et al., 1998b; Dreanno et al., 2006c). Several earlier reports (see Rittschof and Cohen, 2004; Rittschof, 1990) and further research have shown that this reproductive strategy comes with a cost, because the same chemical cue acts as a potent feeding stimulant for a barnacle predator, the whelk *Acanthinucella spirata* (Zimmer et al., 2016).

Intense research efforts culminated in the discovery of settlement-inducing protein complex (SIPC), a glycoprotein that induces conspecific larval settlement in *Amphibalanus amphitrite* (Darwin 1854) (Matsumura et al., 1998b; Dreanno et al., 2006a,b,c). Originally shown to comprise 3 subunits of 98, 88 and 76 kDa (Matsumura et al., 1998b), cloning of the cDNA that encodes for SIPC from *A. amphitrite* revealed that SIPC is a single protein of 1547 amino acids with a calculated molecular mass of 171.7 kDa (Dreanno et al., 2006c). It is firmly established that SIPC is an N-linked glycan-containing protein and that glycans, such as mannose, can induce larval settlement behaviour (Pagett et al., 2012; Dreanno et al., 2006b; Thiyagarajan et al., 2002). Research on another barnacle species, *Balanus glandula*, has identified another orthologous protein termed MULTIFUNCin (Ferrier et al., 2016), which shares 78% nucleotide sequence identity with SIPC and is expressed as three subunits of 98, 88 and 76 kDa in addition to the 199 kDa of the full-length protein (Ferrier et al., 2016). Positive-settlement-directed purification of MULTIFUNCin and its subunits indicated that only the 199 kDa and the 98 kDa proteins were inducing settlement of the cyprids of this species (Ferrier et al., 2016). These authors reported failure to express MULTIFUNCin in *Escherichia coli* and *Pichia pastoris* heterologous systems (Ferrier et al., 2016). Another study, however, reported the expression of a recombinant SIPC from *A. amphitrite* in a baculovirus expression system (Zhang et al., 2016). This recombinant SIPC from *A. amphitrite* was lacking its N-terminal signal peptide sequence, had an apparent molecular mass of around 200 kDa and did not induce cyprid settlement (Zhang et al., 2016). This inability of heterologously expressed *A. amphitrite* SIPC to induce settlement of the cyprids is in disagreement with results from another study, which showed that native SIPC, purified from *A. amphitrite* homogenates, induces cyprid settlement with a half maximal effective concentration (EC_{50}) of 102 nmol l^{-1} (Pagett et al., 2012).

Throughout the extensive literature pertaining to the identification of SIPC or other orthologous proteins (Dreanno et al., 2007; Ferrier et al., 2016; Matsumura et al., 1998b), emphasis

[1]Department of Biology, National and Kapodistrian University of Athens, Athens 157 84, Greece. [2]Institute of Biosciences and Applications, National Center for Scientific Research 'Demokritos', Agia Paraskevi, Athens 15310, Greece. [3]Biomedical Sciences Research Center 'Alexander Fleming', Fleming 34, 16672 Vari, Greece.

*Author for correspondence (sdedos@biol.uoa.gr)

M.K., 0000-0001-6949-8490; M.P., 0000-0002-9030-1469; S.M., 0000-0002-0388-3369; D.V., 0000-0003-2695-0795; I.K., 0000-0003-3232-0167; S.G.D., 0000-0002-0432-338X

Received 23 May 2018; Accepted 1 October 2018

Journal of Experimental Biology (2018) 221, jeb185348. doi:10.1242/jeb.185348

is given to positive-settlement-directed purification, i.e. to identifying and reporting only subunits that exhibit a settlement-inducing response; therefore, it is critical to note that such positive-settlement assay designs may not allow for the identification of any negative settlement cues. Therefore, comparing the results from *B. glandula* MULTIFUNCin (Ferrier et al., 2016) and the *A. amphitrite* SIPC (Pagett et al., 2012), there is ambiguity as to which of the SIPC subunits or the full-length protein exert the settlement behaviour. Is it the N-terminal 98 kDa subunit (Pagett et al., 2012; Ferrier et al., 2016; Matsumura et al., 1998b; Dreanno et al., 2006c) that exerts the settlement cues? Are the post-translational modifications (PTMs) of SIPC, which are critical for its activity as a pheromone cue (Aldred and Clare, 2008; Matsumura et al., 1998b; Zhang et al., 2016; Dreanno et al., 2007, 2006a), solely responsible for its biological activity or is the reported EC_{50} of 102 nmol l^{-1} (Pagett et al., 2012) of native SIPC a composite effect brought about by the simultaneous presence of all three subunits (Matsumura et al., 1998b)?

Attempting to answer these confounding questions, in this study we engineered, expressed and purified a recombinant form of SIPC from *A. amphitrite* as well as a systematic series of SIPC truncated fragments to identify whether SIPC, its putative subunits (Matsumura et al., 1998b; Dreanno et al., 2006a) or its PTMs (Pagett et al., 2012) are responsible for its biological activity. We engineered a recombinant form of SIPC and its truncated fragments to contain an N-terminal myc-tag and a C-terminal 6×His-tag to be able to detect and purify the full-length protein or its fragments from heterologous expression systems. By developing a systematic series of truncated SIPC fragments that were overexpressed and purified from heterologous expression systems, we were aiming to identify, in settlement bioassays of *A. amphitrite* cyprids (Rittschof et al., 1984, 1992), the minimal domain(s) that confer the species-specific biological activity of SIPC, be it protein domain(s) or domain-specific glycosylations (Pagett et al., 2012; Dreanno et al., 2007).

We show here that SIPC is a single glycoprotein, with a molecular mass substantially larger than its calculated molecular mass, that can transduce both settlement preference and settlement avoidance cues and thus guide a cyprid's decision to settle on a substrate or not. We also show that domains in the N-terminal 600 amino acids of SIPC, possibly through their extensive glycosylations, act as settlement preference cues. When these 600 amino acids are sequentially removed in recombinant truncated fragments of SIPC, the remaining domains of SIPC dose-dependently transduce only settlement avoidance cues. Although puzzling at a first impression, the settlement avoidance cues conferred by high concentrations of the full-length recombinant SIPC may serve to reduce reproductive competition (Crisp, 1961) by conspecifics or may actually be a covert predator avoidance strategy (Zimmer et al., 2016).

MATERIALS AND METHODS

Animals

Adult *A. amphitrite* (Cirripedia, Balanidae) were collected from boat docks at Port Mikrolimano and Floisvos, Athens, Greece. Animals were cleaned of epibionts with a small hard brush and meticulously identified as individuals of the species (Bernard and Lane, 1962; Costlow, 1956). Adult barnacles were kept in separate aerated glass tanks (20 l) containing 200-µm-filtered natural seawater at 27°C and a 12 h:12 h light:dark photoperiod. Tanks were fed 24-h-hatched *Artemia* sp. (class: Branchiopoda) nauplii and *Tetraselmis suecica* (class: Chlorodendrophyceae) and *Skeletonema costatum* (class: Bacillariophyceae) algae each day, and seawater was changed on alternate days. Upon stress to induce oviposition (24 h exposure to air or immersing in fresh water for 5 h), adults were returned to seawater for larval release. Hatched nauplii were attracted to a point light source, collected and placed into a beaker containing 2 or 3 l of 0.7 µm GF/F (Whatman)-filtered natural seawater at a density of approximately 1–2 nauplii ml^{-1} with gentle aeration. Nauplii were maintained at 27°C at a 12 h:12 h light:dark photoperiod on a diet of *Chaetoceros gracilis* (class: Bacillariophyceae) provided at a density of 2×10^5 cells ml^{-1} (Matsumura et al., 1998a; Thiyagarajan et al., 2002). Cultured in these conditions, nauplii metamorphosed to cyprids in 5 days. Aliquots of these cyprids were collected with a wide-mouthed Pasteur pipette and aged at 4°C (Phang et al., 2009) for 1 day prior to use in settlement assays. Only batches of cyprids that were active and had numerous oil cells, representing energy reserves (Lucas et al., 1979), were used in settlement assays (Rittschof et al., 1984, 1992).

Construction of recombinant SIPC and truncated SIPC fragments

We obtained a plasmid containing the full-length cDNA of the *A. amphitrite SIPC* gene cloned in pCR4-TOPO (Dreanno et al., 2006c) (Thermo Fisher Scientific). The full-length cDNA was sequenced and then transferred as an *Eco*RI (NEB) fragment into pENTR1α (Thermo Fisher Scientific). The sequencing results revealed differences between the obtained clone and the original cDNA deposited in GenBank (acc. no.: AAR33079.1; see Tables S1–S3). Differences were also observed between this clone and the nucleotide and amino acid sequence of SIPC deposited by Zhang et al. (2016) (GenBank acc. no.: AMR58954.1; see Tables S1–S3), which notably lacks the signal peptide sequence entirely (Zhang et al., 2016). Sequence alignment of the cDNA we obtained and the two other cDNA sequences deposited in GenBank (Zhang et al., 2016; Dreanno et al., 2006c) showed that there are differences between the three protein sequences (Tables S1–S3). Bioinformatic analysis (see below) of the deduced amino acid sequence, encoded by the cDNA we obtained, revealed that the amino acid differences that we observed were not affecting any putative PTM site as judged by its primary structure.

We proceeded to correct the three amino acid differences (Tables S2 and S3) that we observed in the signal peptide sequence using a forward primer (primer 1; Table S4) and a reverse primer (primer 2; Table S4) designed to remove the 3′-untranslated region (UTR) and thus obtain the open reading frame (ORF) of the gene. The PCR reaction conditions were: 95°C for 3 min as an initial step followed by 35 cycles of 95°C for 30 s, 46°C for 30 s and 68°C for 5 min using *Pfx* polymerase (Thermo Fisher Scientific) according to the manufacturer's instructions. The resulting PCR product of ∼4730 bp was cloned as an *Eco*RI fragment in pENTR1α (Thermo Fisher Scientific) to generate clone pENTR1α–SIPC–ORF.

To introduce a myc-tag (N–EQKLISEEDL–C) sequence immediately after the cleavage site of the signal peptide, we used primers 3 and 4 (see Table S4). Clone pENTR1α–SIPC–ORF was used as template and PCR conditions were: 95°C for 3 min as an initial step followed by 35 cycles of 95°C for 30 s, 46°C for 30 s and 68°C for 5 min using *Pfx* polymerase according to the manufacturer's instructions. Primers 3 and 4 were added at a final concentration of 125 nmol l^{-1} and primers 1 and 2 were added at a final concentration of 300 nmol l^{-1} to the PCR reaction. The resulting PCR product of ∼4760 bp was cloned as an *Eco*RI fragment in pENTR1α to create pENTR1α–SIPC–Myc–ORF.

Next, we engineered a C-terminal His-tag on recombinant SIPC by using a reverse primer (primer 5; Table S4) that introduced a

6×His tag just before the stop codon, and a forward primer (primer 6; Table S4). The PCR conditions were identical to those described above but with shorter extension at 68°C for 1 min. The resulting PCR product of 950 bp was digested with *Xba*I (NEB) and cloned into pENTR1α–SIPC–Myc–ORF to create pENTR1α–SIPC–Myc–ORF–6×His.

To this clone, which coded for an N-terminal myc-tagged and C-terminal His-tagged SIPC, we introduced a *Sph*I restriction site just after the myc-tag using primers 7 and 8 (Table S4) and primers 1 and 5. The PCR conditions were: 95°C for 3 min as an initial step followed by 35 cycles of 95°C for 30 s, 46°C for 30 s and 68°C for 5 min using *Pfx* polymerase according to the manufacturer's instructions. Primers 7 and 8 were added at a final concentration of 125 nmol l^{-1} and primers 1 and 5 were added at a final concentration of 300 nmol l^{-1} to the PCR reaction. The resulting PCR product of ∼4760 bp was digested with *Eco*RI–*Xho*I (NEB) and the resulting fragment of ∼2825 bp replaced a fragment of similar size in the previous clone to create pENTR1α–SIPC–Myc–SphI–ORF-6×His.

To generate the Venus variant of EGFP (Nagai et al., 2002), the pEGFP-N1 plasmid (Clontech) was used as template and the ORF of EGFP was amplified by PCR using primers 9 and 10 (Table S4) and cloned as an *Eco*RI fragment in pENTR1α. This clone was used as a template in PCR reactions that progressively mutated the ORF of EGFP to its Venus variant using primers 11–22 (Table S4). The PCR conditions were: 95°C for 3 min as an initial step followed by 35 cycles of 95°C for 30 s, 46°C for 30 s and 68°C for 1 min using *Pfx* polymerase according to the manufacturer's instructions. Primers for mutagenesis (primers 11–22) were added at a final concentration of 125 nmol l^{-1} and primers 9 and 10 were added at a final concentration of 300 nmol l^{-1} to the PCR reaction. Mutations introduced to EGFP to change it to its Venus variant were: F47L, T66G, V69L, Q70M, S73A, M154T, V164A, S176G and T204Y (Nagai et al., 2002).

Upon completion of the introduced mutations, the resulting ORF of Venus was subcloned as a *Sph*I (NEB) fragment into pENTR1α–SIPC–Myc–SphI–ORF–6×His to create pENTR1α–SIPC–Myc–SphI–Venus–ORF–6×His.

To express the neurosecretory-vesicle-localised PTPRN2 (Caromile et al., 2010), first-strand cDNA was synthesised from 5 µg total RNA from rat cerebellum (Takara) with 200 U Superscript®III reverse transcriptase (Thermo Fisher Scientific) in 20 µl reaction volume using an oligo$(dT)_{20}$ primer (Thermo Fisher Scientific) plus random hexamers (Thermo Fisher Scientific) (50 ng per 5 µg total RNA) according to the manufacturer's instructions. The resulting cDNA was diluted with nuclease-free water (Thermo Fisher Scientific) before use in PCR. Using the NCBI reference sequence NM_031600.1 (Wasmeier and Hutton, 1996) as a guide, first a 3.02 kb PCR product, which contained the ORF of *PTPRN2*, was amplified from the cDNA using primers 23 and 24 (Table S4). The PCR conditions were: 95°C for 3 min as an initial step followed by 35 cycles of 95°C for 30 s, 46°C for 30 s and 68°C for 4 min using *Pfx* polymerase according to the manufacturer's instructions, and the PCR product was digested with *Eco*RI–*Age*I (NEB). Next, the ORF of mTurquoise from mTurquoise2-ER (Goedhart et al., 2012) was amplified by PCR using primers 25 and 26 (Table S4). The PCR conditions were: 95°C for 3 min as an initial step followed by 35 cycles of 95°C for 30 s, 46°C for 30 s and 68°C for 1 min using *Pfx* polymerase according to the manufacturer's instructions. The PCR product was cloned as an *Eco*RI–*Xho*I fragment in pENTR1α and then the resulting clone was digested with *Eco*RI–*Age*I and the 3.02 kbp PCR product of the *PTPRN2* ORF was cloned in frame to generate pENTR1α–PTPRN2–mTurquoise.

To replace the signal peptide sequence of SIPC, we used primer 27 (Table S4) and primer 5 and pENTR1α–SIPC–Myc–SphI–ORF–6×His as template. The PCR conditions were: 95°C for 3 min as an initial step followed by 35 cycles of 95°C for 30 s, 46°C for 30 s and 68°C for 5 min using *Pfx* polymerase according to the manufacturer's instructions. The resulting PCR product of ∼4760 bp was digested with *Eco*RI–*Xho*I and a fragment of ∼2825 bp replaced a fragment of similar size in pENTR1α–SIPC–Myc–SphI–ORF–6×His to create pENTR1α–SpSIPC–Myc–SphI–ORF–6×His. To this clone, the ORF of Venus was inserted in frame at the *Sph*I site to create pENTR1α–SpSIPC–Myc–SphI–Venus–ORF–6×His.

To generate truncated fragments of SIPC, we used a structural model-based approach (see below) in conjunction with the bioinformatic analysis of the primary structure of SIPC (see below). Using clone pENTR1α–SIPC–Myc–SphI-ORF–6×His as template, we generated first 2 truncated SIPC fragments (Tr.2 and Tr.7) using primer 28 and 29 (Table S4), respectively, as forward primers and primer 5 as reverse primer. The PCR conditions were: 95°C for 3 min as an initial step followed by 35 cycles of 95°C for 30 s, 46°C for 30 s and 68°C for 4 min for Tr.2 and 2 min for Tr.7, using *Pfx* polymerase according to the manufacturer's instructions. The PCR products were cloned as *Sph*I–*Xho*I fragments in pENTR1α–SIPC–Myc–SphI–ORF–6×His to generate truncated SIPC fragments Tr.2 and Tr.7. Subsequently, and based on the results we obtained from these truncated fragments, we generated truncated fragments Tr.1 and Tr.3 using primers 30 and 31 (Table S4) in an identical approach as above. Next, we generated truncated fragments Tr.4, Tr.5 and Tr.6 using primers 32, 33 and 34 (Table S4) in an identical approach as above by simply adjusting the PCR extension time at 68°C to 3 min. Finally, we generated truncated fragments Tr.8 and Tr.9 using primers 35 and 36 (Table S4) as forward primers and primer 37 (Table S4) as reverse primer in an identical approach as above by simply adjusting the PCR extension time at 68°C to 2 min. These last two PCR products were digested with *Sph*I–*Spe*I (NEB) and cloned into pENTR1α–SIPC–Myc–SphI–ORF–6×His to generate truncated fragments Tr.8 and Tr.9.

For expression in HEK293 cells, all constructs described above were transferred by recombination to the mammalian expression vector pcDNA3.2/V5-DEST (Thermo Fisher Scientific) using the Gateway LR Clonase II enzyme mix (Thermo Fisher Scientific) according to the manufacturer's instructions.

For baculovirus-mediated expression of SIPC in insect Sf9 cells, constructs (1) pENTR1α–SIPC–Myc–SphI–ORF–6×His (hereafter termed SIPC), (2) pENTR1α–SpSIPC–Myc–SphI–ORF–6×His (hereafter termed SpSIPC), (3) pENTR1α–SIPC–Myc–SphI–Venus–ORF–6×His (hereafter termed VSIPC) and (4) pENTR1α–SpSIPC–Myc–SphI–Venus–ORF–6×His (hereafter termed SpVSIPC) were recombined with baculovirus linear DNA using the BaculoDirect™ C-Term Expression Kit (Thermo Fisher Scientific) and the Gateway LR Clonase II enzyme mix (Thermo Fisher Scientific) according to the manufacturer's instructions.

All constructs were sequenced using the described primers and additional primers 38–46 (Table S4) before transfer to expression vectors. Notably, we observed no additional change or mutation in the nucleotide sequence of *SIPC* during the construct engineering process, an indication that the *SIPC* cDNA does not exhibit any inherent instability during its propagation in *E. coli* that would justify the presence of the observed nucleotide changes.

Journal of Experimental Biology (2018) 221, jeb185348. doi:10.1242/jeb.185348

Expression of recombinant SIPC in mammalian and baculovirus expression systems

HEK293 cells were seeded in culture flasks and transfected with expression vectors coding for SIPC, VSIPC and the various truncated SIPC fragments when in density of 1×10^6 cells ml^{-1} with 10 µg of DNA per 1×10^6 cells ml^{-1} of culture medium, using Escort IV transfection reagent (Merck) at a 1:1 ratio of µg DNA:µl of transfection reagent, according to the manufacturer's instructions. Medium and cells were separately harvested at 72 h post-transfection. Medium was initially centrifuged at 90 ***g*** for 10 min at 4°C to remove cells and then at 10,000 ***g*** for 10 min to remove cellular debris, and then carefully aspirated and stored at −80°C until use. Cells were removed by scraping in PBS and centrifuged twice in the presence of PBS at 90 ***g*** for 10 min at 4°C to remove any remaining medium, then resuspended in lysis buffer (50 mmol l^{-1} Tris pH 7.6, 150 mmol l^{-1} NaCl, 2 mmol l^{-1} EDTA, 1% Triton X-100) and a protease inhibitor cocktail tablet/10 ml lysis buffer (Roche Diagnostics) and stored at −80°C until use.

Recombinant baculoviruses expressing the various forms of full-length recombinant SIPC were propagated in Sf9 cells (Thermo Fisher Scientific) and, after 3 re-infections of Sf9 cells, the viral titre was determined using the end-point dilution assay. For large-scale protein expression, Sf9 cells were cultured in Sf-900™ III SFM medium and infected at a multiplicity of infection of 10. For each batch, a total of 10^8 cells were infected. Medium and cells were harvested at 72 h post-infection, and samples were processed separately as described above and stored at −80°C until use.

Protein concentration in samples was determined with the *RC DC*™ Protein Assay kit II (Bio-Rad) using bovine serum albumin (BSA) as standard. Protein samples were pre-incubated with 250 mmol l^{-1} dithiothreitol (DTT) before resolution in one-dimensional (1D) SDS-PAGE gels of 5 or 8% acrylamide/bis-acrylamide at 10 µg protein/lane for cell samples and 20 µl/lane for culture medium samples.

For western blots, gels were wet transferred onto a nitrocellulose membrane (Whatman) in transfer buffer [25 mmol l^{-1} Tris, 192 mmol l^{-1} glycine, 20% (v/v) methanol] at 100 V constant at 4°C for 90 min. Nitrocellulose membranes were blocked by incubation with 5% BSA in TBST (25 mmol l^{-1} Tris-HCl, pH 7.5, 150 mmol l^{-1} NaCl, 0.1% Tween 20) for 1 h at room temperature. Membranes were then incubated for 1 h at room temperature with a Myc-Tag (9B11) mAb (Cell Signaling Technology) at 1:1000 dilution in TBST with 5% BSA, washed with TBST (3×5 min) and incubated with an HRP-conjugated anti-mouse antibody (1:1000, Cell Signaling Technology) for 1 h at room temperature. Immunoreactive bands were detected using the 20× LumiGLO® chemiluminescence kit (Cell Signaling Technology) in an Alpha Innotech FluorChem 8800 imaging system.

Fluorescence confocal microscopy

HEK293 cells were seeded on poly-D-lysine (Merck)-coated coverslips in 6-well plates at a density of 0.5×10^6 cells/well and transiently transfected with expression vectors for VSIPC and PTPRN2–mTurquoise as described above. At 72 h post-transfection, cells were washed 3 times with PBS (pH 7.4), fixed in 1 ml of 4% PFA for 15 min at room temperature followed by 3 washes with PBS, and sealed with 10 µl Duolink® *In Situ* Mounting Medium with DAPI (Merck). Cells were imaged on a multiphoton confocal microscope Leica TCS SP8 MP equipped with an Argon laser (excitation lines at 458, 476, 488, 496 and 514 nm) and an IR MaiTai DeepSee Ti:Sapphire laser (Spectra-Physics) for multiphoton applications. Images were acquired with the spectral detector of the microscope using appropriate emission wavelength ranges: PTPRN2–mTurquoise was excited at 458 nm with the Argon laser and emission was recorded between 465 and 500 nm, VSIPC was excited at 514 nm with the Argon laser and emission recorded at 520–600 nm. DAPI was excited at 780 nm with the MaiTai DeepSee laser and emission captured at 400–450 nm. Images were acquired with the LAS X software (Leica Microsystems CMS GmbH). For imaging, suitable doubly transfected cells were selected, and stacks of 20–30 optical sections were recorded for each cell. Images are presented without any post-processing.

Purification of recombinant SIPC from insect cell culture medium

The 6×His-tagged recombinant SIPC (SpSIPC) recovered from the insect cell culture medium at 72 h post-infection was purified using Ni Sepharose 6 Fast Flow (GE Healthcare Life Science) and the following protocol: Ni beads prepared to 50% slurry with binding buffer (20 mmol l^{-1} sodium phosphate, 0.5 mol l^{-1} NaCl, 20 mmol l^{-1} imidazole, pH 7.4) were added at 1.5 ml per 30 ml of culture medium and incubated at 4°C on a rotating shaker for 1 h. Beads were then washed (2×1 ml) with binding buffer and bound protein(s) were eluted from the beads with 4×0.5 ml elution buffer (20 mmol l^{-1} sodium phosphate, 0.5 mol l^{-1} NaCl, 500 mmol l^{-1} imidazole, pH 7.4). Samples of the eluted fractions were analysed by western blots and then fractions from several batches (n=6) were pooled and dialysed against a HEPES buffer (20 mmol l^{-1} HEPES, 200 mmol l^{-1} NaCl, pH 7.4) using a Spectra/Por™ cellulose ester dialysis membrane (molecular mass cut-off 100 kDa; Thermo Fisher Scientific). Samples were concentrated to 100 µl using the Amicon Ultra-15 centrifugal filter units (Merck) at 4000 ***g*** for 15 min at 4°C. A fraction of these samples was reserved for liquid chromatography tandem-mass spectrometry (see below) and the remaining was dialysed against sterile artificial seawater (ASW) (25‰) overnight in 4°C using Slide-A-Lyzer™ dialysis cassettes (molecular mass cut-off 100 kDa; Thermo Fisher Scientific). Protein concentration was again determined using the *RC DC*™ Protein Assay kit II (Bio-Rad) and BSA as standard, and samples saved for use in settlement assays.

Purification of recombinant SIPC from mammalian cell culture medium

The 6×His-tagged recombinant SIPC (rSIPC) and its truncated fragments recovered from HEK293 cell culture medium were purified using Ni Sepharose 6 Fast Flow (GE Healthcare Life Science) and the following protocol: Ni beads prepared to 50% slurry with binding buffer (20 mmol l^{-1} sodium phosphate, 0.5 mol l^{-1} NaCl, 20 mmol l^{-1} imidazole, pH 7.4) were added at 1.5 ml per 30 ml of culture medium and incubated at 4°C on a rotating shaker for 1 h. Beads were then washed (2×1 ml) with binding buffer and bound protein(s) were eluted from the beads with 4×0.5 ml elution buffer (20 mmol l^{-1} sodium phosphate, 0.5 mol l^{-1} NaCl, 500 mmol l^{-1} imidazole, pH 7.4). Samples of the eluted fractions were analysed by western blot and then fractions from several batches (n=4–7) were pooled and dialysed against ddH_2O for 2 h, then sterile sea water (36‰) for 2 h and finally sterile ASW (25‰) overnight in 4°C using Spectra/Por™ cellulose ester dialysis membrane (molecular mass cut-off 100 kDa; Thermo Fisher Scientific). For truncated fragments of SIPC smaller than 100 kDa, a dialysis membrane with 14 kDa cut-off was used (Merck). Protein concentration was determined using the

RC DC™ Protein Assay kit II (Bio-Rad) and BSA as standard, and samples saved for use in settlement assays.

Analysis of SIPC expression at *A. amphitrite* developmental stages

To identify whether the recombinant SIPC that we engineered has similar mobility in SDS-PAGE gels as the endogenous SIPC of *A. amphitrite*, we generated a rabbit polyclonal antibody against a peptide (N–IHKELKGGTERGGE–C) corresponding to amino acids 1186–1199 of SIPC (GenBank acc. no.: AAR33079.1). With a cysteine attached to its N-terminus, the 15 amino acid peptide was conjugated to keyhole limpet hecocyanin (KLH) and the conjugate was used to immunise rabbits. The antibody was produced and purified by GenScript (NJ, USA).

Monoclonal antibodies against the same peptide were produced according to a modified method of Köhler and Milstein (1975). Five BALB/c mice of 5 weeks of age were immunised intraperitoneally (i.p.) with 25 µg of the KLH-conjugated peptide (peptide synthesis and conjugation by GenScript). All immunisation and animal handling were in accordance with animal care guidelines as specified in EU Directive 2010/63/EU. After 5 cycles of immunisation, mice were euthanised and spleenocytes were collected and fused with the P3X63Ag8.653 cell line (ATCC® CRL1580™) according to the Köhler and Milstein (1975) fusion protocol. Positive clones and antibody specificity were determined through immunoblotting and immunosorbent assays and, among the several positive clones, one was further propagated and used.

Nauplii, cyprids and freshly settled adults were cultured as described above. Nauplii were collected with a wide-mouthed Pasteur pipette, whereas early attached cyprids of *A. amphitrite* (24 h old) and metamorphosed juveniles (48 h old) were gently scraped off the dishes. Animals were anaesthetised in 0.5 mol l^{-1} $MgCl_2$ and this solution was quickly replaced with 1 µl/animal lysis buffer (50 mmol l^{-1} Tris, pH 7.6, 150 mmol l^{-1} NaCl, 2 mmol l^{-1} EDTA, 1% Triton X-100) and a protease inhibitor cocktail tablet/10 ml lysis buffer (Roche Diagnostics) and stored at −80°C until use. After homogenisation with a small pestle, samples were sonicated 3 times (30 s with 2.5 min rest on ice) using an MSE Soniprep 150 ultrasonic disintegrator (Sanyo). Samples were centrifuged at 15,000 ***g*** for 10 min at 4°C and the supernatant was incubated with 250 mmol l^{-1} DTT and resolved in a 5% acrylamide/bis-acrylamide SDS-PAGE gel.

Gels were wet transferred onto a nitrocellulose membrane in transfer buffer [25 mmol l^{-1} Tris, 192 mmol l^{-1} glycine, 20% (v/v) methanol] at 100 V constant at 4°C for 90 min. Nitrocellulose membranes were blocked by incubation with 5% BSA in TBST (25 mmol l^{-1} Tris-HCl, pH 7.5, 150 mmol l^{-1} NaCl, 0.1% Tween 20) for 1 h at room temperature. Membranes were then incubated for 1 h at room temperature with the purified polyclonal anti-SIPC rabbit antibody at 1:1000 dilution in TBST with 5% BSA or with the mouse anti-SIPC monoclonal antibody at 1:100 dilution of the cloned cell line culture supernatant in TBST with 5% BSA. Membranes were then washed with TBST (3×5 min) and incubated with an HRP-conjugated anti-rabbit or HRP-conjugated anti-mouse antibody (1:1000, Jackson Laboratories) for 1 h at room temperature. Immunoreactive bands were detected using the 20× LumiGLO® chemiluminescence kit (Cell Signaling Technology) in an Alpha Innotech FluorChem 8800 imaging system.

Settlement bioassays

Assays were conducted by adding 12 cyprids into individual wells of a 24-well polystyrene sterile microplate (Orange Scientific) containing 2 ml of ASW (25‰) and various concentrations of recombinant SIPC. These experiments were repeated 3 times with 5 replicates for each concentration of recombinant SIPC (n=180 cyprids for each concentration of recombinant SIPC). An identical procedure was used for settlement assays in the presence of BSA (molecular biology grade from NEB).

Plates were covered and sealed with Parafilm® to avoid evaporation, incubated at 25°C away from any light source, and examined after 24, 48, 72 and 96 h. Each animal was inspected under a stereomicroscope and its condition recorded. On each day, cyprids that did not move, had extended thoracopods and did not respond after a light touch with a Pasteur pipette were regarded as dead (Lau and Qian, 2000; Rittschof et al., 1992; Hellio et al., 2004). Both permanently attached and metamorphosed individuals were counted as settled or juveniles (Hellio et al., 2005; Rittschof et al., 2003). The remainder were counted as 'no settlement'. Results were expressed as percentage of settlement, mortality and no settlement. For SIPC and its truncated fragments expressed in HEK293 cells, a similar approach was used but with fewer replicates (n=48–96 cyprids) and fewer concentrations of each protein per assay due to the substantially decreased yields of this expression system relative to the baculovirus expression system.

Adhesive properties of recombinant SIPC and truncated SIPC fragments

Adhesion assays of recombinant SIPC or its truncated fragments (Tr.1 to Tr.9) were conducted by adding 300 µl of each protein sample (50 nmol l^{-1}) into individual wells of a 24-well polystyrene sterile microplate (Orange Scientific). Plates were covered and sealed with Parafilm® to avoid evaporation and incubated at 25°C away from any light source for 24 h. Then, samples were aspirated and wells were blocked by incubation with 5% BSA in TBST (25 mmol l^{-1} Tris-HCl, pH 7.5, 150 mmol l^{-1} NaCl, 0.1% Tween 20) for 1 h at room temperature. Wells were then incubated for 1 h at room temperature with a Myc-Tag (9B11) mAb (Cell Signaling Technology) at 1:1000 dilution in TBST with 5% BSA, washed with TBST (3×5 min) and incubated with an HRP-conjugated anti-mouse antibody (1:1000, Cell Signaling Technology) for 1 h at room temperature. Wells were then washed with TBST (3×5 min) and immunostaining was detected using the 20× LumiGLO® chemiluminescence kit (Cell Signaling Technology) in a Bio-Rad ChemiDoc XRS+ Gel Photo Documentation system. Acquired high-resolution images were analysed with ImageJ 1.48v (National Institutes of Health) and data were exported and statistically analysed with GraphPad Prism v.6. These experiments were repeated 3 times with 2 replicates for each sample.

In experiments where the adhesive properties of native SIPC was assessed, settlement bioassays were conducted by adding 10 cyprids into individual wells of a 24-well polystyrene sterile microplate (Orange Scientific) containing 2 ml of ASW (25‰). These experiments were repeated 6 times with 3 replicates per experiment. Plates were covered and sealed with Parafilm® to avoid evaporation and incubated at 25°C away from any light source for 96 h. Then, the 2 ml of ASW was aspirated from each well and saved, and all settled and non-settled cyprids were counted and removed as well. Wells were then blocked by incubation with 5% BSA in TBST (25 mmol l^{-1} Tris-HCl, pH 7.5, 150 mmol l^{-1} NaCl, 0.1% Tween 20) for 1 h at room temperature. Wells were then incubated for 1 h at room temperature with the mouse anti-SIPC monoclonal antibody at 1:100 dilution of the cloned cell line culture supernatant in TBST with 5% BSA, washed with TBST (3×5 min) and incubated with an HRP-conjugated anti-mouse antibody

(1:1000, Cell Signaling Technology) for 1 h at room temperature. Wells were then washed with TBST (3×5 min) and immunostaining was detected using the 20× LumiGLO® chemiluminescence kit (Cell Signaling Technology) in a Bio-Rad ChemiDoc XRS+ Gel Photo Documentation system.

In parallel, the 2 ml of ASW sample from each well was subjected to trichloroacetic acid (TCA) precipitation to identify native SIPC, according to the following protocol. A 0.5 ml volume of 100% (w/v) TCA was added to 2 ml of each ASW sample and incubated at 4°C for 10 min. Following centrifugation of each sample at 20,800 ***g*** for 5 min at 4°C, supernatant was removed and 200 µl ice-cold acetone was added to the pellet. After centrifugation of each sample at 20,800 ***g*** for 5 min at 4°C, 200 µl ice-cold acetone was added to the pellet and the process was repeated one more time. After evaporation of acetone, 20 µl of 1× SDS-PAGE sample buffer was added to each pellet, and samples were boiled at 95°C for 5 min and subjected to SDS-PAGE electrophoresis in an 8% acrylamide/bis-acrylamide gel. Western blots, with the mouse anti-SIPC monoclonal antibody at 1:100 dilution of the cloned cell line culture supernatant were carried out as described above.

Bacteria counting assays

Assays was conducted by adding 12 cyprids into individual wells of a 24-well polystyrene sterile microplate (Orange Scientific) containing 2 ml of ASW (25‰) and various concentrations of recombinant SIPC (SpSIPC; 10 nmol l^{-1}, 50 nmol l^{-1}, 150 nmol l^{-1}) with 3 replicates per experiment. Plates were covered and sealed with Parafilm® to avoid evaporation and incubated for 24 h away from any light source at 25°C. Then, samples from each well were taken to assess the number of existing heterotrophic marine bacteria using the spread plate counting technique (Zobell, 1941). Zobell's marine-agar-containing plates were inoculated with 10 µl of each ASW sample or 10 µl of its 10^{-1}, 10^{-2} and 10^{-3} serial dilutions using a sterile glass rod. Colonies that developed on the marine agar plates were counted after incubation for 24 h at 25°C and the counts were expressed as colony forming units ml^{-1}.

Bioinformatic screening for PTMs of SIPC

The full-length recombinant SIPC (1568 amino acids) was screened for PTMs using relevant tools provided in the dbPTM server (http://dbptm.mbc.nctu.edu.tw/#; Huang et al., 2016) and ModPred (http://www.modpred.org; Pejaver et al., 2014). For tools available at the dbPTM server, default parameters were used. The results of bioinformatic annotation of PTMs of SIPC are shown in Table S5.

Crystal structure simulation of SIPC

To design the truncated fragments of recombinant SIPC, we used the i-TASSER algorithm at http://zhanglab.ccmb.med.umich.edu/I-TASSER/ (Zhang, 2008), in conjunction with the bioinformatic screening of recombinant SIPC for PTMs. Using the PDB ID:2PN5 crystal structure of thioester-containing protein isoform TEP1 of *Anopheles gambiae* (Baxter et al., 2007), which has 26% identity to SIPC (Pagett et al., 2012), as template, the best scoring model had a C-score=−1.51. This value measures the relative clustering structural density and consensus significance score of multiple threading templates (Zhang, 2008). The estimated accuracy of the model was 0.53±0.15 (TM-score or structural similarity measurement) with a root mean square deviation (RMSD) of 13.7±4.0 Å. The model was visualised, annotated and coloured with PyMOL v.1.7.4.4 to determine the sites of truncation on recombinant SIPC that would eliminate distinct structural domains.

Deglycosylation assays

Full-length SIPC and its fragments were deglycosylated using the Protein Deglycosylation Mix, which contains PNGase F, *O*-glycosidase, neuraminidase, β1–4 galactosidase and β-*N*-acetylglucosaminidase (NEB), in denaturing and non-denaturing reaction conditions according to the manufacturer's instructions. Assay performance was assessed by using bovine fetuin as a positive control. Deglycosylated samples were run on a 5 or 8% acrylamide/bis-acrylamide gel and western blots were carried out as described above. The migration of recombinant SIPC or its truncated fragments was analysed using ImageJ software, and calibrated with the standard molecular mass of the protein ladder from several western blots (n=3–7).

Proteomic analysis of recombinant SIPC

A purified 20 µl sample of SpSIPC expressed in Sf9 cells was subjected to SDS-PAGE electrophoresis in a 5% acrylamide/bis-acrylamide gel. The gel was then fixed for 30 min (45% methanol, 1% acetic acid) and subjected to silver staining using the Pierce™ Silver Stain for mass spectrometry (Thermo Fisher Scientific) according to the manufacturer's instructions. Upon band visualisation, a gel slice corresponding to the 217 kDa molecular mass of SpSIPC (Fig. S4F) was excised, cut into smaller pieces, transferred into a tube, washed, dehydrated with acetonitrile and subjected to in-gel enzymatic digestion with trypsin (Gold, mass spec grade, Promega) in 50 mmol l^{-1} ammonium bicarbonate solution overnight at 37°C. Peptides were eluted and dried to completeness upon speed-vac-assisted solvent removal. The dried peptide mixtures were reconstituted in 2% acetonitrile, 0.1% formic acid in water and analysed by LC-MS/MS. Peptides were pre-concentrated with a flow of 3 µl min^{-1} for 10 min on a C18 trap column (Acclaim PepMap100, 100 µm×2 cm, Thermo Fisher Scientific) and then loaded onto a 15 cm C18 column (75 µm ID, particle size 2 µm, 100 Å, Acclaim PepMap RSLC, Thermo Fisher Scientific). The binary pumps of the HPLC (RSLCnano, Thermo Fisher Scientific) consisted of solution A [2% (v/v) acetonitrile in 0.1% (v/v) formic acid] and solution B [80% (v/v) acetonitrile in 0.1% (v/v) formic acid]. The peptides were separated using a linear gradient of 4% solution B up to 40% in 50 min for a 1 h gradient run with a flow rate of 300 nl min^{-1}. The column was placed in an oven operating at 35°C. The eluted peptides were ionised by a nano-electrospray source and analysed by MS/MS on an LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific). Full scan MS spectra were acquired in the Orbitrap (m/z 300–1600) in profile mode and data-dependent acquisition, with the resolution set to 60,000 at m/z 400 and automatic gain control target at 10^6 ions. The 6 most intense ions were sequentially isolated for collision-induced (CID) MS/MS fragmentation and detection in the linear ion trap. Dynamic exclusion was set to 1 min and activated for 90 s. Ions with single charge states were excluded. Lock masses of m/z 445, 120 and 025 were used for internal calibration. Xcalibur (Thermo Fisher Scientific) was used to control the system and acquire the raw files. Peptides were identified using the Proteome Discoverer 1.4 software (Thermo Fisher Scientific). The Orbitrap raw data, with peak S/N threshold set to 1.5, were searched in Proteome Discoverer 1.4 using the SEQUEST HT against the SpSIPC FASTA sequence of SIPC as well as the UniProt entries of the human proteome (89,663 sequences), the UniProt entries of the *Drosophila melanogaster* proteome (18,540 sequences) and the UniProt entries of *Bombyx mori* proteins (19,062 sequences) with strict trypsin specificity and with a maximum of two missed cleavages and variable modifications of methionine oxidation, deamidation of glutamine

and asparagines residues and acetylation of the protein N-terminus. The identified peptides were filtered based on their Xcorr values versus peptide charge states (XCorr>2 for charge state +2 and XCorr>2.5 for charge state +3). The mass spectrometry proteomics data and search results have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaíno et al., 2016) partner repository with the dataset identifier PXD006858 and 10.6019/PXD006858.

Statistical analysis

All data were fitted to a four-parameter logistic curve according to the following model: $y=\text{Bottom}+(\text{Top}-\text{Bottom})/(1+10^{[(\text{LogEC50}-x)\times\text{HillSlope}]})$ in GraphPad Prism v.6. To measure the relative distance of immunoreactive bands in western blots, images were analysed with ImageJ 1.48v and data were exported and statistically analysed with GraphPad Prism v.6. All other statistical analyses were carried out with GraphPad Prism v.6.

RESULTS

Expression of recombinant SIPC

It was previously reported that SIPC contains a signal peptide that directs its secretion (Dreanno et al., 2006a,c), and that finding agrees with our results where, upon expression of recombinant SIPC in HEK293 cells, we identified a secreted protein in the medium (Fig. 1A). The same was true for the Venus-containing recombinant SIPC (VSIPC), which was found to be secreted from HEK293 cells upon its expression (Fig. S1A). We observed that the expressed protein in HEK293 cells had a molecular mass>200 kDa, much higher than the calculated molecular mass of 171.45 kDa of secreted recombinant SIPC. The Venus-containing VSIPC had a molecular mass>230 kDa (Fig. S1A), much higher than the calculated molecular mass of 199.16 kDa of the secreted recombinant VSIPC. Fluorescence confocal microscopy revealed that recombinant VSIPC, expressed in HEK293 cells, had an endoplasmic-reticulum-like localisation distinct from the secretory-vesicle-localised PTPRN2 (Caromile et al., 2010; Fig. 1B), an indication that SIPC is secreted through transport vesicles.

Assuming that the less complex glycosylations of insect cells (Shi and Jarvis, 2007) will result in a protein of reduced molecular mass, we expressed recombinant SIPC in Sf9 cells and observed that it had a similar molecular mass (>200 kDa) to that observed in HEK293 cells (Fig. 1A,C). In this case, however, the protein was not secreted from Sf9 cells (Fig. 1C). To obtain a secreted form of SIPC from Sf9

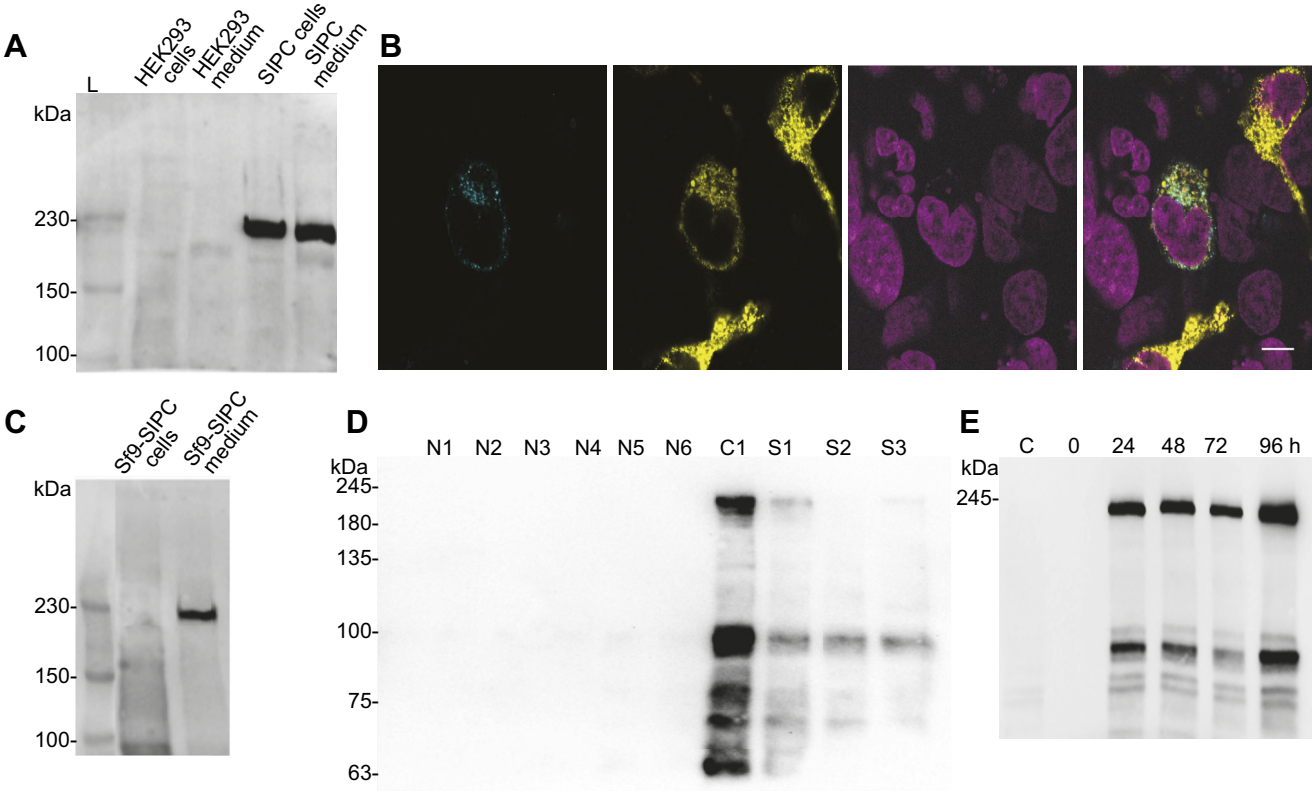

Fig. 1. Expression of recombinant SIPC from *Amphibalanus amphitrite*. (A) Recombinant settlement-inducing protein complex (SIPC) is secreted by HEK293 cells. Each lane contains 10 μg of total cellular protein or 20 μl of cell culture medium from untransfected cells (labelled as HEK293 cells and HEK293 medium) or cells transfected with the recombinant SIPC expression vector (labelled as SIPC cells and SIPC medium). Western blots were probed with an anti-myc antibody and a typical image of several blots (n=10) is shown. L, protein ladder. (B) Recombinant VSIPC (pENTR1α–SIPC–Myc–SphI–Venus–ORF–6×His) is not localised in secretory vesicles. Fluorescence confocal imaging of recombinant VSIPC in HEK293 cells co-transfected with PTPRN2–mTurquoise. Image 1 (far left): PTPRN2–mTurquoise localisation in secretory vesicles. Image 2: VSIPC localisation in endoplasmic-reticulum-like structures. Image 3: DAPI localisation in cell nuclei. Image 4 (far right): merge of images 1, 2 and 3 (scale bar: 10 μm). (C) The endogenous signal peptide of SIPC does produce a secreted recombinant SIPC in insect Sf9 cells. Gel loading and western blot conditions are identical to A. A typical image of several blots (n=3) is shown. (D) Detection of native SIPC in extracts from *A. amphitrite* at different developmental stages, with a mouse anti-SIPC monoclonal antibody. Each lane contains 10 μg of total cellular protein from different developmental stages of this species. A typical image of several blots (n=3) is shown. N1–N6, nauplii stage 1 to stage 6; C1, cyprid day 1; S1–S3, settled day 1 to day 3. (E) Time course of secretion of recombinant SpSIPC (pENTR1α–SpSIPC–Myc–SphI–ORF–6×His) from Sf9 cells after infection with recombinant SpSIPC-expressing baculoviruses. Each lane contains 20 μl of Sf9 cell culture medium. C, control Sf9 cell culture medium. Western blots were probed with an anti-myc antibody and a typical image of several blots (n=3) is shown.

Journal of Experimental Biology (2018) 221, jeb185348. doi:10.1242/jeb.185348

cells, we engineered a construct that contained the signal peptide of sarcotoxin 1A of the flesh fly *Sarcophaga peregrina* that has been shown to direct secretion of baculovirus-encoded proteins (O'Reilly et al., 1995). This change in the signal peptide resulted in the expression and secretion of recombinant SIPC from Sf9 cells (Fig. S1B). These cells can be cultured in a serum-free and protein-free medium and this was convenient for the purification of recombinant SIPC from the medium. Upon purification of SIPC from Sf9 cell culture medium, to eliminate any possibility that the immunoreactive protein was not SIPC, we probed the overexpressed proteins with a polyclonal antibody generated against a peptide sequence in the C-terminus of SIPC (Fig. S1C). This antibody recognised a single band in western blots from Sf9 cell culture medium at a molecular mass>200 kDa, only when a different signal peptide replaced the endogenous signal peptide of SIPC, in agreement with our previous results (Fig. 1A,C, Fig. S1A–C). As shown in the western blots in Fig. S1A,B, bands of smaller molecular mass, similar to those described before (Matsumura et al., 1998b; Dreanno et al., 2006a), were immunoreactive with the N-terminally located myc-epitope but not with the anti-SIPC antibody (Fig. S1C).

We tested this antibody against lysates from *A. amphitrite* nauplii, cyprids and adults of different developmental stages and observed that a protein of molecular mass much higher than the calculated molecular mass of endogenous SIPC immunoreacted with the anti-SIPC antibody (Fig. S1D). When probed in the same western blot, the antibody against SIPC recognised a band of identically high molecular mass (Fig. S1E). Because the western blot signals we were observing with the polyclonal antibody were generally weak (see Fig. S1C–E), we generated a mouse monoclonal antibody (mAb clone 1C4) against the same C-terminal epitope of SIPC and used it in western blot assays against lysates from *A. amphitrite* nauplii, cyprids and adults of different developmental stages (Fig. 1D). The results showed that this mAb reacted strongly with a band of ∼96 kDa and less with a band of ∼210 kDa, in addition to several other bands of smaller molecular masses (Fig. 1D). Immunoreactive bands were recorded exclusively at the cyprid stage and not in the nauplii stage (Fig. 1D). This observation effectively ended all our preliminary attempts to purify from cyprid homogenates and use the native SIPC in settlement assays because it showed that native SIPC exists in multiple subunits in cyprids as previously described (Matsumura et al., 1998b; Dreanno et al., 2006c; Ferrier et al., 2016). Realising that settlement assay results with the purified, native *A. amphitrite* SIPC can be potentially marred by the biological role of each of the 3 reported SIPC subunits (Matsumura et al., 1998b) versus the biological role of the full-length protein, we expressed the full-length SIPC in heterologous expression systems. To obtain large amounts of the protein for settlement assays with *A. amphitrite* cyprids, we expressed the two forms of the recombinant protein (SpSIPC and SpVSIPC) in Sf9 cells and observed that both were secreted from these cells (Fig. 1E, Fig. S1F–H). Once again, bands of smaller molecular masses, similar to those described before (Matsumura et al., 1998b; Dreanno et al., 2006a), were immunoreactive with the N-terminally located myc-epitope, in agreement with results from another study (Zhang et al., 2016).

Settlement assays with recombinant SIPC

We tested the biological activity of purified recombinant SIPC (SpSIPC) in settlement assays with cyprids of *A. amphitrite* (Fig. 2A) using an established protocol (Dreanno et al., 2007). The results show that *A. amphitrite* cyprids behaved in a bimodal fashion to the presence of SIPC, preferring to settle and metamorphose at lower concentrations of SIPC (EC_{50}=3.73 nmol l^{-1}) or preferring to avoid settlement at higher concentrations of SIPC (IC_{50}=101 nmol l^{-1}) (Fig. 2B–D). The settlement or avoidance response of the animals was evident within 24 h in the presence of SIPC (Fig. 2A), so the calculations of effective concentrations of SIPC were done with the 24 h data (Fig. 2C,D). Peak settlement preference responses were recorded at a concentration of 25 nmol l^{-1} SpSIPC (Fig. 2A) and settlement avoidance responses were evident from a concentration of 50 nmol l^{-1} SpSIPC onwards (Fig. 2A). Cyprids that did not metamorphose to juveniles were actively seeking a substrate to settle even after 96 h in the wells of the culture dishes (Fig. S2A) and there was no increase in mortality of the animals neither time-wise nor SIPC-concentration-wise (Fig. S2B). By plotting the percentage of cyprids that settled and metamorphose to juveniles (Fig. 2B, Fig. S2C), it is obvious that SIPC does not curtail, in any way, the developmental progression of barnacles. To rule out that the observed settlement avoidance by the animals at the highest range of SpSIPC concentrations (Fig. 2A,D) was not a potential artefact of our assay design, we carried out the same assays in the presence of BSA (Fig. S3A–D). The results showed that BSA at concentrations as high as 1 µmol l^{-1} had no settlement-inducing effect on cyprids and neither affected metamorphosis nor mortality (Fig. S3A–D). At a concentration of 10 µmol l^{-1}, BSA inhibited settlement of cyprids (Fig. S3A,B) with a calculated EC_{50} of 3.69 µmol l^{-1} at 24 h, a 36.5-fold higher concentration than the settlement avoidance response of the cyprids to SpSIPC. Because native SIPC has been shown to be a cyprid cuticle-bound protein (Zhang et al., 2016; Dreanno et al., 2006a), we tested whether recombinant SIPC is recognised by the cyprids as a surface-adhered moiety. When polystyrene plates were incubated with or without recombinant SIPC, only the wells that contained SpSIPC showed immunostaining (Fig. 2E). This immunostaining (Fig. 2E) was evident when cyprids were incubated for 96 h in polystyrene plates whereupon immunoreactive traces of native SIPC deposited by the cyprids could be detected on the surface of the plates (Fig. 2F). Western blots of the ASW where cyprids were incubated for 96 h showed that native SIPC is not a diffusible molecule (Fig. S4A). Furthermore, bacterial growth in the presence of recombinant SIPC did not mediate the settlement preference or avoidance behaviour of the cyprids (Fig. S4B).

To clarify whether PTMs of SIPC and not SIPC itself are responsible for the behaviour of cyprids, in the presence of SIPC, and identify which of its multiple domains is responsible for its biological activity, we generated truncated fragments of SIPC, maintaining the endogenous signal peptide, and expressed and purified the fragments from culture medium of HEK293 cells (Fig. 3A–D). The design of the fragments was based on a crystal structure simulation that we generated through i-TASSER (Zhang, 2008) that aided the choice of truncated fragment length. The process was iterative and informed by the bioinformatics analysis on the PTMs of SIPC (see Table S5 and Fig. S4C). The cumulative results (Fig. 3A–D) show that truncation of the first 212 amino acids results in a poorly secreted protein (Fig. 3C,D), an effect that was remedied when further truncations were made to the full-length SIPC (Fig. 3B–D). The apparent molecular masses of some of the various truncated SIPC fragments differed from their calculated molecular masses substantially but, as the size of the truncated fragments decreased, the apparent and calculated molecular masses became similar (Fig. 3A–D, Fig. S4D). As with the recombinant SIPC expressed in Sf9 cells (Fig. 2E), we tested the ability of all the

Journal of Experimental Biology (2018) 221, jeb185348. doi:10.1242/jeb.185348

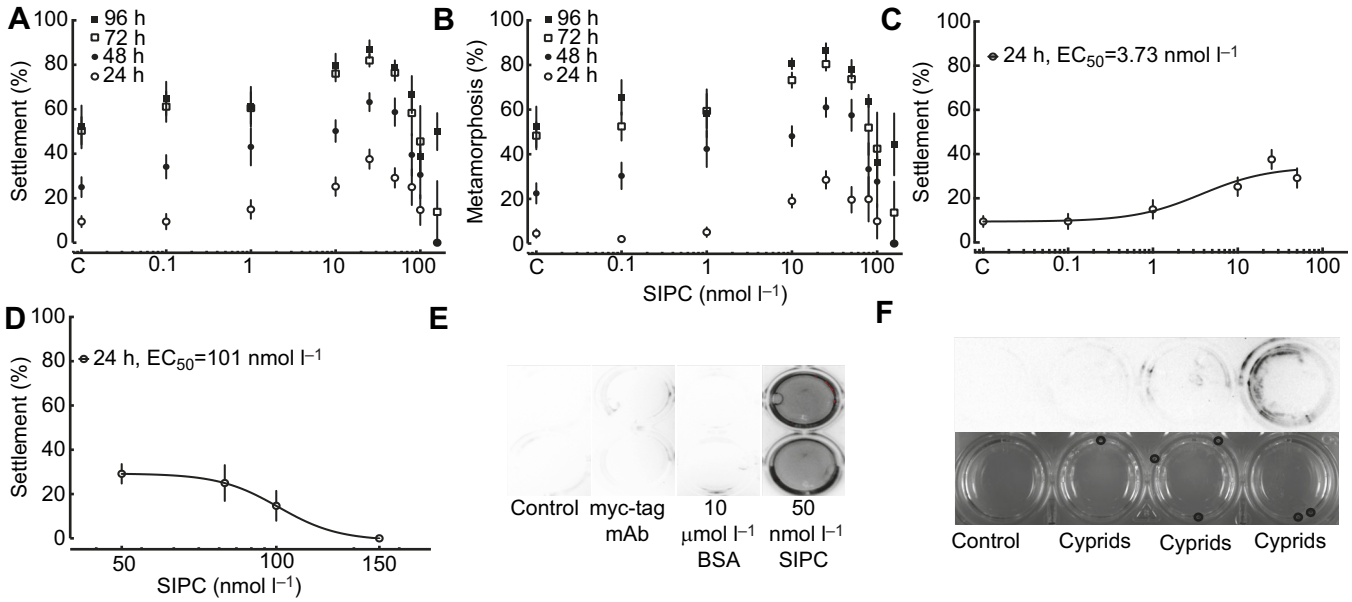

Fig. 2. Recombinant SIPC from *A. amphitrite* induces gregarious settlement preference or avoidance by cyprids of this species. (A) Recombinant SIPC transduces gregarious settlement preference and settlement avoidance behaviour by *A. amphitrite* cyprids in a dose-dependent manner. Values are expressed as the percentage of animals that settled from the total number of animals placed in a well of a 24-well tissue culture plate. The cumulative results from 3 independent experiments with 5 replicates each is shown. Results of one-way ANOVA with the Tukey's multiple comparisons test between matched observations for the various time points revealed no statistical significance only between the datasets at 72 and 96 h ($P>0.05$). Results of one-way ANOVA with the Bonferonni's multiple comparisons test between control and the various concentrations of recombinant SIPC (SpSIPC) revealed statistical significance only between the 10 nmol l^{-1}, 25 nmol l^{-1} and 50 nmol l^{-1} concentrations of SpSIPC and the control group ($P<0.05$) in all 4 tested time points. (B) Recombinant SIPC promotes metamorphosis of the settled *A. amphitrite* cyprids in a dose-dependent manner. Assay conditions and replicates are identical to A. For the percentage of animals that underwent metamorphosis, results of one-way ANOVA with the Tukey's multiple comparisons test between matched observations for the various time points revealed no statistical significance only between the datasets at 72 and 96 h ($P>0.05$). (C) Effective concentration of recombinant SIPC that induces gregarious settlement preference behaviour by *A. amphitrite* cyprids within 24 h. Assay conditions and replicates are identical to A and plotted separately for clarity. Results of one-way ANOVA with the Bonferonni's multiple comparisons test between control and the various concentrations of recombinant SIPC (SpSIPC) revealed statistical significance only between the 10 nmol l^{-1}, 25 nmol l^{-1} and 50 nmol l^{-1} concentrations of SpSIPC and the control group ($P<0.05$). (D) Effective concentration of recombinant SIPC that induces settlement avoidance behaviour by *A. amphitrite* cyprids within 24 h. Assay conditions and replicates are identical to A and plotted separately for clarity. Results of one-way ANOVA with the Bonferonni's multiple comparisons test between control and the various concentrations of recombinant SIPC (SpSIPC) revealed statistical significance only between the 50 nmol l^{-1} and 150 nmol l^{-1} concentrations of SpSIPC ($P<0.05$). (E) Recombinant SIPC adheres to the surface of polystyrene tissue culture plates used in settlement bioassays. Sterile tissue culture plates were incubated at 25°C empty (control) or with 0.3 ml of sterile artificial sea water (ASW) (Myc-Tag mAb) or with 0.3 ml of 10 μmol l^{-1} BSA in sterile ASW or with 0.3 ml of 50 nmol l^{-1} SpSIPC in sterile ASW. Solutions were then aspirated and western blotting carried out on the wells with a Myc-Tag (9B11) mAb. A typical image of several experiments ($n=3$) is shown. (F) A mouse anti-SIPC monoclonal antibody can detect native SIPC deposited by cyprids of *A. amphitrite* on the surface of polystyrene tissue culture plates used in settlement bioassays. Sterile tissue culture plates were incubated at 25°C without (control) or with 10 cyprids each in 2 ml sterile ASW for 96 h. Solutions and settled and non-settled cyprids were then removed and western blots were carried out on the wells with a mouse anti-SIPC monoclonal antibody. The bright-field images on the lower row contain drawn circles that show the position of the settled cyprids on each well, whereas the upper row shows the corresponding immunostained image. A typical image of several experiments ($n=6$) is shown.

truncated SIPC fragments to adhere to polystyrene plates (Fig. S4E). The results revealed that truncated SIPC fragments Tr.1 to Tr.3 adhere to the surface of polystyrene plates, while Tr.6 to Tr.9 do not (Fig. S4E).

We tested the settlement of cyprids against the truncated fragments of SIPC and observed that removal of the N-terminal 482 amino acids made the cyprids avoid settlement (Table 1, Fig. S5), whereas truncated fragments up until the N-terminal 482 amino acids made cyprids exhibit settlement preference rather than avoidance (Table 1, Fig. S5). The settlement avoidance behaviour was consistent even for the small C-terminal fragment of SIPC (Table 1, Fig. S5), which contains the receptor-binding domain of α2-macroglobulin [*A2M_recep* (PF07677.13)], the protein that is phylogenetically related to SIPC (Dreanno et al., 2006c) (Fig. S4C).

PTMs of recombinant SIPC

We plotted the differences in the apparent and calculated molecular mass of recombinant SIPC and its truncated fragments from several western blots in an attempt to map the location of PTMs on the primary structure of SIPC (Fig. 4A). Informed by a previous analysis of the *N*-linked glycosylations on native SIPC (Pagett et al., 2012), and by our bioinformatics analysis on the location of any PTMs on SIPC (see Table S5), we carried out deglycosylation assays on recombinant SIPC and selected truncated SIPC fragments (Fig. 4B–D) that would allow us to determine where such PTMs are located. We used a deglycosylation assay that provides more extensive removal of glycans (Fig. 4D) than previously reported with PNGase F, which removes only *N*-linked glycans (Pagett et al., 2012). Deglycosylation of the full-length SIPC reduced the molecular mass of the protein by 22.27±1.35 kDa (Fig. 4B, Table S6), while deglycosylation of Tr.5, Tr.6 and Tr.7 reduced the molecular mass of the fragments by 11.33±2.17, 13.91±1.83 and 7.04±1.24 kDa, respectively (Fig. 4C). We re-analysed the data presented in a previous study, in which the molecular nature and relative abundance of *N*-linked glycans, cleaved from native SIPC with PNGase F, was determined (Pagett et al., 2012), and found that *N*-linked glycans present on the native protein have a molecular mass of 14.7 kDa (Table S6).

Journal of Experimental Biology (2018) 221, jeb185348. doi:10.1242/jeb.185348

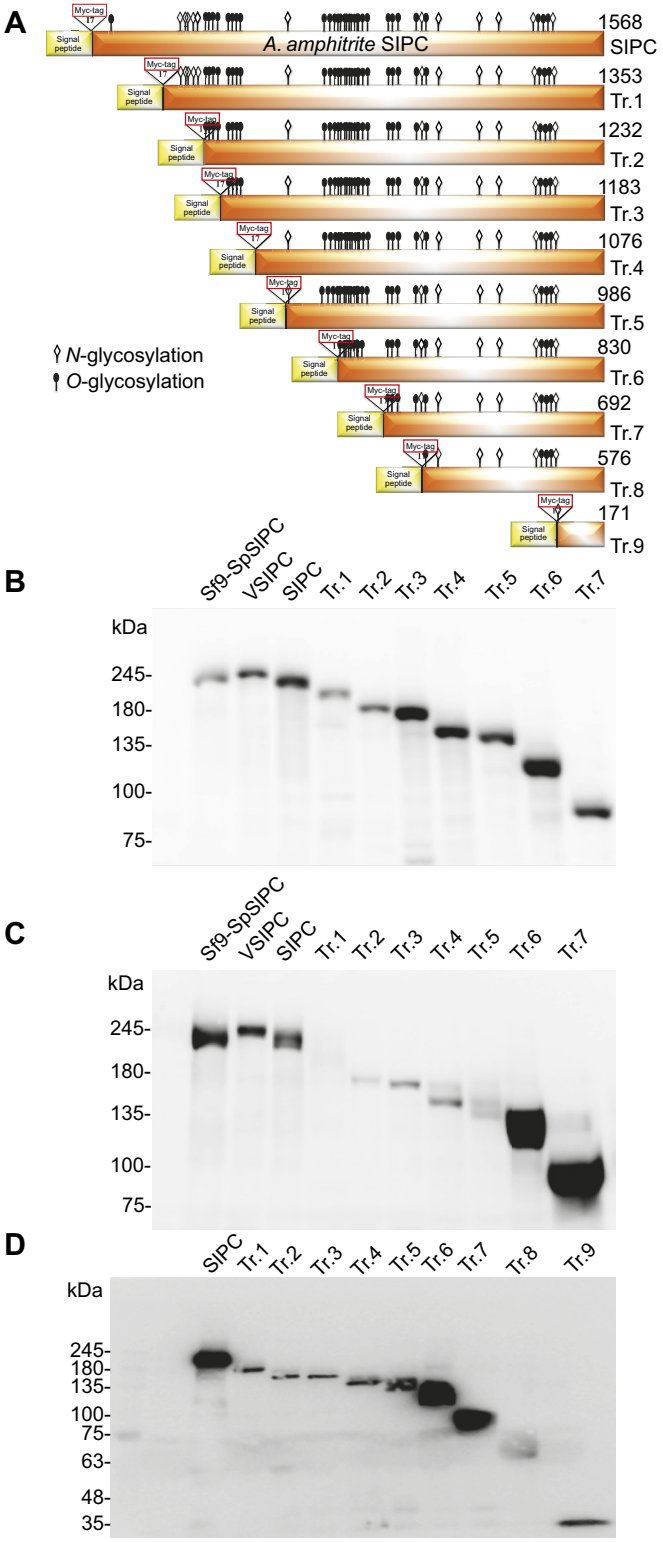

Fig. 3. Design and expression of truncated SIPC fragments. (A) Graphical representation of the truncated SIPC fragments showing the protein length and the post-translational modifications identified by our bioinformatic analysis. Tr.1–Tr.9, truncated SIPC fragments 1 to 9. (B) Expression of recombinant SIPC and its truncated fragments in HEK293 cells. Each lane contains 10 μg of total cellular protein. Western blots were probed with an anti-myc antibody and a typical image of several blots (n=3) is shown. (C) Secretion of recombinant SIPC and its truncated fragments by HEK293 cells. Each lane contains 20 μl of cell culture medium. Western blots were probed with an anti-myc antibody and a typical image, which corresponds to the samples shown in B, is shown. (D) Secretion of recombinant SIPC and its truncated fragments by HEK293 cells. Each lane contains 20 μl of cell culture medium. Western blots were probed with an anti-myc antibody and proteins were resolved less than C to include truncated fragments Tr.8 and Tr.9. Sf9-SpSIPC, recombinant SIPC secreted from Sf9 cells; VSIPC, Venus-tagged recombinant SIPC secreted from HEK293 cells.

By plotting these data (Pagett et al., 2012) with our deglycosylation results, the apparent molecular masses on western blots and the bioinformatic analysis, we estimate that *N*-linked glycans present on full-length SIPC and Tr.1 (Figs 3A and 4A) as well as Tr.6 and Tr.7 (Figs 3A and 4A) actually contribute 15 kDa to the increased molecular mass of SIPC, and less to its truncated fragments (see Table S6 for details). Numerous *O*-linked glycosylation sites are predicted on truncated SIPC fragments that are larger than Tr.4 (i.e. Tr.2 and Tr.3) and also in a region in the middle of the primary structure of the protein (Fig. 3A). Such *O*-linked glycans may contribute an additional 8 kDa cumulatively to the full-length SIPC (see Tables S5 and S6 for details). The overall results, shown as a Circos figure (Krzywinski et al., 2009) (Fig. 4E), depict the relative contribution of each putative *N*-linked, *O*-linked or other simple glycans on the increase in the molecular mass of SIPC and its truncated fragments. The plot was created by a combination of bioinformatic analysis, western blots analysis, deglycosylation assays and previously published data (Pagett et al., 2012) to show how *N*-linked or *O*-linked glycans alter the molecular mass of full-length SIPC or its truncated fragments. The data table that was used to generate the Circos plot and the plot parameters are provided in Dataset 1.

DISCUSSION

SIPC (Dreanno et al., 2007; Matsumura et al., 1998b) serves as a mate-discerning (Dreanno et al., 2007) and mate-assessing (Matsumura et al., 1998b; Pagett et al., 2012) chemical cue but also serves as a predator-attracting chemical cue (Ferrier et al., 2016; Zimmer et al., 2016). In this context, and balancing the trade-off between reproduction and predation, having both settlement preference and avoidance cues conveyed by the same protein, i.e. SIPC, may be advantageous for barnacles because it will inform them of mate availability when the chemical cue is sparse enough to

Table 1. Recombinant SIPC (rSIPC) or its truncated fragments transduce gregarious settlement preference or settlement avoidance behaviour by *A. amphitrite* cyprids

rSIPC and truncated SIPC fragments (see Fig. 3A)	Settlement preference EC_{50} (nmol l^{-1})	Settlement avoidance EC_{50} (nmol l^{-1})
SIPC (rSIPC)	1.56	–
Tr.1	1.19	–
Tr.2	1.49	–
Tr.3	26.6	–
Tr.4	14.1	–
Tr.5	–	10.8
Tr.6	–	12.4
Tr.7	–	4.75
Tr.8	–	14.2
Tr.9	–	550

Settlement preference and avoidance EC_{50} values are derived from dose–response curves shown in Fig. S5 and were calculated using GraphPad Prism v.6. Values were computed from the percentage of animals that settled within 24 h from the total number of animals placed in a well of a 24-well tissue culture plate. The cumulative results from 2 independent experiments with 4 replicates each is shown. See Fig. S5 for further details.

Journal of Experimental Biology (2018) 221, jeb185348. doi:10.1242/jeb.185348

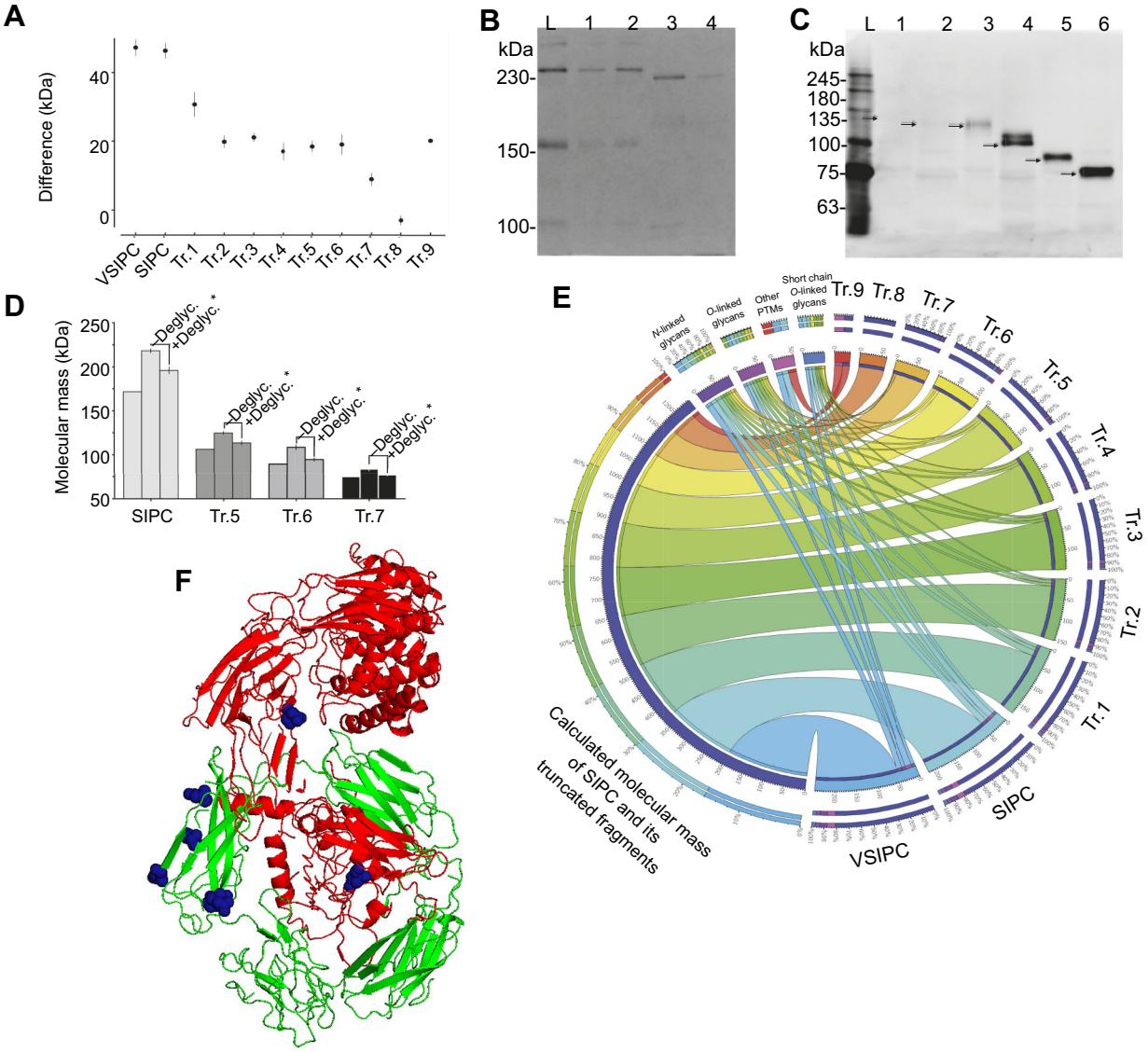

Fig. 4. Post-translational modifications of SIPC. (A) Differences between the apparent and calculated molecular mass of SIPC and its truncated fragments. Analysis by ImageJ software of several western blots (*n*=3–7) is shown. (B) Deglycosylation assays of recombinant SIPC secreted by Sf9 cells. Each lane contains 20 µl of cell culture medium. Western blots were probed with an anti-myc antibody and a typical image of several blots (*n*=3) is shown. L, protein ladder; 1 and 2, SpSIPC culture medium; 3, SpSIPC culture medium after deglycosylation assay in non-denaturing reaction conditions; 4, SpSIPC culture medium after deglycosylation assay in denaturing reaction conditions. (C) Deglycosylation assays in non-denaturing reaction conditions of recombinant SIPC fragments secreted by HEK293 cells. Each lane contains 20 µl of cell culture medium. Western blots were probed with an anti-myc antibody and a typical image of several blots (*n*=3) is shown. L, protein ladder; 1, Tr.5 culture medium; 2, Tr.5 culture medium after deglycosylation; 3, Tr.6 culture medium; 4, Tr.6 culture medium after deglycosylation; 5, Tr.7 culture medium; 6, Tr.7 culture medium after deglycosylation. Arrows indicate the migration of the bands before and after the deglycosylation reaction. (D) Differences between the calculated, apparent and deglycosylated molecular masses of SIPC and selected truncated SIPC fragments. Analysis by ImageJ software of the migration of recombinant SIPC or the truncated SIPC fragments and calibration with the standard molecular mass of the protein ladder from several western blots (*n*=3) is shown. For each protein (*x*-axis), its calculated molecular mass is the left bar, its apparent molecular mass is the middle bar and its deglycosylated molecular mass is the right bar. Results from paired *t*-tests with Welch's correction [shown with an asterisk (*)] indicate that deglycosylation resulted in statistically significant ($P<0.05$) decreases in the apparent molecular mass of all tested truncated fragments and the full-length SIPC. (E) Circos plot showing the relative contribution of different post-translational modifications (PTMs), identified on SIPC or its truncated fragments, to the difference between the apparent and calculated molecular mass of each protein. Percentage values are related to kDa. See Dataset 1 for details. (F) Crystal structure simulation of SIPC showing the locations of *N*-linked glycosylations (dark blue spheres) that could be experimentally predicted in this study. The N-terminal domains of SIPC that induce gregarious settlement behaviour by *A. amphitrite* cyprids are shown in green. The C-terminal domains of SIPC that induce settlement avoidance by *A. amphitrite* cyprids are shown in red.

ensure reproductive success without reproductive competition, and will inform them of reproductive competition and potential predation when the chemical cue is abundant.

We show here that a recombinant form of SIPC from *A. amphitrite* serves such a dual role because it transduces both settlement preference and settlement avoidance cues to cyprids of this species in a dose-dependent manner in laboratory conditions (Fig. 2). The settlement preference behaviour in the presence of endogenous SIPC has been recorded before (Matsumura et al., 1998b; Dreanno et al., 2006c; Pagett et al., 2012) with a reported EC_{50} of 102 nmol l^{-1} for settlement preference (Pagett et al., 2012), but possibly the amounts of protein needed to produce a full-range

dose response precluded the observation of settlement avoidance behaviour. In addition, positive-settlement-directed purification of *B. glandula* MULTIFUNCin, an orthologue of *A. amphitrite* SIPC (Ferrier et al., 2016), showed that the 199 kDa full-length MULTIFUNCin and its 98 kDa N-terminal subunit possess settlement-inducing biological activity, but no data are available for the 88 kDa and the 76 kDa subunits of either MULTIFUNCin (Ferrier et al., 2016) or SIPC (Matsumura et al., 1998b; Dreanno et al., 2007).

By engineering a secreted recombinant SIPC, we obtained a complete picture of the response of *A. amphitrite* cyprids to the presence of SIPC (Figs 1 and 2). To achieve this, we had to replace the signal peptide sequence of endogenous SIPC (Fig. 1, Fig. S1), a result that agrees with previous suggestions that SIPC may not be an actively secreted protein in *A. amphitrite* but rather a deposited protein (Zhang et al., 2016; Dreanno et al., 2006a) present on their cuticle (Crisp and Meadows, 1962; Dreanno et al., 2006a,c) as well as a part of the temporary adhesive that these animals use in their plankto-benthic transition (Dreanno et al., 2006b).

The presence of PTMs on SIPC has been postulated ever since its discovery (Dreanno et al., 2006c) and further research has shown that glycans, such as mannose (Pagett et al., 2012), stimulate settlement of *A. amphitrite* cyprids (Matsumura et al., 1998a). Furthermore, the large difference between the apparent and calculated molecular mass of SIPC is an argument in favour of data that showed that PTMs (i.e. *N*-linked glycans) bring about the settlement preference response of *A. amphitrite* cyprids (Pagett et al., 2012) (Fig. 2A), so it can be argued that recombinant SIPC itself does not have a biological activity. Bioinformatic analysis of the SIPC primary sequence showed that there are several putative *N*-linked glycosylation sites and many *O*-linked glycosylation sites on SIPC (Table S5). We also identified several other putative PTM consensus sequences on SIPC (such as SUMOylation, ubiquitination or palmitoylation sites), but these modifications are reversible and we have not observed either full-length recombinant SIPC or its truncated fragments to migrate in western blots as more than one band (Figs 1 and 3). It should be noted that Tr.9 showed a consistently (n=4) high difference of ∼20 kDa between the observed and calculated molecular mass for reasons we cannot, at present, explain (Fig. 4A,E). The same applies to the difference in the observed and calculated molecular mass of SIPC and VSIPC, which contain an additional ∼15 kDa in mass that seems to be associated with the N-terminal 282 amino acids of the protein. This additional mass of ∼15 kDa may not be a protein because our LC-MS/MS analysis on SDS-PAGE gel-isolated purified recombinant SIPC from Sf9 cells (Fig. S4F) yielded no evidence of conjugation of SIPC with proteins of small molecular mass that would explain this ∼15 kDa difference (Table S6). Nevertheless, this additional entity of ∼15 kDa does not seem to contribute to the settlement response of cyprids because the settlement assays showed that exclusion of the N-terminal 245 amino acids of SIPC does not change the settlement response of cyprids (Table 1, Fig. S5).

Based on the apparent molecular mass of the truncated SIPC fragments, the deglycosylation assays and previous data (Pagett et al., 2012), some of the putative PTMs could not be supported by our experimental results (Tables S5 and S6, Fig. 4A). For example, *N*-linked glycosylation sites at the C-terminus of recombinant SIPC do not seem to contain *N*-linked glycans, because a truncated SIPC fragment (Tr.8) has an apparent molecular mass equivalent to its calculated (Fig. 4A, Table S6), whereas Tr.7, when deglycosylated, does not yield a molecular mass reduction that would justify the existence of *N*-linked glycans at the C-terminus of SIPC (Fig. 4A,D). The same is true for deglycosylated fragments Tr.5 and Tr.6, which contain only two *N*-linked glycosylation sites (Fig. 4A–D).

Research on the PTMs of SIPC showed that high mannose-type oligosaccharides are present on SIPC and mannose can induce settlement by the cyprids (Pagett et al., 2012), an indication that glycans that are present on SIPC influence the settlement preference response. We suggest that, if indeed *N*-linked glycans contribute to the settlement-inducing activity of SIPC, then these glycans should be located at the N-terminal 600 amino acids of SIPC and more specifically at asparagines 263, 285, 289, 302 and 305 of *A. amphitrite* recombinant SIPC (Figs 3A and 4F). Barnacles can discriminate between conspecific and allospecific SIPC in gregarious settlement assays (Dreanno et al., 2007; Crisp and Meadows, 1962) and the SIPC proteins identified so far from various barnacle species (Yorisue et al., 2012) have conserved *N*-linked glycosylation sites. Clearly deserving further investigation by site-directed mutagenesis of the *N*-linked glycosylation sites of SIPC, we posit that *N*-linked glycans on SIPC (Pagett et al., 2012) must play an auxiliary rather than decisive role in the gregarious settlement of the cyprids. Putative *N*-linked glycosylation sites are present at the C-terminus of SIPC (Fig. 4F), but these sites may be non-glycosylated, based on the apparent molecular mass of truncated SIPC fragments (Figs 3A and 4A) and our deglycosylation assays (Fig. 4B,C).

The bimodal response of barnacles to our engineered, recombinant form of SIPC shows that the settlement behaviour, and thereby the reproductive success, of these sessile species is a conspecific cue-guided response. We propose that, to avoid reproductive competition and possible predation, barnacles will not settle if conspecific cues inform of overcrowding but will settle gregariously when conspecific cues are sparse.

To clarify how SIPC transduces such bimodal responses at the molecular level, one would need to identify its endogenous receptor in *A. amphitrite*. Although complex ligand–receptor responses, dependent on receptor expression, have been recently documented in other cellular systems (Jenni et al., 2015; Fathi et al., 2016), the absence of the genomic sequence of *A. amphitrite* and the limited available transcriptomic data make any attempt to explain at the molecular level the bimodal response of cyprids to SIPC highly speculative, at present.

Acknowledgements
We thank Dr R. Kirby and Dr A. Clare for the generous gift of a vector containing the cDNA of SIPC from *A. amphitrite*.

Competing interests
The authors declare no competing or financial interests.

Author contributions
Conceptualization: M.K., S.G.D.; Methodology: M.K., M.P., S.M., M.Z., D.V., I.K., M.S., S.G.D.; Software: M.K., S.G.D.; Validation: M.K., M.P., S.M., S.G.D.; Formal analysis: M.K., M.P., S.M., M.Z., S.G.D.; Investigation: M.K., M.P., S.M., S.G.D.; Resources: M.K., M.Z., D.V., I.K., M.S., S.G.D.; Data curation: M.K., M.P., M.Z., S.G.D.; Writing - original draft: S.G.D.; Writing - review & editing: M.K., S.G.D.; Visualization: M.K., M.P., S.M., M.Z., S.G.D.; Supervision: S.G.D.; Project administration: S.G.D.; Funding acquisition: S.G.D.

Funding
Research reported in this publication was supported from funds by the National and Kapodistrian University of Athens under award number Kapodistrias:11240 and by the GSRT-EPANII-11SYN-5-1274 'Maripaints' research grant from the Greek General Secretariat for Research and Technology (GSRT) and the European Social Fund (ESF).

Data availability
The mass spectrometry proteomics data and search results have been deposited to the ProteomeXchange with the project accession: PXD006858 and project doi:10.6019/PXD006858.

Supplementary information

Supplementary information available online at
http://jeb.biologists.org/lookup/doi/10.1242/jeb.185348.supplemental

References

Aldred, N. and Clare, A. S. (2008). The adhesive strategies of cyprids and development of barnacle-resistant marine coatings. *Biofouling* **24**, 351-363.

Barnes, H. and Barnes, M. (1977). Studies on the reproduction of cirripedes. I. Introduction: copulation, release of oocytes, and formation of the egg lamellae. *J. Exp. Mar. Biol. Ecol.* **27**, 195-218.

Baxter, R. H. G., Chang, C.-I., Chelliah, Y., Blandin, S., Levashina, E. A. and Deisenhofer, J. (2007). Structural basis for conserved complement factor-like function in the antimalarial protein TEP1. *Proc. Natl Acad. Sci. USA* **104**, 11615-11620.

Bernard, F. J. and Lane, C. E. (1962). Early settlement and metamorphosis of the barnacle Balanus amphitrite niveus. *J. Morphol.* **110**, 19-39.

Caromile, L. A., Oganesian, A., Coats, S. A., Seifert, R. A. and Bowen-Pope, D. F. (2010). The neurosecretory vesicle protein phogrin functions as a phosphatidylinositol phosphatase to regulate insulin secretion. *J. Biol. Chem.* **285**, 10487-10496.

Costlow, J. D. (1956). Shell development in Balanus improvisus Darwin. *J. Morphol.* **99**, 359-415.

Crisp, D. J. (1961). Territorial behaviour in Barnacle settlement. *J. Exp. Biol.* **38**, 429-446.

Crisp, D. J. and Meadows, P. S. (1962). The chemical basis of gregariousness in cirripedes. *Proc. R. Soc. Lond. Ser. B. Biol. Sci.* **156**, 500-520.

Darwin, C. (1854). *A monograph on the sub-class Cirripedia, with figures of all the species. The Balanidae (or sessile cirripedes); the Verrucidae, etc., etc., etc.* London: Ray Society.

Dreanno, C., Kirby, R. R. and Clare, A. S. (2006a). Locating the barnacle settlement pheromone: spatial and ontogenetic expression of the settlement-inducing protein complex of Balanus amphitrite. *Proc. R. Soc. B* **273**, 2721-2728.

Dreanno, C., Kirby, R. R. and Clare, A. S. (2006b). Smelly feet are not always a bad thing: the relationship between cyprid footprint protein and the barnacle settlement pheromone. *Biol. Lett.* **2**, 423-425.

Dreanno, C., Matsumura, K., Dohmae, N., Takio, K., Hirota, H., Kirby, R. R. and Clare, A. S. (2006c). An α2-macroglobulin-like protein is the cue to gregarious settlement of the barnacle Balanus amphitrite. *Proc. Natl Acad. Sci. USA* **103**, 14396-14401.

Dreanno, C., Kirby, R. R. and Clare, A. S. (2007). Involvement of the barnacle settlement-inducing protein complex (SIPC) in species recognition at settlement. *J. Exp. Mar. Biol. Ecol.* **351**, 276-282.

Fathi, S., Nayak, C. R., Feld, J. J. and Zilman, A. G. (2016). Absolute ligand discrimination by dimeric signaling receptors. *Biophys. J.* **111**, 917-920.

Ferrier, G. A., Kim, S. J., Kaddis, C. S., Loo, J. A., Ann Zimmer, C. and Zimmer, R. K. (2016). MULTIFUNCin: a multifunctional protein cue induces habitat selection by, and predation on, Barnacles. *Integr. Comp. Biol.* **56**, 901-913.

Goedhart, J., von Stetten, D., Noirclerc-Savoye, M., Lelimousin, M., Joosen, L., Hink, M. A., van Weeren, L., Gadella, T. W. J. and Royant, A. (2012). Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nat. Commun.* **3**, 751.

Hellio, C., Simon-Colin, C., Clare, A. and Deslandes, E. (2004). Isethionic acid and floridoside isolated from the red alga, Grateloupia turuturu, inhibit settlement of Balanus amphitrite cyprid larvae. *Biofouling* **20**, 139-145.

Hellio, C., Tsoukatou, M., Maréchal, J.-P., Aldred, N., Beaupoil, C., Clare, A. S., Vagias, C. and Roussis, V. (2005). Inhibitory effects of mediterranean sponge extracts and metabolites on larval settlement of the barnacle Balanus amphitrite. *Mar. Biotechnol.* **7**, 297-305.

Høeg, J. T. and Møller, O. S. (2006). When similar beginnings lead to different ends: constraints and diversity in cirripede larval development. *Invertebr. Reprod. Dev.* **49**, 125-142.

Huang, K.-Y., Su, M.-G., Kao, H.-J., Hsieh, Y.-C., Jhong, J.-H., Cheng, K.-H., Huang, H.-D. and Lee, T.-Y. (2016). dbPTM 2016: 10-year anniversary of a resource for post-translational modification of proteins. *Nucleic Acids Res.* **44**, D435-D446.

Jenni, S., Goyal, Y., von Grotthuss, M., Shvartsman, S. Y. and Klein, D. E. (2015). Structural basis of neurohormone perception by the receptor tyrosine kinase torso. *Mol. Cell* **60**, 941-952.

Köhler, G. and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**, 495.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J. and Marra, M. A. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639-1645.

Lau, S. C. K. and Qian, P.-Y. (2000). Inhibitory effect of phenolic compounds and marine bacteria on larval settlement of the barnacle Balanus amphitrite amphitrite darwin. *Biofouling* **16**, 47-58.

Lucas, M. I., Walker, G., Holland, D. L. and Crisp, D. J. (1979). An energy budget for the free-swimming and metamorphosing larvae of Balanus balanoides (Crustacea: Cirripedia). *Mar. Biol.* **55**, 221-229.

Matsumura, K., Mori, S., Nagano, M. and Fusetani, N. (1998a). Lentil lectin inhibits adult extract-induced settlement of the barnacle, Balanus amphitrite. *J. Exp. Zool.* **280**, 213-219.

Matsumura, K., Nagano, M. and Fusetani, N. (1998b). Purification of a larval settlement-inducing protein complex (SIPC) of the barnacle, Balanus amphitrite. *J. Exp. Zool.* **281**, 12-20.

Nagai, T., Ibata, K., Park, E. S., Kubota, M., Mikoshiba, K. and Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotech.* **20**, 87-90.

O'reilly, D. R., Kelly, T. J., Masler, E. P., Thyagaraja, B. S., MOY Robson, R., Shaw, T. C. and Miller, L. K. (1995). Overexpression of Bombyx mori prothoracicotropic hormone using baculovirus vectors. *Insect Biochem. Mol. Biol.* **25**, 475-485.

Pagett, H. E., Abrahams, J. L., Bones, J., O'Donoghue, N., Marles-Wright, J., Lewis, R. J., Harris, J. R., Caldwell, G. S., Rudd, P. M. and Clare, A. S. (2012). Structural characterisation of the N-glycan moiety of the barnacle settlement-inducing protein complex (SIPC). *J. Exp. Biol.* **215**, 1192-1198.

Pejaver, V., Hsu, W.-L., Xin, F., Dunker, A. K., Uversky, V. N. and Radivojac, P. (2014). The structural and functional signatures of proteins that undergo multiple events of post-translational modification. *Protein Sci.* **23**, 1077-1093.

Phang, I. Y., Chaw, K. C., Choo, S. S. H., Kang, R. K. C., Lee, S. S. C., Birch, W. R., Teo, S. L. M. and Vancso, G. J. (2009). Marine biofouling field tests, settlement assay and footprint micromorphology of cyprid larvae of Balanus amphitrite on model surfaces. *Biofouling* **25**, 139-147.

Rittschof, D. (1990). Peptide-mediated behaviors in marine organisms Evidence for a common theme. *J. Chem. Ecol.* **16**, 261-272.

Rittschof, D. and Cohen, J. H. (2004). Crustacean peptide and peptide-like pheromones and kairomones. *Peptides* **25**, 1503-1516.

Rittschof, D., Branscomb, E. S. and Costlow, J. D. (1984). Settlement and behavior in relation to flow and surface in larval barnacles, Balanus amphitrite Darwin. *J. Exp. Mar. Biol. Ecol.* **82**, 131-146.

Rittschof, D., Clare, A. S., Gerhart, D. J., Mary, S. A. and Bonaventura, J. (1992). Barnacle in vitro assays for biologically active substances: toxicity and Settlement inhibition assays using mass cultured Balanus amphitrite amphitrite darwin. *Biofouling* **6**, 115-122.

Rittschof, D., Lai, C.-H., Kok, L.-M. and Teo, S. L.-M. (2003). Pharmaceuticals as antifoulants: concept and principles. *Biofouling* **19**, 207-212.

Shi, X. and Jarvis, D. (2007). Protein N-glycosylation in the Baculovirus-insect cell system. *Curr. Drug Targets* **8**, 1116-1125.

Thiyagarajan, V., Harder, T. and Qian, P.-Y. (2002). Effect of the physiological condition of cyprids and laboratory-mimicked seasonal conditions on the metamorphic successes of Balanus amphitrite Darwin (Cirripedia; Thoracica). *J. Exp. Mar. Biol. Ecol.* **274**, 65-74.

Vizcaíno, J. A., Csordas, A., Del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T. et al. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447-D456.

Wasmeier, C. and Hutton, J. C. (1996). Molecular cloning of Phogrin, a protein-tyrosine phosphatase homologue localized to insulin secretory granule membranes. *J. Biol. Chem.* **271**, 18161-18170.

Yorisue, T., Matsumura, K., Hirota, H., Dohmae, N. and 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.

Yule, A. B. and Crisp, D. J. (1983). Adhesion of cypris larvae of the barnacle, Balanus balanoides, to clean and arthropodin treated surfaces. *J. Mar. Biol. Assoc. UK* **63**, 261-271.

Yule, A. B. and Walker, G. (1985). Settlement of balanus balanoides: the effect of cyprid antennular secretion. *J. Mar. Biol. Assoc. UK* **65**, 707-712.

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40.

Zhang, G., Yang, X.-X., Leung, P. M., He, L.-S., Chan, T. Y., Yan, G.-Y., Zhang, Y., Sun, J., Xu, Y. and Qian, P.-Y. (2016). Secretory locations of SIPC in Amphibalanus amphitrite cyprids and a novel function of SIPC in biomineralization. *Sci. Rep.* **6**, 29376.

Zimmer, R. K., Ferrier, G. A., Kim, S. J., Kaddis, C. S., Zimmer, C. A. and Loo, J. A. (2016). A multifunctional chemical cue drives opposing demographic processes and structures ecological communities. *Ecology* **97**, 2232-2239.

Zobell, C. (1941). Studies on marine bacteria. I. The cultural requirements of heterotrophic aerobes. *J. Mar. Res.* **4**, 42-75.

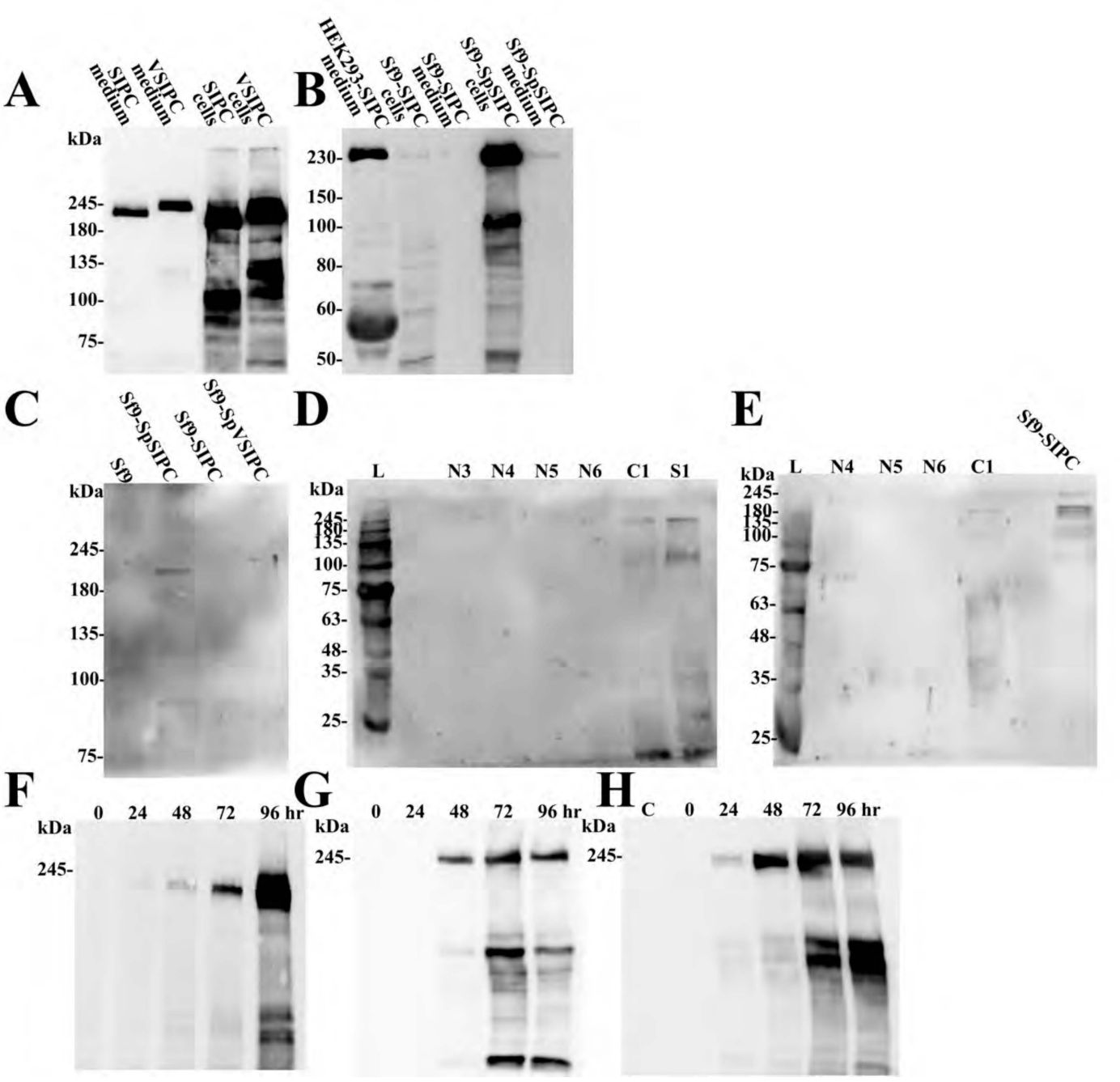

Fig. S1: Expression of recombinant SIPC from *A. amphitrite*.

A: Recombinant SIPC and VSIPC are secreted by HEK293 cells. Gel loading and Western blots conditions are identical to Fig. 1A. A typical image of several blots (n=4) is shown.

B: An engineered signal peptide in recombinant SIPC (SpSIPC) results in secretion of SIPC by insect Sf9 cells. Gel loading and Western blots conditions are identical to Fig. 1A. A typical image of several blots (n=3) is shown.

C: Detection of recombinant SIPC with a rabbit anti-SIPC polyclonal antibody. Each lane contains 20 µl of Sf9 cell culture medium. A typical image of several blots (n=3) is shown.

D: Detection of native SIPC, in extracts from *A. amphitrite* different developmental stages, with a rabbit anti-SIPC polyclonal antibody. Each lane contains 10 µg of total cellular protein from different developmental stages of this species. A typical image of several blots (n=3) is shown. N3 to N6: Nauplii stage 3 to stage 6, C1: Cyprid day 1, S1: Settled day 1, L: Protein Ladder.

E: Native SIPC has a MW similar to recombinant SIPC. Each lane contains 10 µg of total cellular protein from different developmental stages of *A. amphitrite* or 20 µl of Sf9 cell culture medium. A typical image of several blots (n=3) is shown. N4 to N6: Nauplii stage 4 to stage 6, C1: Cyprid day 1.

F: Time course of expression of recombinant SpSIPC by Sf9 cells after infection with recombinant SpSIPC-expressing baculoviruses. Each lane contains 10 µg of total cellular protein from Sf9 cells.

G: Time course of expression of recombinant VSpSIPC by Sf9 cells after infection with recombinant VSpSIPC-expressing baculoviruses. Each lane contains 10 µg of total cellular protein from Sf9 cells.

H: Time course of secretion of recombinant VSpSIPC from Sf9 cells after infection with recombinant VSpSIPC-expressing baculoviruses. Each lane contains 20 µl of Sf9 cell culture medium. Western blots were probed with an anti-myc antibody and a typical image of several blots (n=3) is shown.

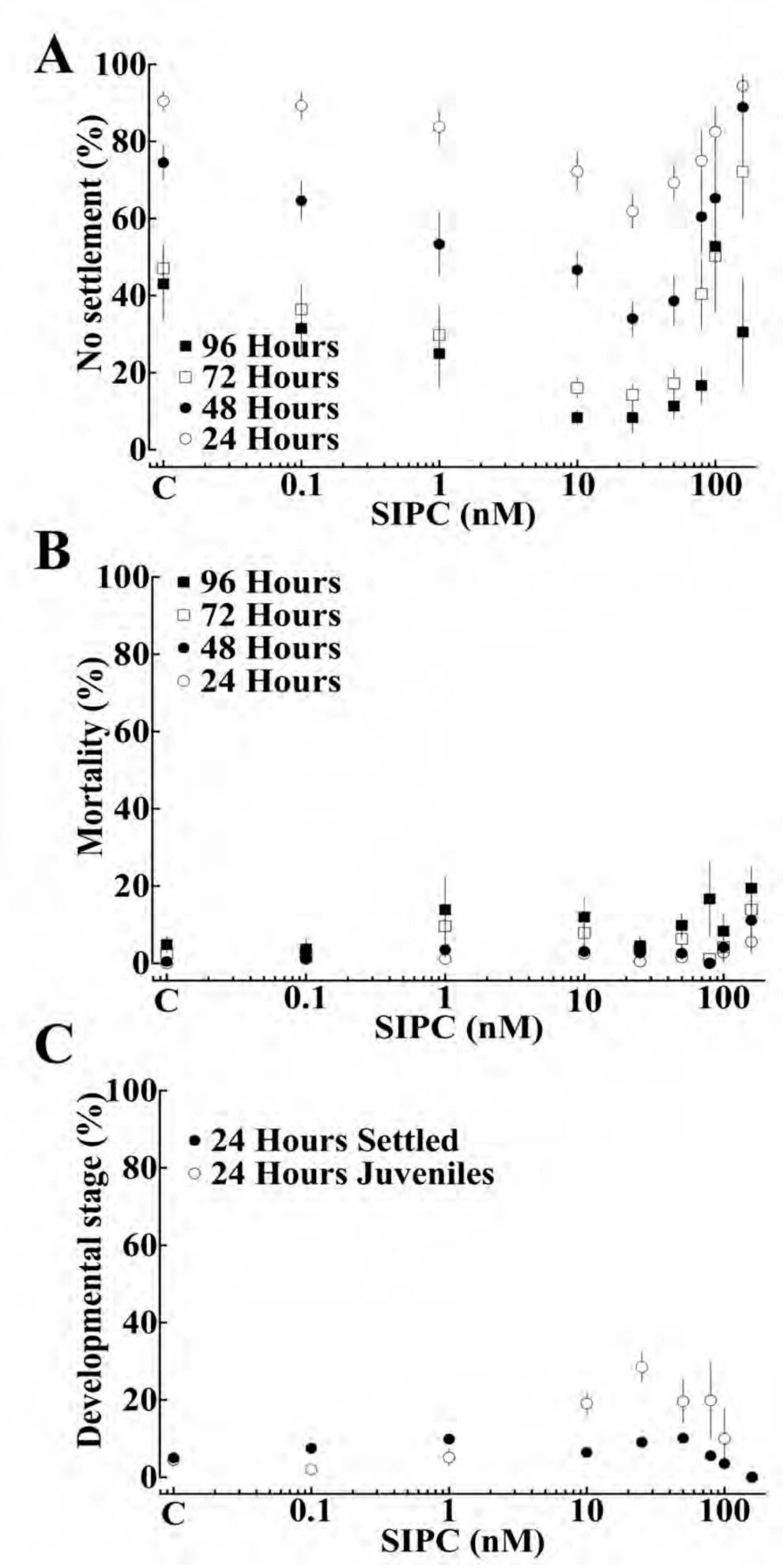

Fig. S2

A: High concentrations of recombinant SIPC induce settlement avoidance by *A. amphitrite* cyprids. Values are expressed as the percentage of animals that settled from the total number of animals placed in a well of a 24-well tissue culture plate. The cumulative results from 3 independent experiments with 5 replicates each is shown. See statistical analysis in legend of Figure 2A.

B: Recombinant SIPC does not have any effect on mortality of *A. amphitrite* cyprids in settlement assays. Values are expressed as the percentage of animals that settled from the total number of animals placed in a well of a 24-well tissue culture plate. The cumulative results from 3 independent experiments with 5 replicates each is shown. Results of one-way ANOVA with the Tukey's multiple comparisons test between matched observations for the various time points revealed no statistical significance between the datasets ($p=0.0753$). Results of one-way ANOVA with the Bonferonni's multiple comparisons test between control and the various concentrations of recombinant SIPC (SpSIPC) revealed no statistical significance ($p>0.05$) in all 4 tested time points.

C: Recombinant SIPC induces settlement and concomitant metamorphosis within 24 hrs in *A. amphitrite* cyprids. Values are expressed as the percentage of animals that settled from the total number of animals placed in a well of a 24-well tissue culture plate. The cumulative results from 3 independent experiments with 5 replicates each is shown. Results of unpaired, one-tailed t-tests between the percentages of settled and metamorphosed animals revealed that SpSIPC did not induce metamorphosis of the animals within 24 hrs ($p=0.06$).

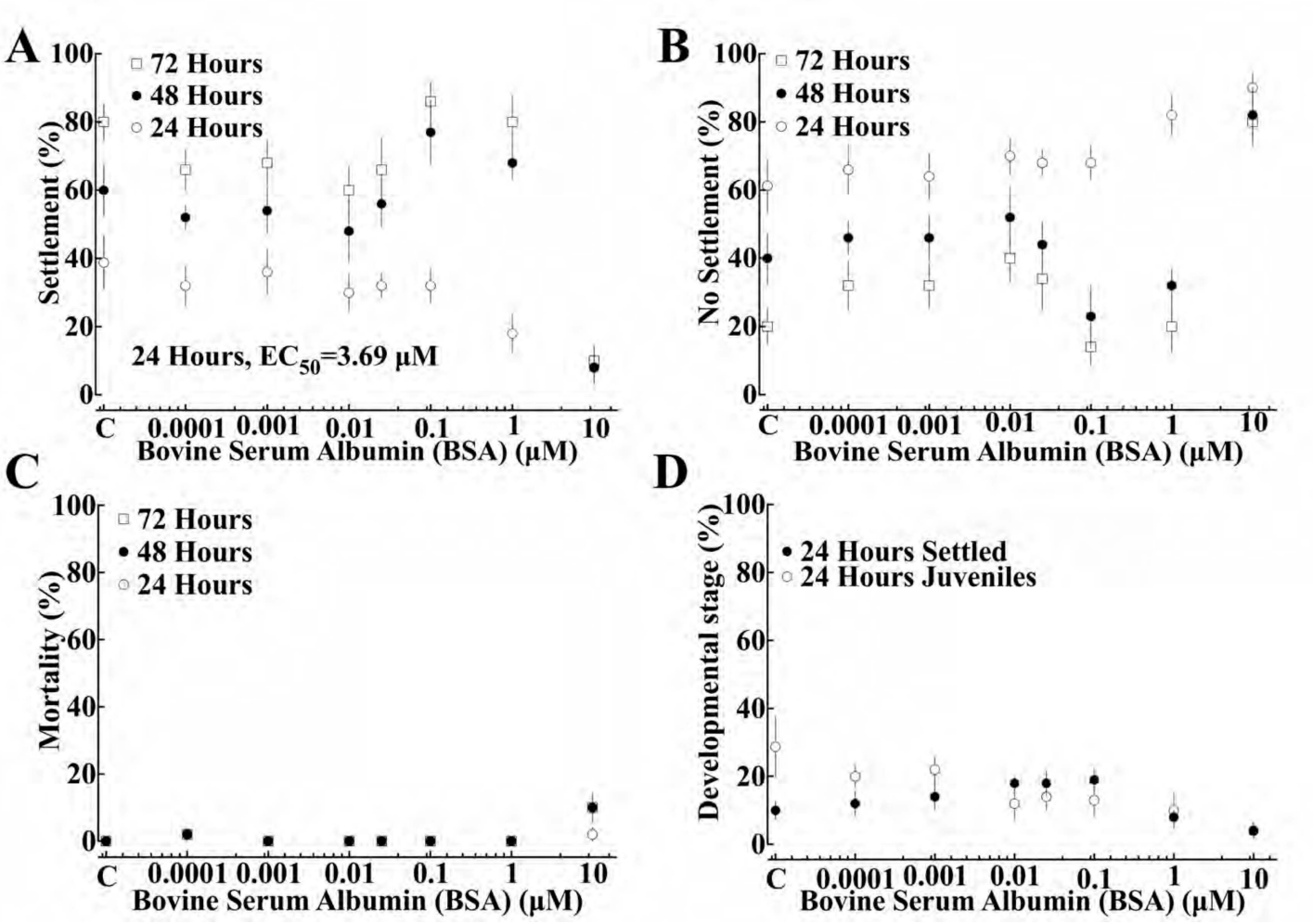

Fig. S3

A: Effect of Bovine Serum Albumin (BSA) on settlement behaviour of *A. amphitrite* cyprids. Values are expressed as the percentage of animals that settled from the total number of animals placed in a well of a 24-well tissue culture plate. The cumulative results from 3 independent experiments with 5 replicates each is shown. Results of one-way ANOVA with the Tukey's multiple comparisons test between matched observations for the various time points revealed statistical significance between all datasets ($p<0.05$). Results of one-way ANOVA with the Bonferonni's multiple comparisons test between control and the various concentrations of BSA revealed that BSA inhibited settlement only at a concentration of 10 µM and at 48 and 72 hrs.

B: Only concentrations of 10 μM of Bovine Serum Albumin (BSA) inhibit settlement of *A. amphitrite* cyprids. Values are expressed as the percentage of animals that settled from the total number of animals placed in a well of a 24-well tissue culture plate. The cumulative results from 3 independent experiments with 5 replicates each is shown. Results of one-way ANOVA with the Tukey's multiple comparisons test between matched observations for the various time points revealed statistical significance between all datasets ($p<0.05$). Results of one-way ANOVA with the Bonferonni's multiple comparisons test between control and the various concentrations of BSA revealed that BSA inhibited settlement only at a concentration of 10 μM at 24, 48 and 72 hrs.

C: Bovine Serum Albumin (BSA) does not have any effect on mortality of *A. amphitrite* cyprids in settlement assays. Values are expressed as the percentage of animals that settled from the total number of animals placed in a well of a 24-well tissue culture plate. The cumulative results from 3 independent experiments with 5 replicates each is shown. Results of one-way ANOVA with the Tukey's multiple comparisons test between matched observations for the various time points revealed no statistical significance between all datasets ($p>0.05$). Results of one-way ANOVA with the Bonferonni's multiple comparisons test between control and the various concentrations of BSA revealed that BSA increased mortality rates of cyprids only at a concentration of 10 μM and at 48 and 72 hrs.

D: Bovine Serum Albumin (BSA) does not have any effect on settlement and concomitant metamorphosis within 24 hrs in *A. amphitrite* cyprids. Values are expressed as the percentage of animals that settled from the total number of animals placed in a well of a 24-well tissue culture plate. The cumulative results from 3 independent experiments with 5 replicates each is shown. Results of unpaired, one-tailed t-tests between the percentages of settled and metamorphosed animals revealed that BSA did not promoted metamorphosis of the animals within 24 hrs ($p=0.225$).

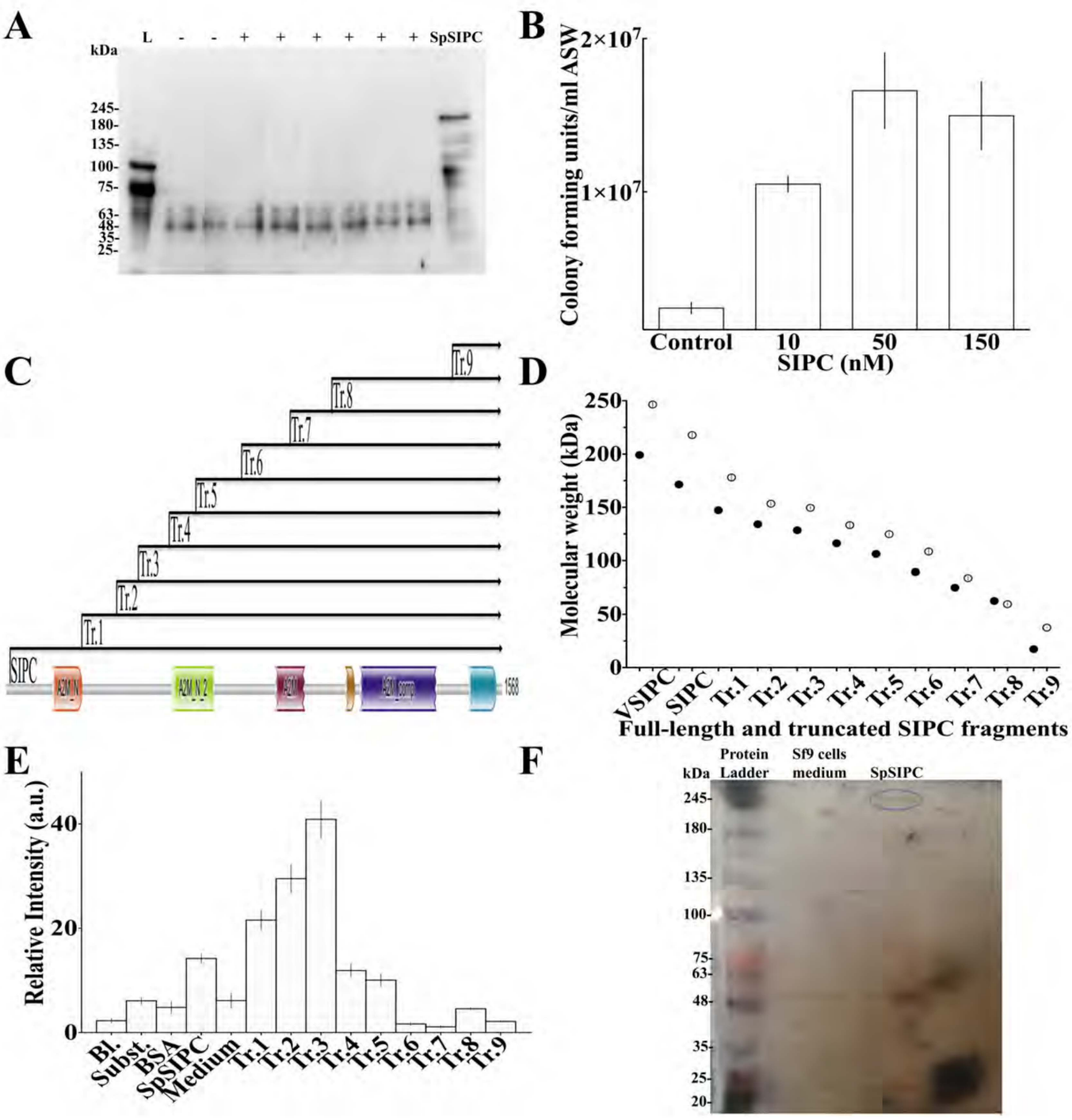

Fig. S4

A: Native SIPC is not an ASW-diffusible moiety. Trichloroacetic acid (TCA)-precipitated proteins from wells of tissue culture plates used in behavioural bioassays were resolved in 8% SDS-PAGE gels and Western blots were carried out with a mouse anti-SIPC monoclonal antibody as described in the Methods section. For each lane the (+) symbol indicates wells that contained 10 cyprids/2 ml of ASW and the (-) symbol indicates wells that contained no cyprids of *A. amphitrite*. An unknown protein of ~ 48 kDa immunoreacts with the mouse anti-SIPC monoclonal antibody. The supernatant of Sf9 cells expressing SpSIPC (20 µl) was used as a positive control. A typical image of several blots (n=3) is shown.

B: Bacterial growth in ASW used in behavioural bioassays in the presence or absence of recombinant SIPC. The spread plate counting technique was used to identify bacterial growth in behavioural bioassays in the absence or presence of the indicated concentrations of SIPC. Results are expressed as colony forming units/ml of ASW. ANOVA analysis showed that bacteria introduced by cyprids do not affect the behavioural response of the cyprids to SIPC ($p>0.05$). The cumulative results from 2 independent experiments with 3 replicates each is shown.

C: Graphical representation of the pfam domains present in SIPC and the various truncated SIPC fragments. Pfam domains were visualized via HMMER at https://www.ebi.ac.uk/Tools/hmmer/search/phmmer and the length of the truncated SIPC fragments in shown.

D: Calculated (closed circles) and apparent (open circles) MW of SIPC and its truncated fragments quantified by western blots. Western blots were probed with an anti-myc antibody and the cumulative data from several blots (n=3-7) is shown.

E: Adhesive properties of recombinant SIPC and truncated SIPC fragments to the surface of polystyrene tissue culture plates used in behavioural bioassays. Sterile tissue culture plates were incubated at 25 °C for 24 hrs empty (Bl.) or with 0.3 ml of BSA (10 µM), SpSIPC (50 nM), HEK293 cell culture medium (medium) or purified truncated SIPC fragments (Tr.1 to Tr.9, 50 nM each). Solutions were then aspirated and western blots were carried out on the wells with a Myc-Tag (9B11) mAb (Cell Signaling Technology) as described in the Methods section. In a set of wells (Subst.) only LumiGLO® chemiluminescence reagent was added. A typical image of several experiments (n=3) is shown. Acquired high resolution images were analysed with ImageJ 1.48v (NIH, USA) and data was exported and statistically analysed with GraphPad Prism v.6. Results are expressed as arbitrary units (a.u.) of relative intensity as determined by the ImageJ 1.48v software.

F: Purified recombinant SIPC is a high molecular weight protein. An SDS-PAGE gel was silver stained (see Methods section) and a band (denoted by a circle) was excised and subjected to LC-MS/MS. See ProteomeXchange Consortium dataset identifier PXD006858 and 10.6019/PXD006858 for detailed results.

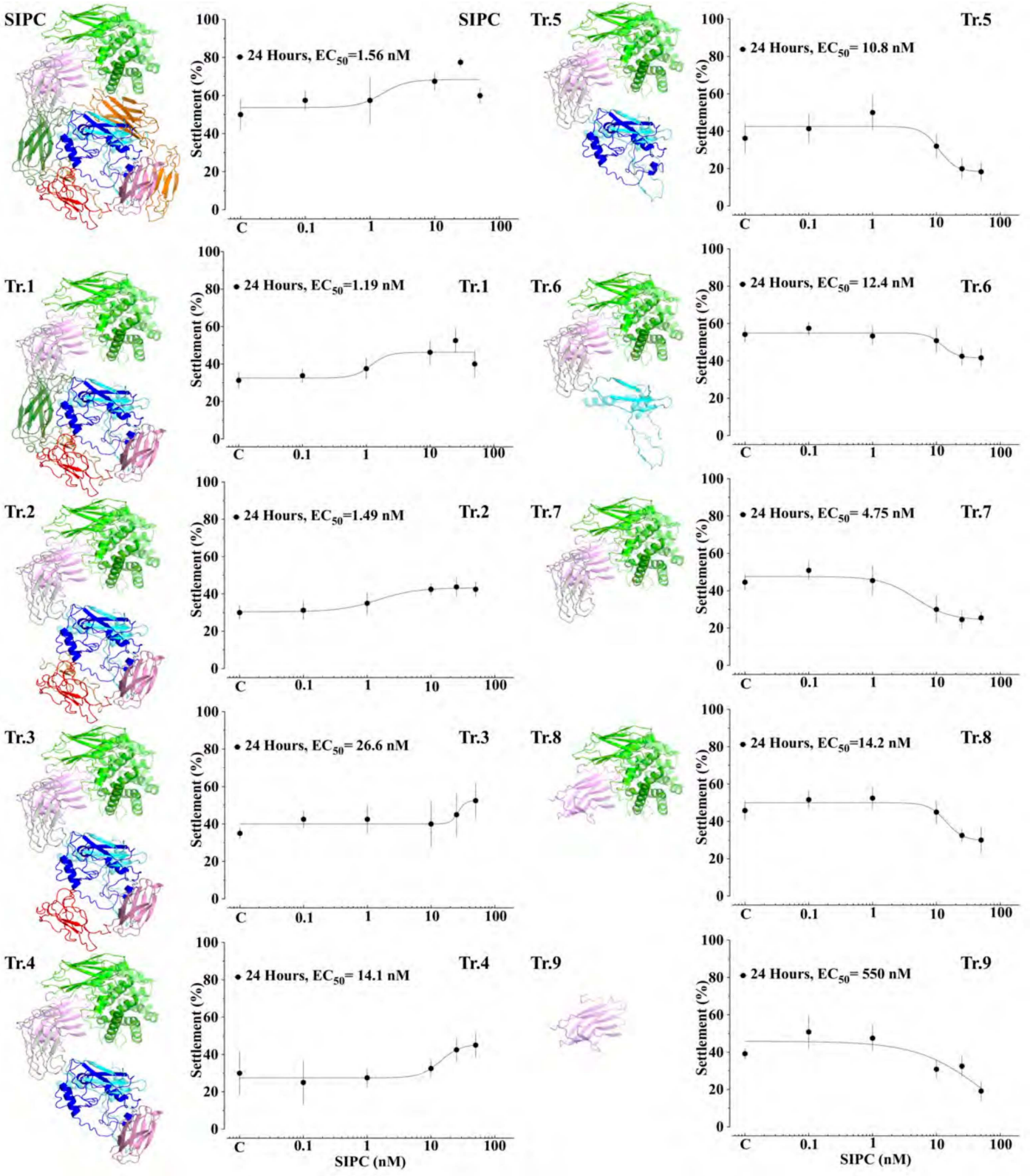

Fig. S5

Recombinant truncated SIPC fragments transduce gregarious settlement preference or settlement avoidance behaviour by *A. amphitrite* cyprids. Values are expressed as the percentage of animals that settled from the total number of animals placed in a well of a 24-well tissue culture plate. The cumulative results from 2 independent experiments with 4 replicates each is shown. On the left of each graph, the crystal structure simulation of SIPC or its truncated fragments is shown for visual clarity.

Table S1: Alignment of the open reading frame from 3 cDNA sequences of A. amphitrite SIPC using the multiple sequence alignment software MAFFT at http://www.ebi.ac.uk/Tools/msa/mafft/. Bold letters in red indicate the differences between the nucleotide sequences.

CLUSTAL format alignment by MAFFT FFT-NS-i (v7.215)

```
Kotsiri et al SIPC      ATGGGTCGGCCATCGTGGTTCTACTGGTCGCTTTGGCGACGGCAAGCGCCGTCAAGGTC
Dreanno et al SIPC      ATGGGTGGTCCCGTCGTCGTTCTACTGGTCGCCTTGGCGACGGCAAGCGCCGTCAAGGTC
Zhang et al SIPC        ATGGGAT----------------------------------------CCGTCAAGGTC
                        *****                                           ***********

Kotsiri et al SIPC      CCCGAAAGCGGGTACCTGTTCACGGCACCCAAAGTACTTCAGGCTGGCACAGATGAACGT
Dreanno et al SIPC      CCCGAAAGCGGGTACCTGTTCACGGCACCCAAAGTACTCCAGGCTGGCACAGATGAACGT
Zhang et al SIPC        CCCGAAAGCGGGTACCTGTTCACGGCACCCAAAGTACTCCAGGCTGGCACAGATGAACGT
                        ************************************.********************

Kotsiri et al SIPC      GCCTGCCTCAGTCTGTTCAACCTACCCGGGCCGAATCGCGCTCTCAAGCTGAAGTTCTAC
Dreanno et al SIPC      GCCTGCCTCAGTCTGTTCAACCTACCCGGACCGAACCGTGCGCTCAAGCTGAAGTTCTAC
Zhang et al SIPC        GCCTGCCTCAGTCTGTTCAACCTACCCGGACCGAACCGTGCGCTCAAGCTGAAGTTCTAC
                        ******************************.*****.**.** ******************

Kotsiri et al SIPC      GAGCGCGACGTTCCAAGCAGTCTGTCGACCACGCTAGATAAGAGCGATTTCCTGCTGTTT
Dreanno et al SIPC      GAGCGCGACGTTCCAAGCAGTCTGTCGACCACGCTAGATAAGAGCGATTTCCTGCTGTTT
Zhang et al SIPC        GAGCGCGACGTTCCAAGCAGTCTGTCGACCACGCTAGATAAGAGCGATTTCCTGCTGTTT
                        ************************************************************

Kotsiri et al SIPC      GAGACAAACACTGCAGTGCCAGACAGCGTAGCAGAGAATGGAGAGTACTGTTTCGACATC
Dreanno et al SIPC      GAGACAAACACAGCAGTGCCGGACAGCGTAGCAGAGAATGGAGAGTACTGCTTCGACATT
Zhang et al SIPC        GAGACAAACACAGCAGTGCCGGACAGCGTAGCAGAGAATGGAGAGTACTGCTTCGACATT
                        *********** ********.******************************.********.

Kotsiri et al SIPC      ACCATTCCATCCAAGGTGGTGGCCCGCAGTGCTGACATGCACATGGAGCTGACGGCTGGC
Dreanno et al SIPC      ACCATTCCATCCAAGGTGGTGGCCCGCAGTGCTGACATGCACATGGAGCTGACGGCTGGT
Zhang et al SIPC        ACCATTCCATCCAAGGTGGTGGCCCGCAGTGCTGACATGCACATGGAGCTGACGGCTGGT
                        ************************************************************.

Kotsiri et al SIPC      GAAGGCGTCTGGAAGGAGGAGTCCGTGGTGACTCTGAAATCGGAGACGTTCCTGACGCTG
Dreanno et al SIPC      GAAGGCGTCTGGAAGGAGGAGTCTGTGGTGACTCTGAAATCGGAGACGTTCCTGACGCTG
Zhang et al SIPC        GAAGGCGTCTGGAAGGAGGAGTCTGTGGTGACTCTGAAATCGGAGACGTTCCTGACGCTG
                        ************************.***********************************

Kotsiri et al SIPC      GTTCAGACGGACAAGTCCAAGTACCAGCCTGGTCAGAAGGTGCTCTTCAGAGTGGTTACG
Dreanno et al SIPC      GTTCAGACGGACAAGTCCAAGTACCAGCCTGGTCAGAAGGTGCTCTTCAGAGTGGTTACC
Zhang et al SIPC        GTTCAGACGGACAAGTCCAAGTACCAGCCTGGTCAGAAGGTGCTCTTCAGAGTGGTTACC
                        ************************************************************

Kotsiri et al SIPC      CTCAGTCACGACCTGACCGCCCTCAACAATGACCTGAATGAGGTGTGGGTCACCACTCCG
Dreanno et al SIPC      CTTAGTCACGACCTGACCGCCCTCAACAATGACCTGAATGAGGTGTGGATCACCACTCCG
Zhang et al SIPC        CTTAGTCACGACCTGACCGCCCTCAACAATGACCTGAATGAGGTGTGGATCACCACTCCG
                        **.*****************************************.***********

Kotsiri et al SIPC      GACAACGTCCGCGTGGCCCAGTGGAAGAATGTCAAAACCAACACTGGCATGGTGCAGCTG
Dreanno et al SIPC      GACAACATCCGCGTGGCCCAGTGGAAGAACGTGAAAACCAACACTGGCATGGTGCAGCTG
Zhang et al SIPC        GACAACATCCGCGTGGCCCAGTGGAAGAACGTGAAAACCAACACTGGCATGGTGCAGCTG
                        ******.**********************.** **************************

Kotsiri et al SIPC      GAACTGCAGCTGACCGAGGAGCCACCTCTGGGCTCGTGGACAATTCACGTGCGGACTACC
Dreanno et al SIPC      GAACTGCAGCTGACCGAGGAGCCACCTCTGGGCTCGTGGACAATTCACGTGCTGACTACC
Zhang et al SIPC        GAACTGCAGCTGACCGAGGAGCCACCTCTGGGCTCGTGGACAATTCACGTGCTGACTACC
                        ************************************************** *******
```

```
Kotsiri et al SIPC      CAGGACACCTACACCAAGCGATTTACGGTGGAGGAGTATGTTCTGCCAACATTTGAGCTG
Dreanno et al SIPC      CAGGACACCTACACCAAGCGATTTACGGTGGAGGAGTATGTTCTGCCAACATTTGAGCTG
Zhang et al SIPC        CAGGACACCTACACCAAGCGATTTACGGTGGAGGAGTATGTTCTGCCAACATTTGAGCTG
                        ************************************************************

Kotsiri et al SIPC      GAGATCGAGGCACCCGAGTCTCTTGAGAGCAACGAGAAGACTGTCACCGTTAAAGTTTGC
Dreanno et al SIPC      GAGATCGAGGCACCCGAGTCTCTTGAGAGCAACGAGAAGACTGTCACCGTTAAAGTTTGC
Zhang et al SIPC        GAGATCGAGGCACCCGAGTCTCTTGAGAGCAACGAGAAGACTGTCACCGTTAAAGTTTGC
                        ************************************************************

Kotsiri et al SIPC      GCAAAGTACACTTTCGGAAAGCCGCTCATTGCTGCTAACGTGAGCATCAATGCCACAGCT
Dreanno et al SIPC      GCAAAGTACACTTTCGGAAAGCCGCTTATTGCTGCCAACGTAAGCATCAACGCCACAGCC
Zhang et al SIPC        GCAAAGTACACTTTCGGAAAGCCGCTTATTGCTGCCAACGTAAGCATCAACGCCACAGCC
                        ***************************.********.*****.********.********.

Kotsiri et al SIPC      AGAGGCATTGGAAGCTGGCAATACAACAATAATAAGGATCTGCTGAGGAATATCTCAGAC
Dreanno et al SIPC      AGAGGCATTGGAAGCTGGCAATACAACAATAATCCGGATCTGCTGAGGAATATCTCAGAC
Zhang et al SIPC        AGAGGCATTGGAAGCTGGCAATACAACAATAATCCGGATCTGCTGAGGAATATCTCAGAC
                        ********************************  ************************

Kotsiri et al SIPC      TATCAGTTCAGCGACGAGCAGGGATGTGCCATCTTTGACCTGGTGGTCAGCAAGATCGGC
Dreanno et al SIPC      TATCAATTCAGCGACGAGCAGGGATGTGCCATCTTTGACCTGGTGGTCAGCAAGATCGGC
Zhang et al SIPC        TATCAATTCAGCGACGAGCAGGGATGTGCCATCTTTGACCTGGTGGTCAGCAAGATCGGC
                        *****.************************************************

Kotsiri et al SIPC      ATCGGTCATCGCAATATTGGAGGCGGAAATACGGTCATCATCACCATCGATGTGGAAGAG
Dreanno et al SIPC      ATCGGTCATCGTAATATTGGAGGCGGAAATACGGTGATCATCACCATCGATGTGGAAGAG
Zhang et al SIPC        ATCGGTCATCGTAATATTGGAGGCGGAAATACGGTGATCATCACCATCGATGTGGAAGAG
                        ***********.************************ ************************

Kotsiri et al SIPC      CAGGGTACAGGACTCCGGCAGGTTGAGGTCAAGGAGGTCAGTCAAGCGTATTCCTTCATC
Dreanno et al SIPC      CAGGGTACAGGACTCCGGCAGGTGGAGGTCAAGGAGGTCAGTCAAGCGTATTCCTTCATC
Zhang et al SIPC        CAGGGTACAGGACTCCGGCAGGTGGAGGTCAAGGAGGTCAGTCAAGCGTATTCCTTCATC
                        ********************** ***********************************

Kotsiri et al SIPC      AACCTCAGGCAGAGCGACAACGCGCAGAAGTTCCTGAAGCCGAAGCTGCCCTTCTACGGA
Dreanno et al SIPC      AACCTCAGGCAGAGCGACAACGCGCAGAAGTTCCTGAAGCCGAAGCTGCCCTTCTACGGA
Zhang et al SIPC        AACCTCAGGCAGAGCGACAACGCGCAGAAGTTCCTGAAGCCGAAGCTGCCCTTCTACGGA
                        ************************************************************

Kotsiri et al SIPC      GAGTACACGCTGAGCATGAGAGATGGTAAGGCTGCGAAGAATGAAATCGTCAAGGTCTGC
Dreanno et al SIPC      GAGTACACGCTGAGTATGAGAGATGGTAAGGCTGCGAAGAATGAAATCGTCAAGGTCTGC
Zhang et al SIPC        GAGTACACGCTGAGTATGAGAGATGGTAAGGCTGCGAAGAATGAAATCGTCAAGGTCTGC
                        **************.*********************************************

Kotsiri et al SIPC      TACACAGCCAAGTACAAGGAGAGGGTAATCAGTGACGAGAAGAAGCCAACTCCCGATGAC
Dreanno et al SIPC      TACACAGCCAAGTACAAGGAGAGGGTCATCAGTGACGAGAAGAAGCCAACTCCCGATGAC
Zhang et al SIPC        TACACAGCCAAGTACAAGGAGAGGGTCATCAGTGACGAGAAGAAGCCAACTCCCGATGAC
                        *************************** ********************************

Kotsiri et al SIPC      CCCGTTTATAGCACACACAAGAAGTACGAGTCGCACGTGAAGACCGAGTTCGGCTACACG
Dreanno et al SIPC      CCCGTTTATAGCACACACAAGAAGTACGAGTCGCACGTGAAGACCGAGTTCGGCTACACG
Zhang et al SIPC        CCCGTTTATAGCACACACAAGAAGTACGAGTCGCACGTGAAGACCGAGTTCGGCTACACG
                        ************************************************************

Kotsiri et al SIPC      CCGTTCTTCTGGGAGACTTCGGAGCCCAATCGTCGCACCACCGGTGGCGAATGTCGCGAG
Dreanno et al SIPC      CCGTTCTTCTGGGAGACTTCGGAGCCCAATCGCCGCACCACCGGTGGCGAATGTCGCGAG
Zhang et al SIPC        CCGCTCTTCTGGGAGACTTCGGAGCCCAATCGTCGCACCACTGGTGGCGAATGTCGCGAG
                        ***.*************************.********.*****************

Kotsiri et al SIPC      TATAAAACAGACGAAAACGGTCGTATCGTCTACTATATCCCACCACAGGCTGAGGACATC
```

```
Dreanno et al SIPC          TATAAAACAGACGAAAACGGTCGTATCGTCTACTATATCCCACCACAGGCTGAGGACATC
Zhang et al SIPC            TATAAAACAGACGAAAACGGTCGTATCGTCTACTATATCCCACCACAGGCTGAGGACATC
                            ************************************************************

Kotsiri et al SIPC          GACTCCATTGACATCTCGACTTCTACCAGCGTCGGTGGCGATTCTGACAGCTCGCACTCC
Dreanno et al SIPC          GACTCCATTGACATCTCGACTTCTACCAGCGTCGGTGGCGATTCTGACAGCTCGCACTCC
Zhang et al SIPC            GACTCCATTGACATCTCGACTTCTACCAGCGTCGGTGGCGATTCTGACAGCTCGCACTCC
                            ************************************************************

Kotsiri et al SIPC          ACGCTAACGGCGTTCTTCTCGCCTTCCCACTCGTACCTAAGCATCGACACGCACGAACTT
Dreanno et al SIPC          ACGTTGACGGCGTTCTTCTCGCCTTCCCACTCGTACCTAAGCATCGACGCGCACGAACTT
Zhang et al SIPC            ACGCTGACGGCCTTCTTCTCGCCTTCCCACTCGTACCTGAGCATCGACGCGCACGAACTT
                            ***.*.***** ************************.*********.***********

Kotsiri et al SIPC          CCCGAACAGCTGCCCTGCTCCGGAGATGTCACTGTGAAACTGCTCTCAACTGAGGAGGGA
Dreanno et al SIPC          CCCGAACAGCTGCCCTGCTCCGGAGATGTCACTGTGAAACTGCTCTCAACTGAGGAGGGA
Zhang et al SIPC            CCTGAACAGCTGCCCTGCTCCGGAGACGTCACTGTGAAACTGCTCTCAACTGAGGAGGGA
                            **.***********************.**********************************

Kotsiri et al SIPC          CCGGTGCCAGCCATGGTTTACAAGATTCTGTCCCGAGGCAAGATCATCAAGGCGGGCAAC
Dreanno et al SIPC          CCGGTGCCAGCCATGGTTTACAAGATTTTGTCCCGAGGCAAGATCATCAAGGCGGGCAAC
Zhang et al SIPC            CCGGTGCCAGCTATGGTTTACAAGATTCTGTCCCGAGGCAAGATCATCAAGGCGGGCAAC
                            ***********.***************.**********************************

Kotsiri et al SIPC          ATGAACACTAACACGCTGACGTTCCCGGTGCTGCCCAAGATGGGCCCCGAGTTCAAGCTG
Dreanno et al SIPC          ATGAACACTAACACGCTGACGTTCCCGGTGCTGCCCAAGATGGGCCCCGAGTTCAAGCTG
Zhang et al SIPC            ATGAACACTAATACGCTGACGTTCCCGGTGCTGCCCAAGATGGGCCCCGAGTTCAAGCTG
                            ***********.************************************************

Kotsiri et al SIPC          CTTGTCTACTACATCAAGGAGAGTGGCGAAGTGGTCAGCGACTCCAGGGTCTTCAAGGTC
Dreanno et al SIPC          CTTGTCTACTACATCAAGGAGAGTGGCGAAGTGGTCAGCGACTCCAGGGTCTTCAAGGTC
Zhang et al SIPC            CTTGTCTACTACATCAAGGAGAGTGGCGAAGTGGTCAGCGACTCCAGGGTCTTCAAGGTC
                            ************************************************************

Kotsiri et al SIPC          GACAAGTGCTTCCCCAACACGGTCCAGGTCAGCTGGGATCAGAAGACGGTCAAGCCGGGA
Dreanno et al SIPC          GACAAGTGCTTCCCCAACACTGTTCAGGTCAGCTGGGATCAGAAGACGGTCAAGCCGGGA
Zhang et al SIPC            GACAAGTGCTTCCCCAACACTGTTCAGGTCAGCTGGGATCAGAAGACGGTCAAGCCGGGA
                            ******************** **.*********************************

Kotsiri et al SIPC          GACTCTGCTTCGTTCACCGTTCGTGCCTCTCCCAACTCCGTCTGTGGAATCAGCGCTGTG
Dreanno et al SIPC          GACTCTGCTTCGTTCACCGTTCGTGCCTCTCCCAACTCCGTCTGTGGAATCAGCGCTGTG
Zhang et al SIPC            GACTCTGCTTCGTTCACCGTTCGTGCCTCTCCCAACTCCGTCTGTGGAATCAGCGCTGTG
                            ************************************************************

Kotsiri et al SIPC          GACAAGAGCACCGAACTTCTGGGCACGTCCAACCAGATCACACTGGATACTGTGTTCAGC
Dreanno et al SIPC          GACAAGAGCACCGAACTTCTGGGCACGTCCAACCAGATCACACTGGATACCGTGTTCAGC
Zhang et al SIPC            GACAAGAGCACCGAACTTCTGGGCACGTCCAACCAGATCACACTGGATACCGTGTTCAGC
                            ************************************************.*********

Kotsiri et al SIPC          AAGCTGCAGCAGTTCATCATTAATAGCTTCGAGAGCCCCAACCAGGTCCGCAGCGATGGC
Dreanno et al SIPC          AAGCTGCAGCAGTTCATCATTAATAGCTTCGAGAGCCCCAACCAGGTCCGCAGCGATGGC
Zhang et al SIPC            AAGCTGCAGCAGTTCATCATTAATAGCTTCGAGAGCCCCAACCAGGTCCGCAGCGATGGC
                            ************************************************************

Kotsiri et al SIPC          GATTACTGCAGGGAGCTGCAGCTTTCACTGGTAGACACTCTGCGAAGCGGTGGCGCCACG
Dreanno et al SIPC          GATTACTGCAGGGAACTGCAGCTTTCACTGGTAGACACTCTGCGAAGCGGTGGCGCCACG
Zhang et al SIPC            GATTACTGCAGGGAACTGCAGCTTTCACTGGTAGACACTCTGCGAAGCGGTGGCGCCACG
                            **************.*********************************************

Kotsiri et al SIPC          GCAGCTGAGCTTACTGGTCAGTCCACCCCGGAGGGCACGCCAGAGTCTGAAACCAGCGGA
Dreanno et al SIPC          GTGGCCGAGCTTACTGGTCAGTCCACCCCGGAGGGCACGCCAGAGTCTGAAACCAGCGGA
```

```
Zhang et al SIPC          GTGGCCGAGCTTACTGGTCAGTCCACCCCGGAGGGCACGCCAGAGTCTGAAACCAGCGGA
                          *..**.**************************************************

Kotsiri et al SIPC        GCTGCCCAGTCGTCGCTCTTCATCCCTCCACCGACACGGTCTCAGCGATTCCGCACAGAT
Dreanno et al SIPC        GCTGCCCATTCGTCGCTCTTCATCCCTCCACCGACACGGTCTCAGCGATTCCGCACGGAT
Zhang et al SIPC          GCTGCCCATTCGTCGCTCTTCATCCCTCCACCGACACGGTCTCAGCGATTCCGCACGGAT
                          ******** ***********************************************.***

Kotsiri et al SIPC        CGCGATGACGCCATCAAGCCTTTCGATGAAGCCGGTTTCCTCGTTCTCTCCAACCTGGCC
Dreanno et al SIPC        CGCGAAGACGCCATCAAGCCTTTCGATGAAGCCGGTTTCCTCGTTCTCTCCAACCTGGCC
Zhang et al SIPC          CGCGAAGACGCCATCAAGCCTTTCGATGAAGCCGGTTTCCTCGTTCTCTCCAACCTGGCC
                          ***** ***************************************************

Kotsiri et al SIPC        CTGGAGACTCGGCCCTGCTACAAGAGAGTTCAGGCCAAGGAACTGCCGGAGCTCACCGAG
Dreanno et al SIPC        CTGGAGACTCGGCCCTGCTACAAGAGAGTTGAGGCCAAGGAACTGCCGGAGCTCACCGAG
Zhang et al SIPC          CTGGAGACTCGGCCCTGCTACAAGAGAGTTCAGGCCAAGGAACTGCCGGAGCTCACCGAG
                          ***************************** ****************************

Kotsiri et al SIPC        GATAAAATCCAGGCGTCTAGGGGCGGCGAAGAAGAGCTTCTCGACGACCTGGATTCGCCG
Dreanno et al SIPC        GATAAAATCCAGGCGTCTAGGGACGGCGAAGAAGAGCTTCTCGACGACCTGGATTCGCCG
Zhang et al SIPC          GATAAAATCCAGGCGTCTAGGGACGGCGAAGAAGAGCTTCTCGACGACCTGGATTCGCCG
                          **********************.************************************

Kotsiri et al SIPC        GTACCAGCGTTGGCTTTCTCCAAAGAGTCAGCGGACGCGTCACGGTTCGCCGCAGAGGGT
Dreanno et al SIPC        GTACCGGCGTTGGCTTTCTCCAAAGAGTCAGCGGACGCGTCACGGTTCGCCGCAGAGGGT
Zhang et al SIPC          GTACCGGCGTTGGCTTTCTCCAAAGAGTCAGCGGACGCGTCACGGTTCGCCGCAGAGGGT
                          *****.***************************************************

Kotsiri et al SIPC        GGTGTCTCAGGTGGAGGAGGCGCGGCGCCGCCTCAAGAGGATCAGGTGCGCGACTTCTTC
Dreanno et al SIPC        GGTGTCTCAGGTGGGGGTGGAGCAGCGCCGCCTCAAGAGGATCAGGTGCGCGACTTCTTC
Zhang et al SIPC          GGTGTCTCAGGTGGGGGTGGAGCAGCGCCGCCTCAAGAGGATCAGGTGCGCGACTTCTTC
                          **************.** ** **.************************************

Kotsiri et al SIPC        CCAGAGGCCTTCCTCTTCAGCATCGAGACACTAGATGCCGAGGGTGTGAAGACTGTCACC
Dreanno et al SIPC        CCAGAGGCCTTCCTCTTCAGCATCGAGACACTAGATGCGGAGGGTGTGAAGACTGTCACG
Zhang et al SIPC          CCAGAGGCCTTCCCCTTCAGCATCGAGACACTAGATGCGGAGGGTGTGAAGACTGTCACC
                          *************.********************** ********************

Kotsiri et al SIPC        TCCGAGATGCCGGACACCATCACCTCCTGGGTCGGATCAGCCATCTGCACCAACAGCAAG
Dreanno et al SIPC        TCCGAGATGCCAGACACCATCACCTCCTGGGTCGGATCAGCCATCTGCACCAACAGCAAG
Zhang et al SIPC          TCCGAGATGCCGGACACCATCACCTCCTGGGTCGGATCAGCCATCTGCACCAACAGCAAG
                          ***********.************************************************

Kotsiri et al SIPC        GACGGATTCGGCATCTCTAACAAGACCAGCATCACCACCTTCAAGCCGTTCTTCACAGAG
Dreanno et al SIPC        GACGGATTCGGCATCTCTAACAAGACCAGCATCACCACCTTCAAGCCGTTCTTCACCGAA
Zhang et al SIPC          GACGGATTCGGCATCTCTAACAAGACCAGCATCACCACCTTCAAGCCGTTCTTCACAGAG
                          ************************************************************ **.

Kotsiri et al SIPC        GTGTCTCTGCCGTACTCGATGAAGCGTGGCGAGATCCTGAGCATGTCCGTCAGCGTCTTC
Dreanno et al SIPC        GTGTCTCTGCCGTACTCGATGAAGCGTGGTGAGATCCTGAGCATGTCTGTCAGCGTCTTC
Zhang et al SIPC          GTGTCTCTGCCGTACTCGATGAAGCGTGGCGAGATCCTGAGCATGTCCGTCAGCGTCTTC
                          ****************************.****************.************

Kotsiri et al SIPC        AACTTCCTCGACTCGAGTCTCTCGGTTTACCTTGAAGTTGGTGCCTCTGACCAGTATGAG
Dreanno et al SIPC        AACTTCCTCGACTCCAGTCTCTCGGTTTACCTCGAAGTTGGTGCCTCTGACCAGTATGAG
Zhang et al SIPC          AACTTCCTCGACTCGAGTCTCTCGGTTTACCCTGAAGTTGGTGCCTCTGACCAGTATGAG
                          ************** ****************..**************************

Kotsiri et al SIPC        ATTAGCGGAGAGGTCGCCATGGGTCTCTGCATTGCCGCCGGTCGCACTGAAGTCAAGTCG
Dreanno et al SIPC        ATTAGCGGAGAGGTCGCCATGGGTCTCTGCATTGCCGCCGGTCGCACTGAGGTCAGGTCG
Zhang et al SIPC          ATTAGCGGAGAGGTCGCCATGGGTCTCCGCATTGCCGCGGGTCGCACTGAGGTCAGGTCG
```

Journal of Experimental Biology • Supplementary information

```
                          **************************.********** ***********.****.****

Kotsiri et al SIPC        TTCCCGGTCAACTTCCTGGGTCTGGGCGAGGTCAACATCACAGTCACGGCCAGGGCACAG
Dreanno et al SIPC        TTCCCGGTCAACTTCCTGGGTCTGGGCGAGGTCAACATCACAGTCACGGCCAGGGCACAG
Zhang et al SIPC          TTCCCGGTCAACTTCCTGGGTCTGGGCGAGGTCAACATCACAGTCACGGCCAGGGCACAG
                          ************************************************************

Kotsiri et al SIPC        GACGGATACTGCGACGAAGGCAACACCATCGCGCCGGGCAGTGACACCGTCATCCGACCA
Dreanno et al SIPC        GACGGATACTGCGACGAAGGCAACACCATCGCGCCGGGCAGTGACACCGTCATCCGACCA
Zhang et al SIPC          GACGGATACTGCGACGAAGGCAACACCATCGCGCCGGGCAGTGACACCGTCATCCGACCA
                          ************************************************************

Kotsiri et al SIPC        ATCGTCGTCAAACCAGAGGGATTCCCTCAAGAGGTGACCCACTCTCGCTTCATCTGTCTT
Dreanno et al SIPC        ATCGTCGTCAAACCAGAGGGATTCCCTCAAGAAGTGACCCACTCTCGCTTCATCTGTCTC
Zhang et al SIPC          ATCGTCGTGAAACCAGAGGGATTCCCTCAAGAAGTGACCCACTCTCGTTTCATCTGCCTC
                          ******** ***********************.**************.********.**.

Kotsiri et al SIPC        GACAAGGATGACGACCACCACACGGAGACTGTGAACCTGCCGGTGCCGGAAGGCCTGGTG
Dreanno et al SIPC        GACAAGGATGACGACCACCACACGGAGACGGTGAACCTGCCGGTGCCGGAAGGCCTGGTG
Zhang et al SIPC          GACAAGGATGACGACCACCACACGGAGACGGTGAACCTGCCGGTGCCGGAAGGTCTGGTA
                          ***************************** ***********************.******.

Kotsiri et al SIPC        CCCGACTCTCAGCGCGCCTACTTCTCCGTCATCGGAGATCTTCTGGGACCGACCTTCCAG
Dreanno et al SIPC        CCCGACTCTCAGCGCGCCTACTTCTCCGTCATCGGAGATCTTTTGGGACAGACCTTCCAG
Zhang et al SIPC          CCCGACTCTCAGCGCGCCTACTTCTCCGTCATCGGAGATCTTCTGGGACAGACCTTCCAG
                          ******************************************.****** **********

Kotsiri et al SIPC        GGTCTTGAGGGAGGTCTCATTAAGTCGCCTACCGGCGCCGGTGAGCCCAACATGATCACT
Dreanno et al SIPC        GGTCTGGAGGGAGGTCTCATTAAGTCGCCTACCGGCGCCGGTGAGCCCAACATGATCACT
Zhang et al SIPC          GGTCTTGAGGGAGGTCTCATTAAGTCGCCTACCGGCGCCGGTGAGCCCAACATGATCACT
                          ***** ************************************************************

Kotsiri et al SIPC        CTGGTGCCCAATATCTACATCCGTCGCTACCTGGAGACAACTGGTCAGCTCAACGAGCGT
Dreanno et al SIPC        CTGGTGCCCAATATCTACATCCGTCGCTACCTGGAGACAACTGGTCAGCTCAACGAGCGT
Zhang et al SIPC          CTGGTGCCCAATATCTACATCCGTCGCTACCTGGAGACAACTGGTCAGCTCAACGAGCGT
                          ************************************************************

Kotsiri et al SIPC        CAGAGACGGCAGCTAGAACACAACATGAAGAGCGGTTACCAGCGCCAGTTGCGCTTCAGG
Dreanno et al SIPC        CAGAGACGACAGCTAGAGCACAACATGAAGAGCGGCTACCAGCGCCAGTTGCGCTTCAGG
Zhang et al SIPC          CAGAGGCGACAGCTAGAGCACAACATGAAGAGCGGCTACCAGCGCCAGTTGCGCTTCAGG
                          *****.**.********.*****************.************************

Kotsiri et al SIPC        CGGTACGATGGCTCGTTCTCGTCGTACGGAAATGAGGACCCTCAGGGCTCCATGTGGCTC
Dreanno et al SIPC        CGGTACGATGGCTCGTTCTCGTCGTACGGAAATGAGGACCCTCAGGGCTCCATGTGGCTC
Zhang et al SIPC          CGGTACGATGGTCCGTTCTCGTCGTACGGAAATGAGGACCCTCAGGGCTCCATGTGGCTC
                          ***********..************************************************

Kotsiri et al SIPC        ACTGCCTTCGTTGTCAAGGCCTTCCGCGAGGCGTCCGAGTACATCGAAATCGATGAGACT
Dreanno et al SIPC        ACTGCCTTCGTTGTCAAGGCCTTCCGCGAGGCGTCCGAGTACATCGAAATCGATGAGACT
Zhang et al SIPC          ACCGCCTTTGTTGTCAAGGCCTTCCGCGAGGCGTCCGAGTACATCGAAATCGATGAGACT
                          **.*****.************************************************

Kotsiri et al SIPC        ATTATCAACAAGGCTAAGGACTGGATTCTGAAGAAACAGAACACTACTGGCTGTTTCCCG
Dreanno et al SIPC        ATTATCAACAAGGCTAAGGACTGGATTCTGAAGAAACAGAACACTACTGGCTGTTTCCCG
Zhang et al SIPC          ATTATTAACAAGGCTAAGGACTGGATTCTGAAGAAACAGAACACTACTGGCTGTTTCCCG
                          *****.************************************************

Kotsiri et al SIPC        AGGTTCGGCGAGCTTATTCACAAGGAGCTGAAGGGTGGCACCGAGCGAGGCGGTGAAGCG
Dreanno et al SIPC        AGGTTCGGCGAGCTTATTCACAAGGAGCTGAAGGGTGGCACCGAGCGAGGCGGTGAAGCG
Zhang et al SIPC          AGGTTCGGCGAGCTTATTCACAAGGAGCTGAAGGGTGGCACCGAGCGAGGCGGTGAAGCG
                          ************************************************************
```

```
Kotsiri et al SIPC      GCCCTCACGGCCTTCGTCATGTTGGCTCTGAAGGACATCGCAACCACTAATGAGCTGGCC
Dreanno et al SIPC      GCCCTCACGGCCTTCGTCATGTTGGCTCTGAAGGACATCGCAACCACTAATGAGCTGGCC
Zhang et al SIPC        GCCCTCACGGCCTTCGTCATGCTGGCTCTGAAGGACATCGCAACCACTAATGAGCTGGCC
                        **********************.************************************

Kotsiri et al SIPC      AACGGCTTCGCCTGCCTAGAGGACGGTCTTCTGCTCCCCAACAAGACCCTGTATTCGGAG
Dreanno et al SIPC      AACGGCTTCGCCTGCCTAGAGGACGGTCTTCTGCTCCCCAACAAGACCCTGTATTCGGAG
Zhang et al SIPC        AACGGTTTCGCCTGTCTAGAGGACGGTCTTCTGCTCCCCAACAAGACCCTGTATTCGGAG
                        *****.********.********************************************

Kotsiri et al SIPC      ATTCTTTTGGCGTACACATACCTGAATATGGGCCAAGATGTCAAGGGGGAGAGGCTGGTG
Dreanno et al SIPC      ATTCTTTTGGCGTACACATACCTGAATATGGGCCAAGATGTCAAGGGGGAGAGGCTGGTG
Zhang et al SIPC        ATTCTTTTGGCGTACACATACATGAATATGGGCCAAGATGTCAAGGGAGAGAGGCTGGTG
                        ********************* ************************.************

Kotsiri et al SIPC      AACAAGCTCATGTCGAAGGCCAAACGCGAAGGAGATGACATCCTCTACTGGGAGGGCGAC
Dreanno et al SIPC      AACAAGCTCATGTCGAAGGCCAAACGCGAAGGAGATGACATCCTCTACTGGGAGGGCGAC
Zhang et al SIPC        AACAAGCTCATGTCGAAGGCCAAACGCGAGGGAGATGACATCCTCTACTGGGAGGGCGAC
                        ****************************.****************************

Kotsiri et al SIPC      CGTGATTCTCTATTCGGTGGAAGCCGAGCCGTGGACGTGGAGATGACTGCCTACATGGCC
Dreanno et al SIPC      CGCGATTCTCTCTTTGGTGGAAGCCGAGCCGTGGATGTCGAGATGACTGCCTACATGGCT
Zhang et al SIPC        CGCGATTCTCTCTTCGGTGGAAGCCGGCCGTGGACGTGGAGATGACTGCCTACATGGCT
                        **.******** **.***********.********.** ********************.

Kotsiri et al SIPC      CTCTCGCTGATGCACATCTCTGGAAAGGGCAACATGGAGGAGGCAGCGCGTGCCATTCGC
Dreanno et al SIPC      CTCTCGCTGATGCACATCTCCGGAAAGGGCAATATGGAGGAGGCAGCGCGCGCCATTCGC
Zhang et al SIPC        CTCTCGCTGATGCACATCTCTGGAAAGGGTAATATGGAGGAGGCAGCGCGCGCCATTCGC
                        ********************.********.**.*****************.*********

Kotsiri et al SIPC      TGGATCAACACTCAGAGGAACAGCAACGGAGGCTTCAAATCCACCCAGGACACCATTGTT
Dreanno et al SIPC      TGGATCAACACCCAGAGGAACAGCAACGGGGGTTTCAAATCCACTCAGGACACCATTGTT
Zhang et al SIPC        TGGATCAACACGCAGAGGAACAGCAACGGAGGCTTCAAATCCACTCAGGACACCATTGTT
                        *********** *****************.**.***********.***************

Kotsiri et al SIPC      GCTGTGGAAGCTCTGTCAGAGTTCGCATCTCGCACGTTCGCCTCTGATCTGGCCACGAGT
Dreanno et al SIPC      GCTGTGGAAGCCCTGTCAGAGTTCGCGTCTCGCACGTTCGCCTCTGATCTGGCCACGAGT
Zhang et al SIPC        GCTGTGGAAGCTCTGTCAGAGTTCGCGTCTCGCACGTTCGCCTCTGATCTGGCCACGAGT
                        ***********.**************.*********************************

Kotsiri et al SIPC      GTGTCTGTGACTGCTGGAGGAGAGACTGTTCAGCGGATGGTGGATGGAGACAACAGACTG
Dreanno et al SIPC      GTGTCTGTAACTGCTGGAGGAGAGACTGTTCAGCGGATGGTGGATGGAGACAACAGACTG
Zhang et al SIPC        GTGTCTGTGACTGCTGGAGGAGAGACTGTTCAGCGGATGGTGGATGGAGACAACAGACTG
                        ********.***************************************************

Kotsiri et al SIPC      CTGTATCAGGAGTCCAAGGTGCCAGACCTGACGCTGCCTGGCACCATGAACTTCGATGTC
Dreanno et al SIPC      CTGTATCAGGAGTCCAAGGTGCCAGACCTGACGCTGCCTGGCACCATGAACTTCGATGTC
Zhang et al SIPC        CTGTATCAGGAGTCCAAGGTGCCAGACCTGACGCTGCCTGGCACCATGAACTTCGATGTC
                        ************************************************************

Kotsiri et al SIPC      AGTCCGCCTGGCTGCGTGGTGTACCAGAGTATCTTCCGATTCAGCAGCACTCTAGAGGTG
Dreanno et al SIPC      AGTCCGCCTGGCTGCGTGGTCTACCAGAGTATTTTCCGGTTCAGCAGCACTCTCGAGGTG
Zhang et al SIPC        AGTCCGCCTGGCTGCGTGGTCTACCAGAGTATCTTCCGATTCAGCAGCACTCTAGAGGTG
                        ******************* ***********.*****.************** ******

Kotsiri et al SIPC      CCCGACCCTGCCTTCTCTCTCGGTGTGGCTGCCAAGAAGCGAGGCCGTACTGGCTACGAG
Dreanno et al SIPC      CCCGACCCTGCCTTCTCTCTCGGTGTGGCTGCCAAGAAGCGAGGCCGTACTGGCTACGAG
Zhang et al SIPC        CCCGACCCTGCCTTCCCTCTCGGTGTGGCTGCCAAGAAGCGAGGTCGTACCGGCTACGAG
                        ***************.****************************.*****.*********
```

```
Kotsiri et al SIPC          CTGGAAGTATGCACCAGCTTCCTCCGAAACTCTGGCGCCGTGGACCGCGCCATCCTGGAA
Dreanno et al SIPC          CTGGAAGTATGCACCAGCTTCCTCCGAAACTCTGGCGCCGTGGACCGCGCCATTCTGGAA
Zhang et al SIPC            CTGGAAGTCTGCACCAGCTTCCTCCGAAACTCTGGCGCCGTGGACCGCGCCATTCTGGAA
                            ******** ****************************************.******

Kotsiri et al SIPC          ACGGAACTGCCCTCTGGCTATGTTGCTGTGGACAGCACCCTGAGGGACCTGCGCAGAGGC
Dreanno et al SIPC          ACGGAACTGCCCTCCGGCTATGTTGCTGTGGACAGCACCCTGAGGGACCTGCGCAGAGGC
Zhang et al SIPC            ACGGAACTGCCCTCCGGCTATGTTGCTGTGGACAGCACCCTGAGGGACCTGCGCAGAGGC
                            **************.*********************************************

Kotsiri et al SIPC          TCAGCTGTTCGCAGCTATGAGATCAAGGAAGGTAAAGTGATCTTCACACTGCAAGGAGTG
Dreanno et al SIPC          TCAGCTGTTCGCAGCTATGAGATCAAGGAAGGTAAAGTGATCTTCACACTGCAAGGAGTG
Zhang et al SIPC            TCAGCTGTTCGCAGCTATGAGATCAAGGAAGGTAAAGTGATCTTCACACTGCAAGGAGTG
                            ************************************************************

Kotsiri et al SIPC          GCCGAGGATAAGACCTGTCTGGAGTTCCGCATTATTCAGGAGAACGAAGTTGAGCAGCTG
Dreanno et al SIPC          GCCGAGGATAAGACCTGTCTGGAGTTCCGCATTATTCAGGAGAACGAAGTTGAGCAGCTG
Zhang et al SIPC            GCCGAGGATAAGACCTGTCTGGAGTTCCGCATTATTCAGGAGAACGAAGTTGAGCAGCTG
                            ************************************************************

Kotsiri et al SIPC          AAGCCGTCCATTGTGAAGGTGCACGACTTCTACCGTCCTGAGGAGAGGAACATTCAGGAG
Dreanno et al SIPC          AAGCCGTCCATTGTGAAGGTGCACGACTTCTACCGTCCTGAGGAGAGGAACATTCAGGAG
Zhang et al SIPC            AAGCCGTCCATTGTGAAGGTGCACGACTTCTACCGTCCTGAGGAGAGGAACATTCAGGAG
                            ************************************************************

Kotsiri et al SIPC          TACGAGCTGACTCCCGCTGCTTAG
Dreanno et al SIPC          TACGAGCTGACTCCCGCTGCTTAG
Zhang et al SIPC            TACGAGCTGACTCCCGCTGCTTAG
                            ***********************
```

Table S2: Alignment of the amino acid sequences from 3 A. amphitrite SIPC proteins using the multiple sequence alignment software MAFFT at http://www.ebi.ac.uk/Tools/msa/mafft/. Bold letters in red indicate the differences between the amino acid sequences.
CLUSTAL format alignment by MAFFT FFT-NS-i (v7.215)

```
Kotsiri et al SIPC          MGRAIVVLLVALATASAVKVPESGYLFTAPKVLQAGTDERACLSLFNLPGPNRALKLKFY
Dreanno et al SIPC          MGGPVVVLLVALATASAVKVPESGYLFTAPKVLQAGTDERACLSLFNLPGPNRALKLKFY
Zhang et al SIPC            MGS--------------VKVPESGYLFTAPKVLQAGTDERACLSLFNLPGPNRALKLKFY
                            **                ****************************************

Kotsiri et al SIPC          ERDVPSSLSTTLDKSDFLLFETNTAVPDSVAENGEYCFDITIPSKVVARSADMHMELTAG
Dreanno et al SIPC          ERDVPSSLSTTLDKSDFLLFETNTAVPDSVAENGEYCFDITIPSKVVARSADMHMELTAG
Zhang et al SIPC            ERDVPSSLSTTLDKSDFLLFETNTAVPDSVAENGEYCFDITIPSKVVARSADMHMELTAG
                            ************************************************************

Kotsiri et al SIPC          EGVWKEESVVTLKSETFLTLVQTDKSKYQPGQKVLFRVVTLSHDLTALNNDLNEVWVTTP
Dreanno et al SIPC          EGVWKEESVVTLKSETFLTLVQTDKSKYQPGQKVLFRVVTLSHDLTALNNDLNEVWITTP
Zhang et al SIPC            EGVWKEESVVTLKSETFLTLVQTDKSKYQPGQKVLFRVVTLSHDLTALNNDLNEVWITTP
                            ************************************************************:***

Kotsiri et al SIPC          DNVRVAQWKNVKTNTGMVQLELQLTEEPPLGSWTIHVRTTQDTYTKRFTVEEYVLPTFEL
Dreanno et al SIPC          DNIRVAQWKNVKTNTGMVQLELQLTEEPPLGSWTIHVLTTQDTYTKRFTVEEYVLPTFEL
Zhang et al SIPC            DNIRVAQWKNVKTNTGMVQLELQLTEEPPLGSWTIHVLTTQDTYTKRFTVEEYVLPTFEL
                            **:********************************** **********************

Kotsiri et al SIPC          EIEAPESLESNEKTVTVKVCAKYTFGKPLIAANVSINATARGIGSWQYNNNKDLLRNISD
Dreanno et al SIPC          EIEAPESLESNEKTVTVKVCAKYTFGKPLIAANVSINATARGIGSWQYNNNPDLLRNISD
Zhang et al SIPC            EIEAPESLESNEKTVTVKVCAKYTFGKPLIAANVSINATARGIGSWQYNNNPDLLRNISD
                            ************************************************** ********

Kotsiri et al SIPC          YQFSDEQGCAIFDLVVSKIGIGHRNIGGGNTVIITIDVEEQGTGLRQVEVKEVSQAYSFI
Dreanno et al SIPC          YQFSDEQGCAIFDLVVSKIGIGHRNIGGGNTVIITIDVEEQGTGLRQVEVKEVSQAYSFI
Zhang et al SIPC            YQFSDEQGCAIFDLVVSKIGIGHRNIGGGNTVIITIDVEEQGTGLRQVEVKEVSQAYSFI
                            ************************************************************

Kotsiri et al SIPC          NLRQSDNAQKFLKPKLPFYGEYTLSMRDGKAAKNEIVKVCYTAKYKERVISDEKKPTPDD
Dreanno et al SIPC          NLRQSDNAQKFLKPKLPFYGEYTLSMRDGKAAKNEIVKVCYTAKYKERVISDEKKPTPDD
Zhang et al SIPC            NLRQSDNAQKFLKPKLPFYGEYTLSMRDGKAAKNEIVKVCYTAKYKERVISDEKKPTPDD
                            ************************************************************

Kotsiri et al SIPC          PVYSTHKKYESHVKTEFGYTPFFWETSEPNRRTTGGECREYKTDENGRIVYYIPPQAEDI
Dreanno et al SIPC          PVYSTHKKYESHVKTEFGYTPFFWETSEPNRRTTGGECREYKTDENGRIVYYIPPQAEDI
Zhang et al SIPC            PVYSTHKKYESHVKTEFGYTPLFWETSEPNRRTTGGECREYKTDENGRIVYYIPPQAEDI
                            **********************:*************************************

Kotsiri et al SIPC          DSIDISTSTSVGGDSDSSHSTLTAFFSPSHSYLSIDTHELPEQLPCSGDVTVKLLSTEEG
Dreanno et al SIPC          DSIDISTSTSVGGDSDSSHSTLTAFFSPSHSYLSIDAHELPEQLPCSGDVTVKLLSTEEG
Zhang et al SIPC            DSIDISTSTSVGGDSDSSHSTLTAFFSPSHSYLSIDAHELPEQLPCSGDVTVKLLSTEEG
                            ************************************:***********************

Kotsiri et al SIPC          PVPAMVYKILSRGKIIKAGNMNTNTLTFPVLPKMGPEFKLLVYYIKESGEVVSDSRVFKV
Dreanno et al SIPC          PVPAMVYKILSRGKIIKAGNMNTNTLTFPVLPKMGPEFKLLVYYIKESGEVVSDSRVFKV
Zhang et al SIPC            PVPAMVYKILSRGKIIKAGNMNTNTLTFPVLPKMGPEFKLLVYYIKESGEVVSDSRVFKV
                            ************************************************************

Kotsiri et al SIPC          DKCFPNTVQVSWDQKTVKPGDSASFTVRASPNSVCGISAVDKSTELLGTSNQITLDTVFS
Dreanno et al SIPC          DKCFPNTVQVSWDQKTVKPGDSASFTVRASPNSVCGISAVDKSTELLGTSNQITLDTVFS
Zhang et al SIPC            DKCFPNTVQVSWDQKTVKPGDSASFTVRASPNSVCGISAVDKSTELLGTSNQITLDTVFS
                            ************************************************************
```

```
Kotsiri et al SIPC       KLQQFIINSFESPNQVRSDGDYCRELQLSLVDTLRSGGATAAELTGQSTPEGTPESETSG
Dreanno et al SIPC       KLQQFIINSFESPNQVRSDGDYCRELQLSLVDTLRSGGATVAELTGQSTPEGTPESETSG
Zhang et al SIPC         KLQQFIINSFESPNQVRSDGDYCRELQLSLVDTLRSGGATVAELTGQSTPEGTPESETSG
                         ****************************************.******************

Kotsiri et al SIPC       AAQSSLFIPPPTRSQRFRTDRDDAIKPFDEAGFLVLSNLALETRPCYKRVQAKELPELTE
Dreanno et al SIPC       AAHSSLFIPPPTRSQRFRTDREDAIKPFDEAGFLVLSNLALETRPCYKRVEAKELPELTE
Zhang et al SIPC         AAHSSLFIPPPTRSQRFRTDREDAIKPFDEAGFLVLSNLALETRPCYKRVQAKELPELTE
                         **:******************:****************************:*********

Kotsiri et al SIPC       DKIQASRGGEEELLDDLDSPVPALAFSKESADASRFAAEGGVSGGGGAAPPQEDQVRDFF
Dreanno et al SIPC       DKIQASRDGEEELLDDLDSPVPALAFSKESADASRFAAEGGVSGGGGAAPPQEDQVRDFF
Zhang et al SIPC         DKIQASRDGEEELLDDLDSPVPALAFSKESADASRFAAEGGVSGGGGAAPPQEDQVRDFF
                         *******.**************************************************

Kotsiri et al SIPC       PEAFLFSIETLDAEGVKTVTSEMPDTITSWVGSAICTNSKDGFGISNKTSITTFKPFFTE
Dreanno et al SIPC       PEAFLFSIETLDAEGVKTVTSEMPDTITSWVGSAICTNSKDGFGISNKTSITTFKPFFTE
Zhang et al SIPC         PEAFPFSIETLDAEGVKTVTSEMPDTITSWVGSAICTNSKDGFGISNKTSITTFKPFFTE
                         **** ************************************************************

Kotsiri et al SIPC       VSLPYSMKRGEILSMSVSVFNFLDSSLSVYLEVGASDQYEISGEVAMGLCIAAGRTEVKS
Dreanno et al SIPC       VSLPYSMKRGEILSMSVSVFNFLDSSLSVYLEVGASDQYEISGEVAMGLCIAAGRTEVRS
Zhang et al SIPC         VSLPYSMKRGEILSMSVSVFNFLDSSLSVYPEVGASDQYEISGEVAMGLRIAAGRTEVRS
                         ****************************** ****************** ********:*

Kotsiri et al SIPC       FPVNFLGLGEVNITVTARAQDGYCDEGNTIAPGSDTVIRPIVVKPEGFPQEVTHSRFICL
Dreanno et al SIPC       FPVNFLGLGEVNITVTARAQDGYCDEGNTIAPGSDTVIRPIVVKPEGFPQEVTHSRFICL
Zhang et al SIPC         FPVNFLGLGEVNITVTARAQDGYCDEGNTIAPGSDTVIRPIVVKPEGFPQEVTHSRFICL
                         ************************************************************

Kotsiri et al SIPC       DKDDDHHTETVNLPVPEGLVPDSQRAYFSVIGDLLGPTFQGLEGGLIKSPTGAGEPNMIT
Dreanno et al SIPC       DKDDDHHTETVNLPVPEGLVPDSQRAYFSVIGDLLGQTFQGLEGGLIKSPTGAGEPNMIT
Zhang et al SIPC         DKDDDHHTETVNLPVPEGLVPDSQRAYFSVIGDLLGQTFQGLEGGLIKSPTGAGEPNMIT
                         ********************************** ************************

Kotsiri et al SIPC       LVPNIYIRRYLETTGQLNERQRRQLEHNMKSGYQRQLRFRRYDGSFSSYGNEDPQGSMWL
Dreanno et al SIPC       LVPNIYIRRYLETTGQLNERQRRQLEHNMKSGYQRQLRFRRYDGSFSSYGNEDPQGSMWL
Zhang et al SIPC         LVPNIYIRRYLETTGQLNERQRRQLEHNMKSGYQRQLRFRRYDGPFSSYGNEDPQGSMWL
                         ********************************************.***************

Kotsiri et al SIPC       TAFVVKAFREASEYIEIDETIINKAKDWILKKQNTTGCFPRFGELIHKELKGGTERGGEA
Dreanno et al SIPC       TAFVVKAFREASEYIEIDETIINKAKDWILKKQNTTGCFPRFGELIHKELKGGTERGGEA
Zhang et al SIPC         TAFVVKAFREASEYIEIDETIINKAKDWILKKQNTTGCFPRFGELIHKELKGGTERGGEA
                         ************************************************************

Kotsiri et al SIPC       ALTAFVMLALKDIATTNELANGFACLEDGLLLPNKTLYSEILLAYTYLNMGQDVKGERLV
Dreanno et al SIPC       ALTAFVMLALKDIATTNELANGFACLEDGLLLPNKTLYSEILLAYTYLNMGQDVKGERLV
Zhang et al SIPC         ALTAFVMLALKDIATTNELANGFACLEDGLLLPNKTLYSEILLAYTYMNMGQDVKGERLV
                         ************************************************:***********

Kotsiri et al SIPC       NKLMSKAKREGDDILYWEGDRDSLFGGSRAVDVEMTAYMALSLMHISGKGNMEEAARAIR
Dreanno et al SIPC       NKLMSKAKREGDDILYWEGDRDSLFGGSRAVDVEMTAYMALSLMHISGKGNMEEAARAIR
Zhang et al SIPC         NKLMSKAKREGDDILYWEGDRDSLFGGSRAVDVEMTAYMALSLMHISGKGNMEEAARAIR
                         ************************************************************

Kotsiri et al SIPC       WINTQRNSNGGFKSTQDTIVAVEALSEFASRTFASDLATSVSVTAGGETVQRMVDGDNRL
Dreanno et al SIPC       WINTQRNSNGGFKSTQDTIVAVEALSEFASRTFASDLATSVSVTAGGETVQRMVDGDNRL
Zhang et al SIPC         WINTQRNSNGGFKSTQDTIVAVEALSEFASRTFASDLATSVSVTAGGETVQRMVDGDNRL
                         ************************************************************
```

```
Kotsiri et al SIPC          LYQESKVPDLTLPGTMNFDVSPPGCVVYQSIFRFSSTLEVPDPAFSLGVAAKKRGRTGYE
Dreanno et al SIPC          LYQESKVPDLTLPGTMNFDVSPPGCVVYQSIFRFSSTLEVPDPAFSLGVAAKKRGRTGYE
Zhang et al SIPC            LYQESKVPDLTLPGTMNFDVSPPGCVVYQSIFRFSSTLEVPDPAFPLGVAAKKRGRTGYE
                            *******************************************.**************

Kotsiri et al SIPC          LEVCTSFLRNSGAVDRAILETELPSGYVAVDSTLRDLRRGSAVRSYEIKEGKVIFTLQGV
Dreanno et al SIPC          LEVCTSFLRNSGAVDRAILETELPSGYVAVDSTLRDLRRGSAVRSYEIKEGKVIFTLQGV
Zhang et al SIPC            LEVCTSFLRNSGAVDRAILETELPSGYVAVDSTLRDLRRGSAVRSYEIKEGKVIFTLQGV
                            ************************************************************

Kotsiri et al SIPC          AEDKTCLEFRIIQENEVEQLKPSIVKVHDFYRPEERNIQEYELTPAA
Dreanno et al SIPC          AEDKTCLEFRIIQENEVEQLKPSIVKVHDFYRPEERNIQEYELTPAA
Zhang et al SIPC            AEDKTCLEFRIIQENEVEQLKPSIVKVHDFYRPEERNIQEYELTPAA
                            *********************************************
```

	Kotsiri et al. SIPC (Nucleotides/Amino acids)	**Dreanno et al. SIPC (Nucleotides/Amino acids)**	**Zhang et al. SIPC (Nucleotides/Amino acids)**
Kotsiri et al. SIPC		101/15	(105+44*)/(11+15*)
Dreanno et al. SIPC	101/15		(61+44*)/(8+15*)
Zhang et al. SIPC	(105+44*)/(11+15*)		

Table S3: Differences in the nucleotide and amino acid sequences between the 2 published sequences of SIPC (GenBank: AAR33079.1 and GenBank: AMR58954.1) and the recombinant SIPC described in this study. *:The nucleotide and amino acid sequences deposited by Zhang et al. in GenBank: AMR58954.1 lacks the signal peptide sequence.

Table S4: List of primers used in this study.

Primer Number	Sequence (5' → 3')
1	CATGAGAATTCGCCACCATGGGTGGTCCCGTCGTCGTCCTACTGG
2	CTGAGAATTCCTAAGCAGCGGGAGTCAGCTCGTACTCCTC
3	CGCTTTCGGGGACCTTGACCAGATCCTCTTCAGAGATGAGTTTCTGCTCGGCGCTTGCCGTCGCCAAG
4	CTTCGCGACGGCAAGCGCCGAGCAGAAACTCATCTCTGAAGAGGATCTGGTCAAGGTCCCCGAAAGCG
5	AGCTATCTAGACTAGTGGTGATGGTGATGATGTCCTGAACCAGCAGCGGGAGTCAGCTCGTACTCCTGAA
6	AGATTCTTTTGGCGTACACATACCTGAATA
7	TCTCTGAAGAGGATCTGGGCATGCTCCCCGTCCCCGAAAGCG
8	CGCTTTCGGGGGACGGGGAGCATGCCCAGATCCTCTTCAGAGA
9	AGCACGAATTCGGCATGCGGCCGGCCATGGTGAGCAAGGGCGAGGAGCTG
10	ATCGAGAATTCGCATGCGTTCGAACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGG
11	GCTGACCCTGAAGCTCATCTGCACCACCGG
12	CCGGTGGTGCAGATGAGCTTCAGGGTCAGC
13	ACCACCCTCGGCTACGGCCTGATGTGCTTCGCCCGCTACC
14	GGTAGCGGGCGAAGCACATCAGGCCGTAGCCGAGGGTGGT
15	ACGTCTATATCACGGCCGACAAG
16	CTTGTCGGCCGTGATATAGACGT
17	ACGGCATCAAGGCGAACTTCAAGAT
18	ATCTTGAAGTTCGCCTTGATGCCGT
19	CATCGAGGACGGCGGCGTGCAGCTCGC
20	GCGAGCTGCACGCCGCCGTCCTCGATG
21	CTACCTGAGCTACCAGTCCGCCC
22	GGGCGGACTGGTAGCTCAGGTAG
23	ACTATGAATTCACCGGTGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGG
24	ATCGACTCGAGTTAGTACAGCTCGTCCATGCCGAGAGTGATCC
25	TACCGAATTCGACGAGATGGGGCTACCGCTCCGGCTTTTGCTGCTGCTGCTGCTGCCGCC
26	ATCATACCGGTCCCTGGGGAAGGGCCTTCAGGATGGCATTCACCTCCTCGGCCACAGCT
27	TCCGGTACCGAATTCGCCACCATGAACTTCCAAAACATATTCATATTCGTGGCGTTAATAT TGGCGGTGTTCGCGGGACAATCTCAGGCGGAGCAGAAACTCATCTCTGAAGAGGATCTG
28	ATCAGGCATGCCCCAAGCGTATTCCTTCATCAACCTCAGGCAG
29	ATCAGGCATGCCCAAGCCGTTCTTCACAGAGGTGTCTCTGCCG
30	GATCAGCATGCCCGTTCTGCCAACATTTGAGCTGGAGATCGA
31	CATATGCATGCCCAAGTACAAGGAGAGGGTAATCAGTGACGA
32	CATATGCATGCCCTCGTACCTAAGCATCGACACGCACGAACTT
33	GATCAGCATGCCCGACAAGTGCTTCCCCAACACGGTCCAGGTC
34	GATCAGCATGCCCTCCAACCTGGCCCTGGAGACTCGGCCCTGC
35	TCTGGGCATGCCCGAGGTGACCCACTCTCGCTTCATCTG

36	TCTGGGCATGCCCAGCACTCTAGAGGTGCCCGACCCTGC
37	ACTGGACTAGTGGTGATGGTGATGATGTCCTGAAC
38	CAAAACCAACACTGGCATGGTGCAGCTGGA
39	TACATCAAGGAGAGTGGCGAAGTGGTCAGC
40	CTTCAAGCCGTTCTTCACAGAGGTGTCTCT
41	AGATTCTTTTGGCGTACACATACCTGAATA
42	ACTGCAGTGTTTGTCTCAAACAGCAGGAAA
43	ATGATCACCGTATTTCCGCCTCCAATATTA
44	CCAGTGAAAGCTGCAGTTCCCTGCAGTAAT
45	CAACAATGGTGTCCTGAGTGGATTTGAAAC
46	TCACTTCTTGAGGGAATCCCTCTGGTTTGA

Table S5

Putiative N-linked glycosylation sites on recombinant SIPC	Probability score provided by the server	Server used:	Supported by experimental evidence?	Comments
N263	0.76	http://www.modpred.org/	Supported	Present only in Tr.1
N285	0.7531	http://www.cbs.dtu.dk/services/Net	Supported	Present only in Tr.1
N289	0.5656	http://www.cbs.dtu.dk/services/Net	Supported	Present only in Tr.1
N302	0.68	http://www.modpred.org/	Supported	Present only in Tr.1
N309	0.6872	http://www.cbs.dtu.dk/services/Net	Supported	Present only in Tr.1
N618	0.64	http://www.modpred.org/	Not supported	Tr.5 after deglycosylation does not have an apparent MW higher than Tr.6
N899	0.7683	http://www.cbs.dtu.dk/services/Net	Supported	Tr.6 after deglycosylation has a higher MW than Tr.7
N984	0.7800	http://www.cbs.dtu.dk/services/Net	Supported	Tr.7 after deglycosylation has an apparent MW similar to its Calculated
N1186	0.65	http://www.modpred.org/	Not supported	Tr.8 does not have an apparent MW higher than its Calculated
N1246	0.7153	http://www.cbs.dtu.dk/services/Net	Not supported	Tr.8 does not have an apparent MW higher than its Calculated
N1341	0.74	http://www.modpred.org/	Not supported	Tr.8 does not have an apparent MW higher than its Calculated

N1462	0.71	http://www.modpred.org/	Not supported	Tr.8 does not have an apparent MW higher than its Calculated

Putiative O-linked glycosylation sites on recombinant SIPC	Probability score provided by the server	Server used:	Supported by experimental evidence?	Comments
S79	0.80	http://www.modpred.org/	Supported	Present only in full-length recombinant SIPC
T429	0.780535	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.3 onwards
S436	0.777751	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.3 onwards
T437	0.5	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.3 onwards
S443	0.635919	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.3 onwards
S498	0.89	http://www.modpred.org/	Supported	Present in Tr.3 onwards
T499	0.83	http://www.modpred.org/	Supported	Present in Tr.3 onwards
S500	0.94	http://www.modpred.org/	Supported	Present in Tr.3 onwards
S502	0.88	http://www.modpred.org/	Supported	Present in Tr.3 onwards
S708	0.656591	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.5 onwards
T712	0.749505	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.5 onwards
T717	0.743219	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.5 onwards
S720	0.922073	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.5 onwards
T721	0.814289	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards
T725	0.530811	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards
S728	0.602923	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards
T730	0.595919	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards
S731	0.648028	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards
S736	0.751889	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards
S737	0.657497	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards
T744	0.858312	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards
S746	0.654746	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards
T751	0.685158	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards

S822	0.669561	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards
S826	0.517458	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards
S835	0.797811	http://www.cbs.dtu.dk/services/Net	Supported	Present in Tr.6 onwards
S873	0.81	http://www.modpred.org/	Supported	Present in Tr.6 onwards
T905	0.74	http://www.modpred.org/	Supported	Present in Tr.7 onwards
S1372	0.523284	http://www.cbs.dtu.dk/services/Net	Not supported	Tr.8 does not have an apparent MW higher than its Calculated
T1376	0.543343	http://www.cbs.dtu.dk/services/Net	Not supported	Tr.8 does not have an apparent MW higher than its Calculated
S1438	0.5	http://www.cbs.dtu.dk/services/Net	Not supported	Tr.8 does not have an apparent MW higher than its Calculated
S1561	0.85	http://www.modpred.org/	Not supported	Tr.8 does not have an apparent MW higher than its Calculated

Putative C-linked glycosylation sites on recombinant SIPC	**Probability score provided by the server**	**Server used:**	**Supported by experimental evidence?**	**Comments**
None	-	http://www.modpred.org/		

Putative Myristoylation sites on recombinant SIPC	**Probability score provided by the server**	**Server used:**	**Supported by experimental evidence?**	**Comments**
None	-	http://mendel.imp.ac.at/myristate/SUPLpredictor.htm		

Putative Prenylation sites on recombinant SIPC	**Probability score provided by the server**	**Server used:**	**Supported by experimental evidence?**	**Comments**
None	-	http://www.modpred.org/		

Journal of Experimental Biology • Supplementary information

Putative GPI-anchor amidation sites on recombinant SIPC	Probability score provided by the server	Server used:	Supported by experimental evidence?	Comments
N1323	0.63	http://www.modpred.org/		
N1341	0.87	http://www.modpred.org/		

Table S6

SIPC or truncated fragment name	Calculated MW of SIPC and its truncated fragments	Calculated MW of SIPC and its truncated fragments from Western blots	SEM	N	Differences between calculated and Calculated MW and its truncated fragments from	SEM	N	Reduction in MW by deglycosylation assays	SEM	N	N-linked glycans	Small glycans	O-linked glycans	Other post-translational modifications
VSIPC	199,159	246,4109	2,37	3	47,2519	2,37	3				15	8	8	16
SIPC	171,453	217,7961	2,2	7	46,34306	2,2	7	196,0638	4,08	3	15	8	8	15
Tr.1	147,276	178,0196	3,56	6	30,74362	3,56	6				15	8	8	0
Tr.2	134,108	153,4121	1,62	7	19,8299	1,81	6				7	4	8	0
Tr.3	128,479	149,4574	1,49	6	21,09269	1,26	7				7	5	9	0
Tr.4	116,219	133,2525	2,47	6	17,0335	2,47	6				7	3	7	0
Tr.5	106,35	124,7959	1,64	6	18,44593	1,64	6	113,4602	2,17	3	7	4	7	0
Tr.6	89,543	108,5631	2,88	6	19,02005	2,88	6	94,64828	1,84	3	7	7	5	0
Tr.7	74,895	83,84607	1,81	7	8,951066	1,81	7	76,80681	1,25	3	5	2	2	0
Tr.8	62,454	59,48821	1,34	4	-2,965797	1,34	4				0	0	0	0
Tr.9	17,316	37,46218	0,44	4	20,14618	0,44	4				0	0	0	20

Dataset 1						
Data	Data	1243	85	49	62	51
Data	Data	MW	Nglyc	Deglyc	Oglyc	Other
246	A	199	15	8	8	16
218	B	171	15	8	8	15
178	C	147	15	8	8	0
153	D	134	7	4	8	0
149	E	128	7	5	9	0
133	F	116	7	3	7	0
124	G	106	7	4	7	0
108	H	89	7	7	5	0
84	I	74	5	2	2	0
60	J	62	0	0	0	0
37	K	17	0	0	0	20