Conserved sequences of sperm-activating peptide and its receptor throughout evolution, despite speciation in the sea star *Asterias amurensis* and closely related species

Mia Nakachi[2,3], *Motonori Hoshi*[2], *Midori Matsumoto*[2] and *Hideaki Moriyama*[1,3]

Department of Biosciences and Informatics, Keio University, Hiyoshi, Kouhoku-ku, Yokohama, Japan and Department of Chemistry, Environmental Toxicology and Biotechnology, University of Nebraska – Lincoln, Lincoln, Nebraska, USA

Summary

The asteroidal sperm-activating peptides (asterosaps) from the egg jelly bind to their sperm receptor, a membrane-bound guanylate cyclase, on the tail to activate sperm in sea stars. Asterosaps are produced as single peptides and then cleaved into shorter peptides. Sperm activation is followed by the acrosome reaction, which is subfamily specific. In order to investigate the molecular details of the asterosap–receptor interaction, corresponding cDNAs have been cloned, sequenced and analysed from the Asteriinae subfamily including *Asterias amurensis, A. rubens, A. forbesi and Aphelasterias japonica,* as well as *Distolasterias nipon* from the Coscinasteriinae subfamily. Averages of 29% and 86% identity were found from the deduced amino acid sequences in asterosap and its receptor extracellular domains, respectively, across all species examined. The phylogenic tree topology for asterosap and its receptor was similar to that of the mitochondrial cytochrome *c* oxidase subunit I. In spite of a certain homology, the amino acid sequences exhibited speciation. Conservation was found in the asterosap residues involved in disulphide bonding and proteinase-cleaving sites. Conversely, similarities were detected between potential asterosap-binding sites and the structure of the atrial natriuretic peptide receptor. Although the sperm-activating peptide and its receptor share certain common sequences, they may serve as barriers that ensure speciation in the sea star *A. amurensis* and closely related species.

Keywords: Acrosome reaction, Guanylate cyclase, Sea star, Sperm-activating peptide

Introduction

Speciation in fertilization is critical in the beginning of development. While mechanisms of fertilization have diverged from invertebrates to humans, several strictly defined steps have been governed by specific molecules from eggs and sperm for the recognition of one another (Vacquier, 1998). These steps include sperm activation and/or attraction to the egg, as well as sperm acrosome reaction upon egg contact. Understanding the mechanisms of these molecular interactions is important, as any disruptions of these processes could cause infertility and destroy species barriers at the molecular level.

In the sea star, *Asterias amurensis,* the asteroidal sperm-activating peptide (asterosap) has been purified and identified from the extracellular jelly coat of the egg. Asterosap is composed of 34 amino acids and has an intramolecular disulphide bond that is essential for its physiological activity (Nishigaki *et al.*, 1996). The asterosap precursor mRNA encodes 10 asterosap molecules that are tandemly connected by spacer sequences (Matsumoto *et al.*, 1999). Asterosap binds to the extracellular domain of its receptor, a membrane-bound guanylate cyclase, in the sperm flagellum to activate its intracellular catalytic domain for the production of cGMP (Nishigaki *et al.*, 2000; Matsumoto *et al.*, 2003). The cooperation of three

---

[1]All correspondence to: Hideaki Moriyama. Department of Chemistry, University of Nebraska – Lincoln, 28 Hamilton Hall, Lincoln, Nebraska 68588–0304, USA. Tel: +1 402 472 5367. Fax: +1 402 472 9402. e-mail: hmoriyama2@unl.edu
[2]Department of Biosciences and Informatics, Keio University, Hiyoshi, Kouhoku-ku, Yokohama 223–8522, Japan.
[3]Department of Chemistry, Environmental Toxicology and Biotechnology, University of Nebraska – Lincoln, 28 Hamilton Hall, Lincoln, Nebraska 68588–0304, USA.

egg jelly coat components, ARIS (acrosome reaction-inducing substance), Co-ARIS and asterosap, induces the acrosome reaction (Hoshi *et al.*, 1994). Asterosap is essential for the induction of the acrosome reaction by elevating the internal pH of sperm (Kawase *et al.*, 2005). In addition, asterosap induces hyperactivation-like flagellar movement in sperm that are firmly attached to a solid surface (Shiba *et al.*, 2006), potentially allowing an effective acrosome reaction on the surface of the jelly coat. Furthermore, asterosap also has sperm-attracting activity (Bohmer *et al.*, 2005). The specificity of asterosap activity is conserved within the superorder level (Nakachi *et al.*, 2006).

In sea urchin, resact (SAP-IIA) from *Arbacia punctulata* is the only peptide known to have sperm-attracting activity among more than 70 sperm-activating peptides identified from the egg jelly coat (Ward *et al.*, 1985). Resact binds directly to a membrane-bound guanylate cyclase receptor in sperm and activates it (Singh *et al.*, 1988; Kaupp *et al.*, 2003).

In order to investigate the molecular details of the interaction between asterosap and its receptor, we cloned the cDNA encoding these proteins from five closely related sea star species. We report that although the amino acid sequences are highly conserved over these species, small differences in the sequences can explain their speciation. The functional mechanism of the asterosap receptor is also discussed by analogy to another membrane-bound guanylate cyclase, the atrial natriuretic peptide receptor.

Materials and methods

Animals and gametes

Sea stars were collected locally during their annual breeding seasons as described previously (Nakachi *et al.*, 2006). The sea stars used in this study were *Asterias amurensis* (Aa), *A. rubens* (Ar), *A. forbesi* (Af), *Aphelasterias japonica* (Apj) and *Distolasterias nipon* (Dn).

Cloning of genes

Total RNA was prepared from fresh testes or ovaries by the acid guanidium–phenol–chloroform (AGPC) method (Sambrook *et al.*, 1989). First-strand cDNA was synthesized using reverse transcriptase with either a random 9-mer using Rever Tra Ace (TOYOBO) or the ImProm-2 reverse transcription system (Promega) and was subsequently used as a template for PCR. DNA fragments of asterosap and spacer sequences were amplified by PCR from ovary cDNA using primers TN1 5′-ATCGCTCGAGTAAGACAA-3′ and TN2 5′-TCGTTTGCTCTGACATTG-3′. Extracellular

domain sequences of the asterosap receptors were amplified by PCR from testes cDNA using primers hani1 5′-AGGTGTTTGATGCTGTCTGTCG-3′, hani2 5′-GCGTAGTCTCCGTTATTGAGACC-3′ MN1′ 5′-AAGGAGACAACAAGGATTTAACGTG-3′ and MN2 5′-GACCGGCATATCTAGAGGTGG-3′. Guanylate cyclase domain sequences of the asterosap receptors were also amplified from testes cDNA using primers P2 5′-GACATCGTCGGGTTTACC-3′ and P6 5′-AGTGTTGACGGTGTCACC-3′. PCR products were separated by size via agarose gel electrophoresis, extracted and cloned into the pGEM-T vector (Promega) and sequenced with the M13 forward and reverse primers using BigDye chemistry on an ABI PRISM 3100 DNA sequencer (Applied Biosystems). We deposited the sequences into GenBank and the accession numbers are listed in the Supplemental materials section.

Sequence analysis

Multiple alignments of nucleotide and amino acid sequences were done using Clustal X (Thompson *et al.*, 1997). Phylogenetic reconstruction of nucleotide sequences was performed using PHYLIP version 3.66 (Felsenstein, 2005) and the phylogenetic trees were viewed using njplot (Perriere & Gouy, 1996). Secondary structures and β–turn structures of the amino acid sequences were predicted using PORTER (Pollastri & McLysaght, 2005) and BetaTPred2 (Kaur & Raghava, 2003), respectively. Hydropathy index and pI of amino acid sequences were calculated by sequence analysis (Gilbert, 2006) using Kyte & Doolittle parameters (Kyte & Doolittle, 1982) in a running average of five amino acids. The coiled-coil structures in the amino acid sequences were predicted by COILS (Lupas *et al.*, 1991). Accession numbers for the guanylate cyclase genes were X12874 from sea urchin (*Arbacia punctulata*; Ap) and AB070354 from Aa. Accession numbers for the mitochondrial cytochrome *c* oxidase 1 sequences other than ours were AF498647 (Af), AF031028 (Ap) and AF525453 (*Hemicentrotus pulcherrimus*, Hp, sea urchin). Sequence alignments referring to their 3D structures were performed by the 3D-PSSM (Kelley *et al.*, 2000) and Swiss model (Schwede *et al.*, 2003) servers. The hormone-bound, atrial natriuretic peptide receptor extracellular domain (Ogawa *et al.*, 2004) was accessed from the Protein Data Bank (ID code 1t34).

Results

In order to investigate the molecular details of the interaction between asterosap and its receptor, we cloned the cDNAs encoding these proteins from five closely related sea star species.

Sequence variation among asterosap

Sea star asterosap sequences exist as 10 consecutive repeats interspersed with short spacers (Matsumoto *et al.*, 1999). These 10 asterosap sequences from the same genome are not identical but vary among them. The cDNAs of randomly chosen seven asterosap isoforms were each cloned including their spacer regions from five sea star species: *Asterias amurensis* (Aa), *A. rubens* (Ar), *A. forbesi* (Af) *and Aphelasterias japonica* (Apj), from the Asteriinae subfamily and *Distolasterias nipon* (Dn) from the Coscinasteriinae subfamily. All these five species belong to the family Asteriidae.

The deduced amino acid sequences of the asterosaps were highly conserved among the species with more variation found in the N-terminal region (Fig. 1). Of the 35 cloned sequences, the asterosaps had 10 completely conserved amino acid sites out of 34, covering 29% of the peptide length. While all asterosaps were of the same length, the spacer regions varied ranging from 12 to 22 amino acids. The spacer regions had four completely conserved amino acids sites out of 12 (excluding gap sites), corresponding to 33% identity.

In *A. amurensis* (Aa), asterosap has an intramolecular disulphide bond between Cys8 and Cys32, which is known to be essential for its activity (Nishigaki *et al.*, 1996). These Cys residues were among the completely conserved amino acids from all asterosap sequences examined. As shown in Fig. 1, some amino acid substitutions were unique in *Asterias,* in Apj or in Dn. All the spacer regions had paired basic residues, Lys–Arg or Lys–Lys, at both ends, which are commonly found at the cleavage sites of pro-hormone precursors (Hook *et al.*, 2004). The results suggest that the asterosap precursor undergoes endoproteolytic processing at either the C- or N-termini of paired basic residues followed by the removal of the remaining basic residues as shown in other pro-hormone processing pathways (Fricker, 1988; Steiner, 1998; Hook *et al.*, 2004). The pI represents the favourable environmental pH range of biologically active molecules such as asterosap. While the distribution of pIs for asterosap was between 4.5 and 10.0 in Aa, Af, Ar and Apj, pIs for asterosaps from Dn range from 3.8 to 6.3.

Every asterosap was predicted to have two amphipathic helices between two conserved Cys residues (Fig. 1). Positive charges (Arg11 and 13) in the first helix and negative charges (Asp20 and Glu21) in the second helix were highly conserved. Then, we assumed that these helical structures are important for asterosap binding to its receptor. Hydropathy index of the asterosaps indicated the overall hydrophilicity, explaining its diffusive characteristic in seawater. Conserved hydrophobic residues in each helix may also be involved in either binding to its receptor or peptide folding. Spacer regions had turn structures and were entirely hydrophilic.

Asterosap receptor

A. amurensis asterosap receptor has four domains: extracellular, transmembrane, kinase-like and guanylate cyclase (Matsumoto *et al.*, 2003). We cloned the cDNAs of the extracellular and guanylate cyclase domains of the asterosap receptor from four additional sea star species and compared these sequences. The deduced amino acid sequences (Fig. 2) shared high similarities in these domains: 86% in the extracellular domain and 94% in the guanylate cyclase domain. Most amino acid differences were found between *Asterias* (Aa, Ar and Af) compared with Apj and Dn: more than 10% of the extracellular domain sequences differed. In the extracellular domain, all four Cys residues and three potential glycosylation sites were conserved among the five species. In the guanylate cyclase domain, all putative catalytic sites (Liu *et al.*, 1997) were conserved (Fig. 2).

Molecular evolution of asterosap and its receptor

In order to compare the evolutionary patterns between asterosap and its receptor sequences, we performed phylogenetic reconstruction at the nucleotide level (Fig. 3). The phylogenetic tree reconstructed using the 35 asterosap sequences showed three major groups, each corresponding to the three subfamilies: Dn, Apj and *Asterias,* although the bootstrap analysis did not support these three clusters (Fig. 3*a*). Interestingly, clustering among *Asterias* asterosap sequences was not consistent with species clustering. For example, the clusters from a to d in Fig. 3*a* included more than two *Asterias* species. Moreover, identical asterosap sequences were found between Af02 and Aa13 and between Apj07 and Af05. More phylogenetics analysis may be required to establish taxonomical relationships among these sea star species, especially the phylogenetic location of Apj. The occurrence of close relationships among *Asterias* species found in phylogenies agrees with our previous results on the gamete-compatibility grouping in sea stars (Nakachi *et al.*, 2006).
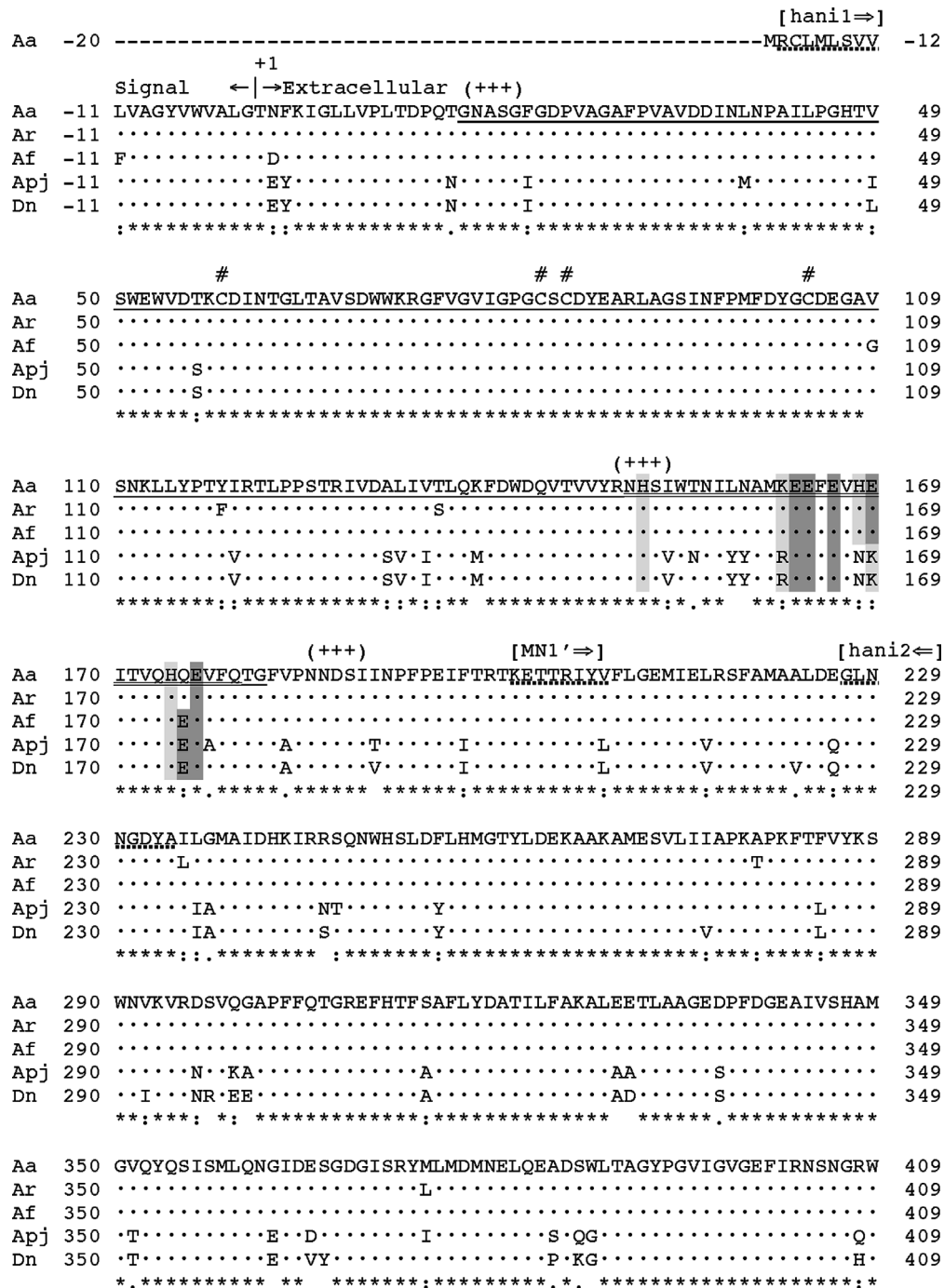
Discussion

Interaction between asterosap and its receptor

The structure of the asterosap receptor and the binding of the asterosap to its receptor were modelled based on the structure of the rat atrial natriuretic peptide and its receptor (ANP and NPR-A; PDB entry 1t34).

```
Number        1       10        20        30 34 +1      +10 -10       -1
Ruler        |--------|---------|---------|---|/|--------|   |--------|
                #[TN1⇒]                     #[TN2⇐]
Aa01(P14)    GQTQFGVCIARVRQQHQGQDETSIFQAILSQCQS-KREITKEPPS-----EDDETKR 51 Aa
Aa03(Cl.29)  ·NKEM·L······N···NGA············-····D···-----······· 51
Aa04(P11-1)  ·NKEM·L······N····A···········-····VQY·S·-----······· 51
Aa06(P13-2)  ············A·····R···-····D···-----······· 51
Aa08(P9-1)   ·DKEM·······A·········-···MSLPLSETQQNQ··EDI·· 56
Aa11         ············L·······-····SS-----QGD·EPE·· 51
Aa13(P13-1)  ·ASEM·L······K·····A···········-·····D···-----······· 51
Af02(P13-1)  ·ASEM·L······K·····A···········-·····D···-----······· 51 Af
Af03         ·DKEM·L····G··N·····A···R···-·····D···-----······· 51
Af04         ··········K·····AF····-·····D···-----······· 51
Af05         ·DKEM·L·······A······N···-··T··D··-----K······· 51
Af08         ·DKEM·······A···L·········-·····D··-----K······· 51
Af09         ·DKEM·L······N·····A···········-····S·D···-----······· 51
Af11(P15)    ··············A·················-----------------· 46
Ar01         ·DKEM·L·······S···A·········-···FS·D···-----······· 51 Ar
Ar02(P13-2)  ············A···R···-···S·D··-----K······· 51
Ar03         ·DKEV·L······N·····A···R·-·····D···-----······· 51
Ar04         ············A··Y·····-·············-----------····V·· 46
Ar05         ·DKEM·L······N·····A···R·-·····D···-----······· 51
Ar06         ·····I·······A··Y·····-············-----------····V·· 46
Ar07         ·DKEM·L······N·····AF·Y····N···-·····DL·-------··G···· 51
Apj01        ·DKEL·L······N···GV··S···MN···-··K···D···-----······· 51 Apj
Apj02        ··KEL·L······N····V··Y···M···R--···I·D···-----······· 51
Apj03        ··KEL·L······N····V··Y···M···R--·K···D···-----······· 51
Apj06        ·DKEL·L······N····V··Y···M···-····I·D···-----······· 51
Apj07        ·DKEM·L······N····A······N···-····D···-----······· 51
Apj10        ·DKEL·L······N····V··Y·······-····I·D···-----······· 51
Apj11        ·DKEL·L···········V··Y···M···-····I·D···-----······· 51
Dn03         ·DKEL·L·····K·N····V··Y···M···-····I·D···-----······· 51 Dn
Dn04         ·D·EL·L····I···SH··GDA··Y·S·V·········-····D···-----·····K 51
Dn05         ·D·EL·L······N····DVY····M···-·····D···-----······· 51
Dn06         ·DAEL·L······N····DV··········-·····D···-----······· 51
Dn07         ·D·EL·L······N····DV··········-····I·D···-----······· 51
Dn08         ·DAEM·L····K··NH····DA······V····-····I·D··I-----······· 51
Dn09         SDAEL·L·····R·N····DV···S·····-·····D···-----······· 51
Common       .  :.*:****: :: :.*. . *  :*: **: **:           .*   *:
Motif        ~~~================~~~=============~~~
Phobic       ----++++||||||||||||||+++++||||||-||||||||||-----|||----
```

Figure 1 Comparison of asterosap and spacer amino acid sequences from five sea star species. Plus (+) and minus (−) position numbers at the top indicate positions within the spacers from the N- and C-termini, respectively. A dot (·) shows the residue identical to Aa01; a minus sign (−) indicates a gap. The slash (/) in the ruler shows the potential cleavage site. Sequences used for PCR primers are underlined by dotted lines in Aa01 and primer names are indicated in square brackets ([]) above with arrows indicating the amplification directions. Two conserved Cys residues are marked by number signs (#) above Aa01. Positively and negatively charged conserved residues in asterosaps are shaded with light and dark grey, respectively. The 'Common' row shows the symbols based on the sequence similarity for each position after CLUSTAL X: an asterisk (∗) for a completely conserved position; a colon (:) for the position conserved among the 'strong' amino acid groups; a period (.) for the position conserved among the 'weaker' amino acid groups. The 'Motif' row shows the predicted secondary structures for Aa04: a tilde (~) for the coil structure; and an equal (=) for the helical structure. The 'Phobic' row shows Kyte & Doolittle hydropathy index from a running average of five amino acids: a plus (+) for hydrophobicity; a vertical bar (|) for hydrophilicity.

The sequence identity between the asterosap receptor and NPR-A sequences was 27%, which is considered sufficient for structural modelling. NPR-A is known to form a dimer (Wilson & Chinkers, 1995). Our native PAGE result showed that the Aa asterosap receptor in sperm appears to be present as a dimer and oligomer (data not shown). In addition, the Aa asterosap receptor has a potential coiled-coil region in the area similar to that of NPR-A, which is important for dimerization (Wilson & Chinkers, 1995). A sequence

```
                                                              [hani1⇒]
Aa  -20 ------------------------------------------------MRCLMLSVV -12
            +1
        Signal   ←|→Extracellular (+++)
Aa  -11 LVAGYVWVALGTNFKIGLLVPLTDPQTGNASGFGDPVAGAFPVAVDDINLNPAILPGHTV 49
Ar  -11 ············································ ·················  49
Af  -11 F··········D···········································  49
Apj -11 ···········EY············N····I··············M·········I  49
Dn  -11 ···········EY············N····I··············M·········L  49
        :*********:*:***********.*****:****************:*********:

            #                            # #                    #
Aa   50 SWEWVDTKCDINTGLTAVSDWWKRGFVGVIGPGCSCDYEARLAGSINFPMFDYGCDEGAV 109
Ar   50 ·················································  109
Af   50 ············································· ·········G 109
Apj  50 ······S·················································· 109
Dn   50 ······S·················································· 109
        *****:******************************************************

                                    (+++)
Aa  110 SNKLLYPTYIRTLPPSTRIVDALIVTLQKFDWDQVTVVYRNHSIWTNILNAMKEEFEVHE 169
Ar  110 ·······F·················S···················· ·····  169
Af  110 ·················································  169
Apj 110 ·········V·········SV·I···M············V·N··YY··R····NK 169
Dn  110 ·········V·········SV·I···M············V····YY··R····NK 169
        *******:;***********:;*.:** ***************:*.** **:*****:;

                    (+++)              [MN1'⇒]              [hani2⇐]
Aa  170 ITVQHQEVFQTGFVPNNDSIINPFPEIFTRTKETTRIYVFLGEMIELRSFAMAALDEGLN 229
Ar  170 ····::···················· ············  229
Af  170 ·····E·················  229
Apj 170 ····E·A····A·····T·····I·········L······V········Q··· 229
Dn  170 ····E·······A·····V·····I·········L······V······V··Q··· 229
        *****:*.*****.****** ******:***********:*******:******.**:*** 229

Aa  230 NGDYAILGMAIDHKIRRSQNWHSLDFLHMGTYLDEKAAKAMESVLIIAPKAPKFTFVYKS 289
Ar  230 ····L··········································T········· 289
Af  230 ·················································  289
Apj 230 ······IA·········NT······Y·················L···· 289
Dn  230 ······IA·········S·······Y·············V········L···· 289
        *****:;.********  :*******:*****************:***;****:**** 

Aa  290 WNVKVRDSVQGAPFFQTGREFHTFSAFLYDATILFAKALEETLAAGEDPFDGEAIVSHAM 349
Ar  290 ········································  349
Af  290 ················································  349
Apj 290 ······N··KA···········A··············AA·····S········· 349
Dn  290 ··I···NR·EE···········A·············AD·····S········· 349
        **:***: *: *************:**************  ******.************ 

Aa  350 GVQYQSISMLQNGIDESGDGISRYMLMDMNELQEADSWLTAGYPGVIGVGEFIRNSNGRW 409
Ar  350 ······················L···················  409
Af  350 ················································  409
Apj 350 ·T··········E··D········I·········S·QG·············Q· 409
Dn  350 ·T··········E··VY·············P·KG·············H· 409
        *.********** **  ******:********.*. ****************:* 
```

Figure 2 Comparison of asterosap receptor amino acid sequences from five sea star species. A dot (·) indicates the amino acid identical to Aa at the corresponding position. Mature proteins start from the residue +1 after the N-terminal signal peptides. Each domain starts at |→ and ends at ←|. Sequences used for PCR primers are underlined with dotted lines in Aa and primer names are given in square brackets ([]) with arrows indicating the amplification directions. An asterisk (∗), a colon (:) and a period (.) at the bottom show the levels of sequence similarities for each position (see the Fig. 1 legend). Extracellular Cys residues are marked by a number signs (#). Potential N-glycosylation sites (Asn–X–Ser/Thr) are marked by a plus (+) in parentheses. Parts of the Aa sequence subjected to the homology modelling are underlined. A potential ligand-binding region is double underlined with positively and negatively charged residues shaded with light and dark grey, respectively. Potential phosphorylation sites in the region equivalent to the rat atrial natriuretic peptide receptor are shaded with light grey in Aa (from positions 500 to 517). A potential coiled-coil region is underlined by short- and long-dotted lines. Potential GTP binding sites in the guanylate cyclase domain are marked by 'g'. Except for Aa, only extracellular and guanylate cyclase domains were cloned and sequenced.

```
                                   447                    466
             (+++)              [MN2⇐] #         #   |→Transmembrane   ←|
Aa   410  TFNATDDYNTPIKWPNDAGPPLDMPVCGYFEEFCPKYGLYFGLGVPIVLLIVGCAVGYFY  469
Ar   410  ·················             428
Af   410  ·····E···V·········             428
Apj  410  ···R····DV·········             428
Dn   410  ···R····DVD········             428
          *** *:**:. ********

Aa   470  YRKIKYEGELDSLVWKINFDDVQAKGKDTNKSGISMKSMVMSTLSVMTNQETQQIFARIG  529

                536
                |→Kinase-like
Aa   530  TYRGNICAIKAVNKHSIDLTRTVRQELKAMHDVRHDNVCQFVGASVDSPHVCILMTYCAK  589

Aa   590  GSLQDILENDDIKLDNMFLASMIADLVKGMIYIHTSMIESHGNLKSSNCVVDNRFVLQIT  649

Aa   650  DYGLHEFKKGQGEDPDLPDDVRYRNLLWRAPELLRMGKKMPLAGTPKGDVYSFAVVLTEM  709

Aa   710  YSRAEPYNLNDDEPEEIVEKVMAGSIPPYRPLLNDVNEKAPECVLKAIRSCWGEDPVERP  769

          771                                                817
          ←|                                                |→Guanylate cyclase
Aa   770  DFFKARTMLAPLQKGLKPNIMDNMITIMERYTNNLEELVDERTOELQKEKAKTEQLLHRM  829

                                                      [P2⇒]
                       g                              g   g
Aa   830  LPPSIASQLIKGISVAPEAFDMVTIFFS-------------------------DIVGFT  863

                  g                    g g g  g
Aa   864  ALSAASTPIQVVNLLNALYTTFDATISNYDVYKVETIGDAYMLVSGLPLRNGNRHAGMIA  923
Ar   864  ············································             923
Af   864  ···········································I·             923
Apj  864  ···················L···I····················             923
Dn   864  ···················L···I····················             923
          ******************* *** ***********************:***********

                                                      [P6⇐]
                       g                    g gggg  gg   g g
Aa   924  SAAWHLLEEVTTFVVPHKRDEKLKLRIGIHSGSCVAGVVGLTMPRYCLFGDTVNTASRME  983
Ar   924  ···········································             972
Af   924  ···········································             972
Apj  924  ···················Q·V·····················             972
Dn   924  ·········S································             972
          **********:*******:* ************************

                  1007
                  ←|
Aa   984  SNGLALKIHVSPECRQVLQELGGYNLVERGLVAMKGKGEILTYWLEGQDPSYKVERNKPP  1043

Aa  1044  KPDL                                                1047
```

Figure 2 (*Cont.*)

alignment between the Aa asterosap receptor and NPR-A, taking their 3-D structures into consideration, indicated that they share key amino acid residues, Trp71 and Asp59 in the asterosap receptor and Trp74 and Asp62 in NPR-A, as well as Arg 13 in asterosap and Arg14 in ANP (Fig. 4). These amino acids have been strictly conserved in all examined sea star sequences.

Resact is a sperm-activating peptide from *Arbacia punctulata* (Ap) and is composed of 14 residues with an intramolecular disulphide bond. The primary structure of the resact receptor was also compared with the asterosap receptor to predict their functional similarity. Their extracellular domains have high sequence similarity. While the resact receptor lacks the large C-terminal portion of the guanylate cyclase domain including almost all putative key catalytic residues, the asterosap receptor and resact receptor may have similar modes of ligand binding but different catalytic regulatory mechanisms.
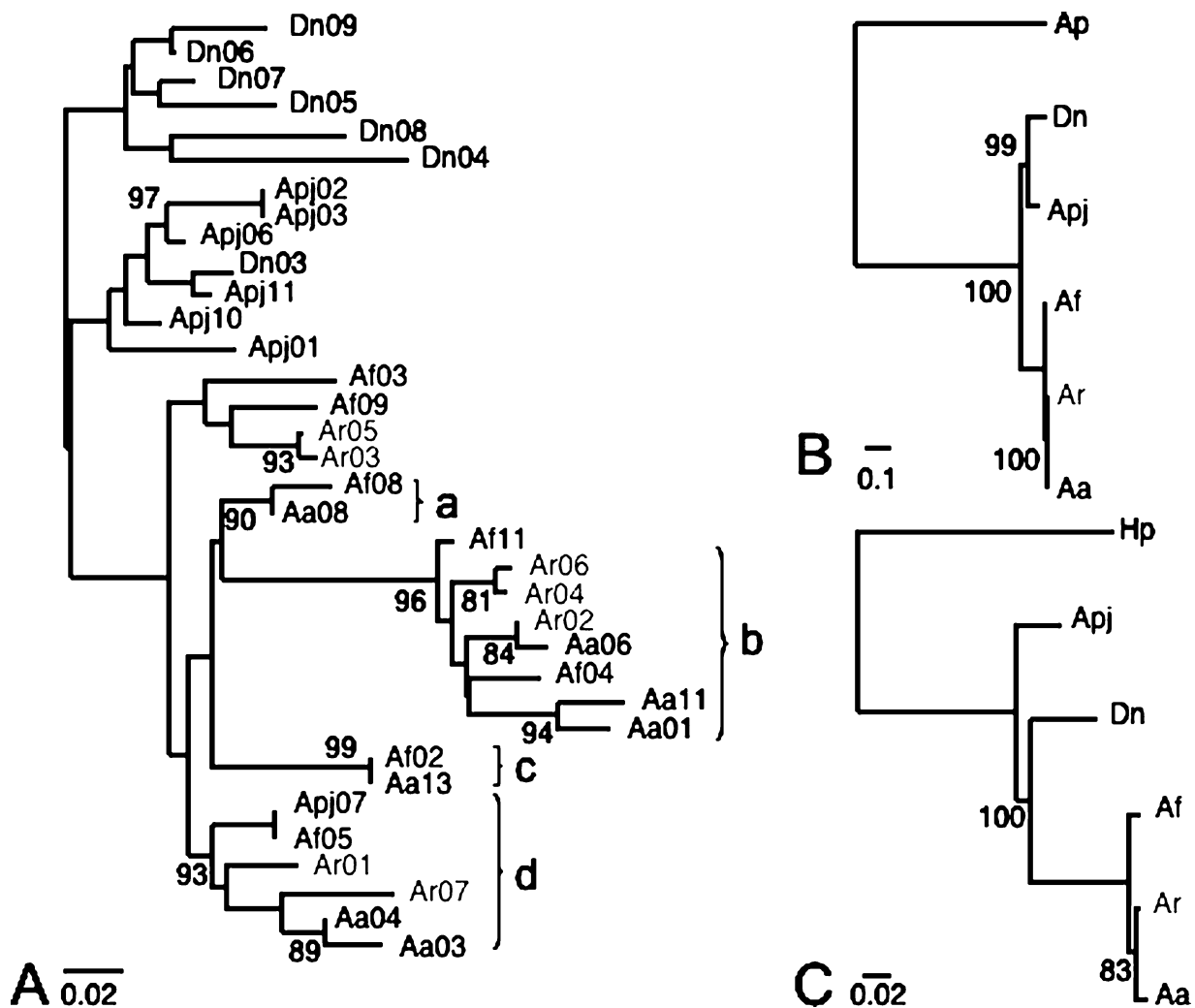
Figure 3 Phylogenetic relationships reconstructed based on the nucleotide sequences of asterosap (a) and its receptor (b and c). The receptor phylogenies were reconstructed using the extracellular domain (B, 1299 nucleotides) and the guanylate cyclase domain (C, 326 nucleotides) separately. Neighbour-joining phylogenetic method (Saitou and Nei, 1987) was used on nucleotide distances estimated by F84 model (Kishino and Hasegawa, 1989). Bootstrap analysis was done with 1000 pseudoreplicates and supporting values greater than 80% are shown on each consensus phylogeny. For asterosap receptor analyses, resact receptor sequences from sea urchin were added as outgroups: Ap, *Arbacia punctulata*; Hp, *Hemicentrotus pulcherrimus*.

Survival strategies of sea stars

The acrosome reaction is an important fertilization process in sea stars, which takes place upon the arrival of sperm onto the egg jelly surface. ARIS, a huge proteoglycan, is the major component of egg jelly in sea stars. A sulphate moiety in ARIS plays an important role in the acrosome reaction. The pH of the egg jelly and the neighbouring seawater can be affected by the content of the sulphate moiety caused by its acidic properties. Each species has several varieties of asterosaps at different pIs. This variation is advantageous for sea stars because a wide distribution of pIs allows with the organism to adjust to any environmental changes. Alternatively, differences in pI can reflect the different sugar components of egg jelly in

sea stars. For example, Dn has fewer sulphate moieties in its egg jelly, which results in a different pH in the environment.

Aa, Apj and Dn live in close proximity off the Japanese shore in the North Pacific Ocean. There is, therefore, a certain risk of cross-fertilization between them. The fertilization season of Apj (fall) is different from that of Aa and Dn (early spring). While Aa and Apj are able to induce cross-acrosome reactions *in vitro* using each other's egg jelly (Nakachi *et al.*, 2006), natural cross-fertilization is prohibited by the season. The cross-acrosome reaction (i.e. fertilization) is, therefore, unlikely to occur between Aa, Apj and Dn. These facts are evident in the phylogenic relationship presented in Fig. 3.

NPR-A/SAPR
W74/W71
D62/D59
ANP/SAP
R14/R13
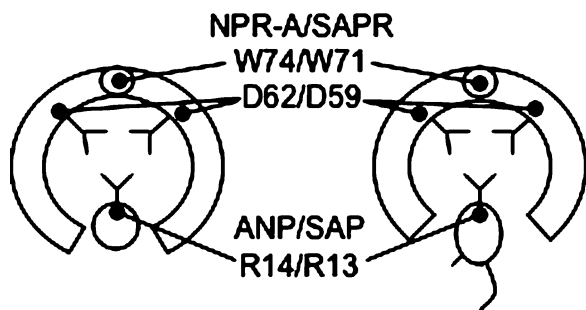
Figure 4 A schematic model of hormone and receptor binding. The rat atrial natriuretic peptide (ANP) and its receptor (NPR-A) are shown on the left. Asterosap and its receptor are shown on the right. Two identical receptor subunits are connected by ionic interactions at the top. A large hydrophobic tryptophan (W) residue from each subunit acts as a spring to secure the hinge. Interactions between the negatively charged aspartic acid (D) and the positively charged arginine (R) play important roles in the binding between ligands and the receptors. The molecular size of asterosap peptide (3700 Da, 34 amino acids) is larger than that of ANP (3100 Da, 28 amino acids).

We have identified homologous guanylate cyclase signalling systems from five different sea star species, presumably the first attempt in marine invertebrates from the North and South Pacific Ocean as well as the Atlantic Ocean. The cloned asterosap and guanylate cyclase sequences both exhibited speciation. Although the number of clones was limited caused by specimen availability, the analysed codon utilization did not have a strong bias. As we accumulate more clones, we will have the capacity to test evolutionary hypotheses, such as positive selection and co-evolution of asterosap and its receptor.

Acknowledgements

References

Bohmer, M., Van, Q., Weyand, I., Hagen, V., Beyermann, M., Matsumoto, M., Hoshi, M., Hildebrand, E. & Kaupp, U.B.
(2005). Ca²⁺ spikes in the flagellum control chemotactic behavior of sperm. *EMBO J.* **24**, 2741–52.

Felsenstein, J. (2005). PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Fricker, L. D. (1988). Activation and membrane binding of carboxypeptidase E. *J. Cell Biochem.* **38**, 279–89.

Gilbert, W. (2006). Sequence Analysis 1.6.0. Informagen.

Hook, V., Yasothornsrikul, S., Greenbaum, D., Medzihradszky, K.F., Troutner, K., Toneff, T., Bundey, R., Logrinova, A., Reinheckel, T., Peters, C. & Bogyo, M. (2004). Cathepsin L and Arg/Lys aminopeptidase: a distinct prohormone processing pathway for the biosynthesis of peptide neurotransmitters and hormones. *Biol. Chem.* **385**, 473–80.

Hoshi, M., Nishigaki, T., Ushiyama, A., Okinaga, T., Chiba, K. & Matsumoto, M. (1994). Egg-jelly signal molecules for triggering the acrosome reaction in starfish spermatozoa. *Int. J. Dev. Biol.* **38**, 167–74.

Kaupp, U.B., Solzin, J., Hildebrand, E., Brown, J.E., Helbig, A., Hagen, V., Beyermann, M., Pampaloni, F. & Weyand, I. (2003). The signal flow and motor response controling chemotaxis of sea urchin sperm. *Nat. Cell Biol.* **5**, 109–17.

Kaur, H. & Raghava, G.P. (2003). Prediction of beta-turns in proteins from multiple alignment using neural network. *Protein Sci.* **12**, 627–34.

Kawase, O., Minakata, H., Hoshi, M. & Matsumoto, M. (2005). Asterosap-induced elevation in intracellular pH is indispensable for ARIS-induced sustained increase in intracellular Ca²⁺ and following acrosome reaction in starfish spermatozoa. *Zygote* **13**, 63–71.

Kelley, L. A., MacCallum, R.M. & Sternberg, M.J. (2000). Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.* **299**, 499–520.

Kyte, J. & Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–32.

Liu, H., Force, T. & Bloch, K.D. (1997). Nerve growth factor decreases soluble guanylate cyclase in rat pheochromocytoma PC12 cells. *J. Biol. Chem.* **272**, 6038–43.

Lupas, A., Van Dyke, M. & Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* **252**, 1162–4.

Matsumoto, M., Briones, A.V., Nishigaki, T. & Hoshi, M. (1999). Sequence analysis of cDNAs encoding precursors of starfish asterosaps. *Dev. Genet.* **25**, 130–6.

Matsumoto, M., Solzin, J., Helbig, A., Hagen, V., Ueno, S., Kawase, O., Maruyama, Y., Ogiso, M., Godde, M., Minakata, H., Kaupp, U.B., Hoshi, M. & Weyand, I. (2003). A sperm-activating peptide controls a cGMP-signaling pathway in starfish sperm. *Dev. Biol.* **260**, 314–24.

Nakachi, M., Moriyama, H., Hoshi, M. & Matsumoto, M. (2006). Acrosome reaction is subfamily specific in sea star fertilization. *Dev. Biol.* **298**, 597–604.

Nishigaki, T., Chiba, K. & Hoshi, M. (2000). A 130-kDa membrane protein of sperm flagella is the receptor for asterosaps, sperm-activating peptides of starfish *Asterias amurensis*. *Dev. Biol.* **219**, 154–62.

Nishigaki, T., Chiba, K., Miki, W. & Hoshi, M. (1996). Structure and function of asterosaps, sperm-activating peptides from the jelly coat of starfish eggs. *Zygote* **4**, 237–45.

Ogawa, H., Qiu, Y., Ogata, C.M. & Misono, K.S. (2004). Crystal structure of hormone-bound atrial natriuretic peptide

receptor extracellular domain: rotation mechanism for transmembrane signal transduction. *J. Biol. Chem.* **279**, 28625–31.

Perriere, G. & Gouy, M. (1996). WWW-query: an on-line retrieval system for biological sequence banks. *Biochimie* **78**, 364–9.

Pollastri, G. & McLysaght, A. (2005). Porter: a new, accurate server for protein secondary structure prediction. *Bioinformatics* **21**, 1719–20.

Qiu, Y., Ogawa, H., Miyagi, M. & Misono, K.S. (2004). Constitutive activation and uncoupling of the atrial natriuretic peptide receptor by mutations at the dimer interface. Role of the dimer structure in signalling. *J. Biol. Chem.* **279**, 6115–23.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Plainview, New York: Cold Spring Harbor Laboratory Press.

Schwede, T., Kopp, J., Guex, N. & Peitsch, M.C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381–5.

Shiba, K., Tagata, T., Ohmuro, J., Mogami, Y., Matsumoto, M., Hoshi, M. & Baba, S.A. (2006). Peptide-induced hyperactivation-like vigorous flagellar movement in starfish sperm. *Zygote* **14**, 23–32.

Singh, S., Lowe, D.G., Thorpe, D.S., Rodriguez, H., Kuang, W.J., Dangott, L.J., Chinkers, M., Goeddel, D.V. & Garbers, D.L. (1988). Membrane guanylate cyclase is a cell-surface receptor with homology to protein kinases. *Nature* **334**, 708–12.

Steiner, D.F. (1998). The proprotein convertases. *Curr. Opin. Chem. Biol.* **2**, 31–39.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. & Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–82.

Vacquier, V.D. (1998). Evolution of gamete recognition proteins. *Science* **281**, 1995–8.

Ward, G.E., Brokaw, C.J., Garbers, D.L. & Vacquier, V.D. (1985). Chemotaxis of *Arbacia punctulata* spermatozoa to resact, a peptide from the egg jelly layer. *J. Cell Biol.* **101**, 2324–9.

Wilson, E.M. & Chinkers, M. (1995). Identification of sequences mediating guanylyl cyclase dimerization. *Biochemistry* **34**, 4696–701.