

Structure and function of sea urchin neuropeptides

Rowe, Matthew Lyndon

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author

For additional information about this publication click this link.

<https://qmro.qmul.ac.uk/jspui/handle/123456789/599>

Information about this research object was correct at the time of download; we occasionally make corrections to records, please therefore check the published record when citing. For more information contact scholarlycommunications@qmul.ac.uk

**STRUCTURE AND FUNCTION OF SEA URCHIN
NEUROPEPTIDES**

By

MATTHEW LYNDON ROWE

**Thesis submitted to the
University of London
for the degree of**

Doctor of Philosophy

September 2009

**School of Biological and Chemical Sciences
Queen Mary, University of London
Mile End Road, London, E1 4NS**

Section B: Declaration of word length and ownership of work

Name of candidate:

Degree for which entered:
(MPhil / PhD / MD(Res))

Date of thesis submission:

Title of thesis:

Name of principal supervisor:

Approximate word length of thesis:

1. Please tick one box:

- I confirm that the above thesis does not exceed the word limit prescribed in the Regulations.
 I confirm that the above thesis exceeds the word limit and has received a suspension of Regulations with attached confirmation.

2. I confirm that the work presented in the thesis is my own and all references are cited accordingly (if the thesis includes any work conducted jointly please attach a statement of the part played by the candidate, certified by the supervisor). I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

Candidate signature:

Date:

Supervisor's signature:

Date:

Submit this form completed with two bound copies and one electronic copy of your thesis to:

Research Degrees Office (Examinations)
E15 Queens Building
Queen Mary, University of London
Mile End Road
London
E1 4NS
+44 (0)20 7882 7942
researchdegrees@qmul.ac.uk

Abstract

The subject of this thesis is the identification and functional characterization of sea urchin neuropeptides. Neuropeptides are important mediators of neural signalling in all known animals with a nervous system, including bilaterians, ctenophorans, and cnidarians. Sea urchin neuropeptides are of particular interest for three significant reasons; echinoderms have a radically different secondarily-derived pentaradial body structure, sea urchins have served as model organisms for research into embryonic development, and thirdly because the genome of a sea urchin, the purple sea urchin *Strongylocentrotus purpuratus* (Stimpson, 1857) has been sequenced (Sodergren et al., 2006). Only one family of neuropeptides, the SALMFamides, has previously been characterized in all classes of the phylum Echinodermata.

The thesis reports the identification of putative neuropeptide GPCRs and at least seven novel sea urchin neuropeptide genes using genomic and Expressed Sequence Tag (EST) analysis. The novel sea urchin neuropeptides identified include putative homologues of vasotocin, the sea cucumber neuropeptide NGIWYamide, thyrotropin-releasing hormone, gonadotropin-releasing hormone, and calcitonin. A further three peptide precursor genes encoding peptides lacking strong homology to any known peptides were also identified and the peptides have been named GKamides and Pedal Peptide-like Neuropeptides. Two of the peptide precursor genes, those encoding peptides homologous to vasotocin and NGIWYamide, also each encode neurophysin domains, which have previously only been identified in association with vasopressin/oxytocin-like peptides.

Biochemical and pharmacological techniques were employed to investigate the occurrence and functions of the putative neuropeptides identified. These included mass spectroscopy and *in vitro* bioassays, the former to detect the putative novel neuropeptides identified in this study and the latter to investigate bioactivity of the peptides in sea urchins. The thesis provides evidence of the neural expression and bioactivity of novel sea urchin neuropeptides and contributes to our understanding of the role of neuropeptides in echinoderm physiology and behaviour.

Contents

List of Tables	9
List of Figures	10
Acknowledgements	13
List of Abbreviations	15
Dedication	16
<u>1: Introduction: Echinoderm Neuropeptides</u>	17
1.1 General introduction	18
1.2 Neuropeptides as neural signalling molecules: structural and functional diversity	19
1.3 Neuropeptide evolution: the FMRFamide-related Peptide Family	24
1.4 Neuropeptides and echinoderm neurobiology and physiology	26
1.5 The biochemical identification and characterization of echinoderm neuropeptides	31
1.6 The sea urchin genome project: an investigation of sea urchin neural genes	38
1.7 The identification of a novel SALMFamide neuropeptide precursor gene in the sea urchin genome	41
1.8 Aims	42
<u>2: The Identification of Putative Sea Urchin Neuropeptide Receptors</u>	44
2.1 Introduction: Neuropeptide GPCR structure and diversity	45
2.2 Methods: Bioinformatic identification of putative neuropeptide receptors	49
2.3 Results	50

2.3.1	Putative sea urchin neuropeptide GPCRs	51
2.3.2	Putative sea urchin neuropeptide GPCRs identified in cDNA libraries	53
2.4	Discussion	71
<u>3: NGFFFamide and Echinotocin: Novel Myoactive Neuropeptides Derived from Neurophysin-containing Precursors</u>		77
3.1	Introduction: identification of neuropeptide precursor genes in <i>Strongylocentrotus</i>	78
3.2	Methods	81
3.2.1	Bioinformatics: Identification of putative sea urchin neuropeptide genes using BLAST homology searching	81
3.2.2	Mass Spectrometry: Development of protocols for the identification of starfish SALMFamides	83
3.2.3	Methods: Immunohistochemistry protocols for the localization of echinotocin	87
3.2.3.1	Immunoblot assay protocol	88
3.2.3.2	Immunohistochemistry: Tissue fixation, sectioning and immunostaining	89
3.2.4	Methods: Pharmacological investigation of echinotocin and NGFFFamide myoactivity	91
3.3	Results	93
3.3.1	Identification of NGFFFamide	94
3.3.2	Identification of the neurophysin domain	95
3.3.3	Identification of echinotocin	95
3.3.4	Mass spectrometry: detection of sea urchin NGFFFamide	99
3.3.5	Results: Immunohistochemistry	105
3.3.6	Results: Characterization of NGFFFamide and echinotocin as potent myoactive peptides	109
3.3.6.1	Acetylcholine and carbachol	109

	3.3.6.2	Echinotocin and NGFFFamide	110
3.4		Discussion	118
	3.4.1	Sea urchin glycoprotein hormones	118
	3.4.2	Neurophysin-encoding precursors in the sea urchin genome	119
	3.4.3	Mass spectrometric analysis: Detection of a putative NGFFFamide peptide	125
	3.4.4	Discussion: Immunohistochemical investigation of echinotocin expression	126
	3.4.5	Discussion: Contractile myoactivity of echinotocin and NGFFFamide	127
		<u>4: Identification of Putative Neuropeptide Precursors from a Sea Urchin Radial Nerve Expressed Sequence Tag (EST) Library</u>	140
4.1		Introduction: identification of sea urchin neuropeptides using a cDNA library	141
4.2		Methods	142
	4.2.1	Identification of novel neuropeptide precursors	142
	4.2.2	Detection of the putative neuropeptides in sea urchin tissue using mass spectrometry	144
4.3		Results	146
	4.3.1	Results: Identification of putative sea urchin neuropeptide precursors	146
	4.3.2	Results: MS analysis of sea urchin body wall and visceral tissue	156
	4.3.3	Results: Pharmacology	162
4.4		Discussion	166
	4.4.1	SpurS8 & SpurS9: two peptides encoded by a second sea urchin SALMFamide neuropeptide precursor	166
	4.4.2	Sea urchin gonadotropin-releasing hormone-like peptide	169

4.4.3	Sea urchin thyrotropin-releasing hormone-like peptide	172
4.4.4	Sea urchin calcitonin/CGRP-like peptide	173
4.4.5	Two novel pedal peptide-like precursor genes	175
4.4.6	GKamide: a novel sea urchin peptide	176
<u>5: Discussion</u>		181
<u>References</u>		192
<u>Appendices</u>		
Appendix A:	Publications	215
Appendix B:	Putative sea urchin voltage-gated ion channel genes	216
Appendix C:	Human neuropeptides organized by receptor family	217
Appendix D:	Mass spectrometry expert parameter settings	220
Appendix E:	34 Putative sea urchin neuropeptide precursors	221

List of Tables

Table 1: Echinoderm neuropeptides (Class Asteroidea)	36
Table 2: Echinoderm neuropeptides (Class Holothuroidea)	37
Table 3: Echinoderm neuropeptides (Class Echinoidea)	43
Table 4: Putative G protein-coupled peptide receptors in <i>S. purpuratus</i>	54
Table 5: Putative <i>S. purpuratus</i> neurotransmitter and neuropeptide receptors identified in EST libraries	55
Table 6: Predicted and observed synthetic peptide masses of echinotocin and NGFFFamide detected using HPLC-MS	101
Table 7: Acetylcholine-induced contraction of sea urchin tissue preparations.	112
Table 8: Standardization of synthetic peptide elution points on an acetonitrile (ACN) gradient, including elution points for echinotocin and NGFFFamide	158
Table 9: Predicted and observed synthetic peptide masses of Sp-GnRH, Sp-TRH, GKamide, PPLN-1 detected using HPLC-MS	159

List of Figures

Figure 1: Human rhodopsin family beta group receptors and putative sea urchin homologues	56
Figure 2: Putative sea urchin peptide receptors homologous to human rhodopsin family beta group receptors	57
Figure 3: Human and putative sea urchin neuromedin B/bombesin, neuromedin U, neurotensin, and growth-hormone secretagogue-like receptors	58
Figure 4: Human and putative sea urchin AVP and GnRH-like receptors	59
Figure 5: Human and putative non-vertebrate tachykinin-like receptors	60
Figure 6: Human and putative non-vertebrate gonadotropin-releasing hormone and vasotocin-like receptors	61
Figure 7: Human and putative non-vertebrate cholecystokinin-like receptors	62
Figure 8: Human and putative non-vertebrate bombesin and neuromedin B-like receptors	63
Figure 9: Human and putative non-vertebrate orexin/hypocretin-like receptors	64
Figure 10: Human and putative non-vertebrate GPR54-like receptors	65
Figure 11: Human and putative non-vertebrate galanin-like receptors	66
Figure 12: Human and putative non-vertebrate GPR7 and GPR8-like receptors	67
Figure 13: Human and putative non-vertebrate melanin-concentrating hormone (MCH)-like receptors	68
Figure 14: Human and putative non-vertebrate gonadotropin and relaxin-like receptors	69
Figure 15: Human and putative non-vertebrate corticotropin-releasing hormone (CRH) and calcitonin/CGRP-like receptors	70
Figure 16: Two putative sea urchin bursicon-like prohormone sequences, compared to two <i>Drosophila</i> bursicon prohormone sequences	96

Figure 17: NGFFFamide precursor nucleotide and amino acid sequence in full	97
Figure 18: Echinotocin precursor nucleotide and amino acid sequence in full	98
Figure 19: SALMFamide S1 (MW=885.4, predicted MW=885.1) detected from tube foot and radial nerve tissue extract	102
Figure 20: SALMFamide S2 (MW=1275.5, predicted MW=1275.4) detected from tube foot and radial nerve tissue extract	103
Figure 21: Detection of NGFFFamide using HPLC-MS	104
Figure 22: Positive control tests for the AB1565 vasopressin antibody	107
Figure 23: Positive control tests for the AB1565 vasopressin antibody	107
Figure 24: Positive control tests for the PC234L vasopressin antibody	107
Figure 25: Sea urchin oesophagus transverse section prepared from paraffin-embedded tissue stained with rabbit anti-vasopressin	108
Figure 26: Acetylcholine induced contraction of <i>Echinus esculentus</i> 'semi-stripped' tube foot preparations measured using an isometric transducer	113
Figure 27: Carbachol induced contraction of <i>Echinus esculentus</i> 'semi-stripped' tube foot preparations measured using an isometric transducer	114
Figure 28: NGFFFamide causes contraction of oesophagus and 'unstripped' tube foot preparations from the sea urchin <i>Echinus esculentus</i> .	115
Figure 29: Echinotocin causes contraction of oesophagus and 'unstripped' tube foot preparations from the sea urchin <i>Echinus esculentus</i> .	116
Figure 30: NGFFFamide induced contraction of <i>Echinus esculentus</i> 'semi-stripped' tube foot preparations measured using an isometric transducer	117
Figure 31: Alignment of vasopressin/oxytocin-like peptides including sea urchin echinotocin and nematode 'nematocin'	136
Figure 32: ClustalX multiple alignment of vasopressin/oxytocin-like peptide precursor sequences	137

Figure 33: Neighbour-joining tree (with bootstrap values) based on a ClustalX multiple alignment of neurophysin (NP) sequences	138
Figure 34: Homology models of echinotocin (A) and NGFFFamide (B) neurophysin molecules	139
Figure 35: Identification of putative neuropeptide precursors from an adult sea urchin (<i>S. purpuratus</i>) radial nerve EST dataset	145
Figure 36: Putative neuropeptide precursor sequences (<i>spp-1</i> to <i>spp-7</i> ; A-G)	149
Figure 37: <i>Spp-1</i> ; the putative SpurS8 and SpurS9 neuropeptide precursor nucleotide and amino acid sequence in full	150
Figure 38: <i>Spp-2</i> ; the putative Sp-GnRH (pEVHHRFSGWRPG-NH ₂) neuropeptide precursor nucleotide and amino acid sequence in full	151
Figure 39: <i>Spp-3</i> ; the putative Sp-TRH (pEYPG-NH ₂) neuropeptide precursor nucleotide and amino acid sequence in full	152
Figure 40: <i>Spp-4</i> ; the putative Sp-CGRP-like neuropeptide precursor nucleotide and amino acid sequence in full	153
Figure 41: <i>Spp-5</i> ; the putative Sp-GKamide (consensus sequence ANMFRSRLRGK-NH ₂) neuropeptide precursor nucleotide and amino acid sequence in full	154
Figure 42: <i>Spp-6</i> ; the putative Sp-PPLN1 (consensus sequence GFNSGAMEPLGSGFI) neuropeptide precursor nucleotide and amino acid sequence in full	155
Figure 43: Detection of PPLN-1 consensus sequence peptide using HPLC-MS	160
Figure 44: Detection of Sp-TRH using HPLC-MS	161
Figure 45: Four representative traces illustrating Sp-GnRH (3 μM) induced contraction of <i>E. esculentus</i> oesophagus tissue	164
Figure 46: Four representative traces illustrating Sp-TRH (3 μM) induced contraction of <i>E. esculentus</i> tube foot tissue	165
Figure 47: A comparison of SALMFamide neuropeptide family sequences	178
Figure 48: A comparison of GnRH-like neuropeptide sequences	179
Figure 49: A comparison of calcitonin and CGRP-like peptide sequences	180

Acknowledgments

I would like to thank my supervisor Professor Maurice Elphick for his guidance, dedication, and enthusiasm. I would also like to thank Richard Melarange, for his invaluable help with preparing pharmacological bioassays, and Andrew Campbell and Greg Elgar, both of Queen Mary, University of London, for their feedback and advice over the duration of my project.

I would also like to thank Alan Scott, Queen Mary, University of London, for his help and guidance with mass spectrometry, Robert Burke, University of Victoria, British Columbia, for providing *Strongylocentotrus purpuratus* tissue extracts, and Michaela Egertová, Queen Mary, University of London, for her experience and help with immunohistochemistry.

I am also grateful to my parents and to the University of London Central Research Fund for their financial support and assistance in synthesizing the peptides investigated in this study, respectively.

List of Abbreviations

7TM	Seven transmembrane [domain]
ACh	Acetylcholine
ACN	Acetonitrile
ACTH	Adrenocorticotropic hormone
API	Atmospheric pressure interface
AVP	Arginine vasopressin
BACs	Bacterial artificial chromosomes
BLAST	Basic local alignment search tool
cDNA	copy Deoxyribonucleic acid
CAL	Calcitonin
CAPPS	Clone-array pooled shotgun sequencing
Carbachol	Carbamylcholine
CCK	Cholecystokinin
CGRP	Calcitonin gene-related peptide
CID	Collision-induced dissociation
CNS	Central nervous system
CPS	Counts per second
CRH	Corticotropin-releasing hormone
DAB	3,3'-Diaminobenzidine
DH	Diuretic hormone
DNA	Deoxyribonucleic acid
DPX	Di-n-butyl phthalate in xylene
DTU	Technical University of Denmark
EST	Expressed sequence tag
EtOH	Ethanol
FaRP	FMRamide-related peptide
FSH	Follicle-stimulating hormone
GABA	γ -amino butyric acid
GAP	Gonadotropin associated peptide
GH	Growth hormone
GI	Gene index
GnRH	Gonadotropin-releasing hormone
GPCR	G-protein coupled receptor
GPS	GPCR proteolytic site
GRAFS	Glutamate rhodopsin adhesion frizzled secretin [system of GPCR nomenclature]
GSS	Gonad-stimulating substance
HMM	Hidden markov model
HPLC-MS	High performance liquid chromatography mass spectrometry
IGF	Insulin-like growth factor
IHC	Immunohistochemistry
IR	Immunoreactivity
JH	Juvenile hormone
LH	Luteinizing hormone
LTP	Long-term potentiation
MALDI-TOF	Matrix-assisted laser desorption/ionization time of flight [mass spectrometry]
MCH	Melanin-concentrating hormone

MECA	Melanocortin, endothelin differentiating, cannabinoid, and adenosin [binding receptors]
MIT	Massachusetts Institute of Technology
mRNA	Messenger ribonucleic acids
MS	Mass spectrometry
NCBI	National Centre for Biotechnology Information
NGS	Normal goat serum
NO	Nitric oxide
NP	Neurophysin
ORF	Open reading frame
PBS	Phosphate buffered saline
PBS-X	PBS containing 0.2% Triton X-100
PDB	Protein Data Bank (also used as a file extension)
PFA	Paraformaldehyde
POMC	Proopiomelanocortin
PPLN	Pedal peptide-like neuropeptide
PSD	Post-source decay
PVN	Paraventricular nucleus
RAMP	Receptor activity-modifying protein
RCP	Receptor coupled protein
RNA	Ribonucleic acid
S1	SALMFamide 1
S2	SALMFamide 2
SON	Supraoptic nucleus
spp	<i>Strongylocentrotus purpuratus</i> peptide
TRH	Thyrotropin-releasing hormone
TSH	Thyroid-stimulating hormone
UTR	Untranslated region
VGIC	Voltage-gated ion channels
WGS	Whole genome shotgun [sequencing]

To my family and friends

Chapter 1

INTRODUCTION: ECHINODERM NEUROPEPTIDES

- 1.1** General introduction
- 1.2** Neuropeptides as neural signalling molecules: structural and functional diversity
- 1.3** Neuropeptide evolution: the FMRFamide-related Peptide (FaRP) Family
- 1.4** Neuropeptides and echinoderm neurobiology and physiology
- 1.5** The biochemical identification and characterization of echinoderm neuropeptides
- 1.6** The sea urchin genome project: an investigation of sea urchin neural genes
- 1.7** The identification of a novel SALMFamide neuropeptide precursor in the sea urchin genome
- 1.8** Aims

1.1 General introduction

The subject of this thesis is the identification and functional characterization of sea urchin neuropeptides. Neuropeptides are important mediators of neural signalling in all known animals with a nervous system, including all bilaterians, ctenophorans, and cnidarians. Sea urchin neuropeptides are of particular contemporary interest for three significant reasons; echinoderms have a radically different secondarily-derived pentaradial body structure, sea urchins have served as model organisms for research into embryonic development and gene regulatory networks, and thirdly because the genome of a sea urchin, the purple sea urchin *Strongylocentrotus purpuratus* (Class Echinoidea; Family Strongylocentrotidae) (Stimpson, 1857) was recently sequenced (Sodergren et al., 2006).

One goal of the thesis has been to utilize the genome to rejuvenate echinoderm neural signalling research. This part of my thesis was completed in collaboration with other researchers with the objective of characterizing the echinoderm nervous system through genomic analysis. The results of the collaboration have been published (Burke et al., 2006). To validate the results of my investigation of sea urchin neuropeptides, further biochemical and pharmacological techniques were employed. These included mass spectrometry and *in vitro* bioassays, the former to detect putative novel neuropeptides identified in this study and the latter to investigate the bioactivity of the peptides in sea urchins.

1.2 Neuropeptides as neural signalling molecules: structural and functional diversity

Neuropeptides are short peptides secreted by nerve cells that have diverse cellular functions in neural signalling. A particular neuropeptide may fulfil roles as a neurotransmitter, a neuromodulator or a neurohormone at different loci throughout an organism. Neurohormones are neurally secreted and transported by the circulatory system to their site of action. Neuromodulators act at the synapse by altering the electrochemical properties of the postsynaptic or presynaptic cell, whereas neurotransmitters are defined as molecules essential for the propagation of a neural signal at synapses (Strand, 1999).

Traditionally endocrinology involved the physiological investigation of hormones secreted systemically into an organism's circulation and acting at sites distant from their site of synthesis and secretion. Neurohormones are similarly secreted, but the same hormones may act as neuromodulators in the extracellular plasma or as neurotransmitters pre or post-synaptically. Neuroendocrine research in vertebrates began with the isolation of gonadotropin-releasing hormone (GnRH) and thyrotropin-releasing hormone (TRH) as hormone releasing factors secreted by neurons originating in the hypothalamus and targeting the anterior pituitary (Boler et al., 1969). Subsequently, invertebrate neuroendocrine pathways were investigated resulting in identification of the arthropod neuropeptide proctolin (Starratt and Brown, 1975) and the molluscan cardioexcitatory peptide, FMRFamide, from bivalve ganglia in 1977 (Price and Greenberg, 1977; Walker, 1986). Subsequent to both important discoveries,

neuropeptides have been found to act as important neurotransmitters, neuromodulators, and neurohormones throughout the animal kingdom (Greenberg and Price, 1983; O'Shea and Schaffer, 1985; De Lange et al., 1997; Strand, 1999; Nassel, 2002). Neuropeptides are even thought to be the primary neurotransmitters in the nervous system of cnidarians, a metazoan phylum basal to bilaterians, suggesting an ancient and evolutionarily conserved role in the mediation of neural signalling (Grimmelikhuijzen et al., 1999; Bosch and Fujisawa, 2001).

The full complement of interneuronal mediators, including neuromodulators and neurotransmitters is by no means limited to peptides and includes monoamines, ions, amino acids, nucleotides, acetylcholine, and γ -amino butyric acid (GABA). These molecular mediators communicate by altering the electrochemical properties of their target cell. The membrane potential of a cell may be indirectly altered via activation of metabotropic G Protein-Coupled Receptors (GPCRs) or directly via activation of ionotropic ligand-gated ion channels. Furthermore, Voltage-Gated Ion Channels (VGICs), stimulated by a change in electrochemical potential, and stretch-activated ion channels, stimulated by a change in the mechanical state of the cell membrane, are also important components to neural signalling, allowing the gated movement of anions and cations, thereby altering the distribution of charge across a neuronal cell membrane.

Because neuropeptides are modified amino acid polymers they are highly diverse in structure. Furthermore, the evolution and functional divergence of a family of neuropeptides is reflected in the substitution and mutation of

peptide sequence structure. Two mammalian neuropeptides, oxytocin (CYIQNCPLG-NH₂) and vasopressin (CYFQNCPRG-NH₂), arose through gene duplication of an ancestral neuropeptide precursor gene similar to the precursor encoding Arg-vasotocin (CYIQNCPRG-NH₂) (Ruppert et al., 1984). Both mammalian neuropeptides have acquired additional roles as neurotransmitters and neuromodulators in the central nervous system alongside their role as neurohormones secreted from cells of the supraoptic nucleus (SON) and the paraventricular nucleus (PVN) of the hypothalamus. Vasopressin has a number of diverse roles including an involvement in water balance via renal tubule receptors (Dibas et al., 1998), ion transport at the blood-brain barrier (Jojart et al., 1984; O'Donnell et al., 2005), and anxiety and aggression in rodents (Bielsky et al., 2004; Blanchard et al., 2005). In the central nervous system vasopressin has been found to potentiate noradrenaline signalling by membrane depolarization and to modulate glutamate excitation and release (Disturnal et al., 1987; Brinton and McEwen, 1989; Syed et al., 2007). Van den Hooff et al. (1989) found vasopressin to be necessary for maintaining Long-Term Potentiation (LTP) between cells of the fimbria fornix and the lateral septum.

Oxytocin has roles as both a myoactive neurohormone important for lactation and parturition in mammals and in the regulation of behaviour associated with trust and pair bonding (Kosfeld et al., 2005). Like vasopressin the neurohormone is expressed in cells originating in the PVN and SON and secreted from the pituitary gland. It is involved in the induction of lactation and uterine contraction, targeting oxytocin GPCRs typically coupled to a phospholipase C-mediated transduction pathway

(Gimpl and Fahrenholz, 2001). However, the peptide is also expressed in the central nervous system. For example the peptide targets vagal motoneurons as an excitatory neuromodulator through the generation of inward sodium currents (Dreifuss et al., 1992).

Unlike hormones derived from lipids and amino acids, neuropeptides are processed in the rough endoplasmic reticulum and packaged and modified in the Golgi apparatus, before transportation in vesicles to the site of Ca^{2+} -dependent exocytosis (Thorndyke and Goldsworthy, 1988; Strand, 1999; Hartenstein, 2006). The translated preprohormone is processed into a prohormone sequence in the rough endoplasmic reticulum by cleavage of the N-terminal signal sequence. The peptide precursors are packaged into 100-300 nm sized vesicles during transportation from the *cis* face to the *trans* face of the Golgi apparatus. The vesicles are several times larger than the vesicles typically used for storing and transporting classical neurotransmitters, such as acetylcholine. The precursors are then processed by prohormone convertases into short peptides, before enzymatic modification which may include pyroglutamate formation and amidation. C-terminal α -amidation, which requires a terminal glycine residue, protects the peptide from degradation. The inactivation of secreted peptides by peptidases occurs in the plasma.

Neuropeptides may be secreted by non-neuronal cell types: for example, somatostatin is also secreted by the gut in mammals (Pradayrol et al., 1978). Many other neuropeptides are also secreted by gut tissue, including galanin, calcitonin gene-related peptide, gastrin-cholecystikinin,

somatostatin, and the tachykinins (Walsh and Dockray, 1994). Although somatostatin is widely expressed in the central nervous system, its name derives from its inhibitory role in the regulation of growth hormone (GH) secreted by the anterior pituitary. Like both vasopressin and oxytocin, somatostatin is also secreted by cells of the PVN and released from the median eminence. The peptide has an atypical precursor structure from which the peptide sequence is differentially processed into two different forms, as opposed to peptide sequence variation in tandem; one of 14 residues and the other of 28, which may either bind the same receptor, albeit with different affinities and activities, or may bind alternate somatostatin receptor subtypes (Epelbaum et al., 1994; Patel, 1999).

Neuropeptide precursors often encode a tandem series of structurally similar peptides with different potencies and functional properties. Insect allatotropin and allatostatin peptides, functional homologues to somatostatin in mammals, respectively stimulate and inhibit the corpora allata, regulating the release of juvenile hormone (JH). Insect allatostatin precursors often encode several peptides of variable myotropic potency (Bendena et al., 1999; Nichols et al., 2002; Taghert and Veenstra, 2003; Veenstra, 2009). The Allatostatin-A precursor encodes as many as 14 peptides, defined by the C-terminal sequence YXFGLamide, in the form of multiple tandem repeats (East et al., 1996; Vanden Broeck et al., 1996; Belles et al., 1999; Stay and Tobe, 2007).

Neuropeptides are likely to have acquired diversity through the processes that drive gene evolution, including gene duplication, whole-genome

duplication events, and random mutation. Although the full sequence of a neuropeptide precursor gene may show extensive sequence variation, the amino acid structures of neuropeptides themselves are well-conserved and the small but significant variations in amino acid sequence have often been shown to coincide with novel functional variation and expansion within a taxon. Furthermore, conservation of neuropeptide structure across several taxa may also be correlated with conservation of function.

1.3 Neuropeptide evolution: the FMRFamide-related Peptide family

The identification of molluscan FMRFamide, and the subsequent identification of structurally related peptides in other invertebrates, has led to the classification of a family of peptides defined by a C-terminal amidated Arginine-Phenylalanine (RF) sequence, the FaRP (FMRFamide-related peptide) family (Dockray, 2004). An endogenous opioid encoded by the Proenkephalin A precursor, [Met⁵]enkephalin-Arg⁶-Phe⁷, also possesses an amidated Arg-Phe C-terminus, which may be indicative of an evolutionary relationship between RFamide peptides and enkephalins or may reflect convergent molecular evolution (Greenberg et al., 1983). Opioid peptide homologues have not yet been identified in invertebrates. Their absence, and the presence of highly diversified vertebrate FaRP gene family members, suggests the opioid peptides arose late in evolutionary history.

The vertebrate endogenous opioids include endorphins (e.g. β -endorphin, 31 residues), enkephalins (e.g. Met and Leu-enkephalin, 5 residues each),

and dynorphins (C-terminal extensions of Leu-enkephalin) derived from three different precursor genes in the human genome, respectively proopiomelanocortin (POMC), proenkephalin, and prodynorphin (Strand, 1999). Endogenous opioids act presynaptically by inhibiting or disinhibiting synaptic output and neurotransmitter release, postsynaptically as a neurotransmitter, and interestingly as a co-transmitter in glutamatergic hippocampal pathways activated by high frequency impulses (Morris and Johnston, 1995; Olson et al., 1995).

Unlike endogenous opioids, the RFamide family of peptides have been identified in a number of diverse bilaterian phyla, including cnidarians (Darmer et al., 1991), nematodes (Li et al., 1999a), arthropods (Vanden Broeck, 2001b; Mousley et al., 2004), platyhelminths (Mousley et al., 2004) and vertebrates (Thorndyke and Goldsworthy, 1988; Strand, 1999; Dockray, 2004; Chartrel et al., 2005). The seven human RFamides include neuropeptide AF/FF, prolactin-releasing peptide, neuropeptide SF/VF, kisspeptin, and 26RFamide, and are encoded by five FaRP precursor genes *farpl-5*. They form a small disparate group for which roles in nociception, vascular regulation, hormone secretion, appetite and feeding behaviour have been identified (Allard et al., 1995; Panula et al., 1999; Samson et al., 2003; Dockray, 2004; Takayasu et al., 2006). A role for RFamide peptides in feeding behaviour is not limited to vertebrates. In an example of functional homology and structural modification accompanied by variation in response, a study of nematode behaviour by de Bono and Bargmann (1998) demonstrated that a single amino acid variation in a *C. elegans*

RFamide receptor produces two behavioural phenotypes distinguishable by social and asocial feeding.

RFamide peptides have been localized in sensory neurons in both invertebrate and vertebrate organisms, including cnidarians (Martin, 1992; Mackie et al., 2003), the nematode *C. elegans* (Troemel et al., 1995; Rogers et al., 2003), the annelid *Platynereis* (Tessmar-Raible et al., 2007), the mollusc *Ischnochiton* (Voronezhskaya et al., 2002) and the zebrafish *Danio* (Pinelli et al., 2000). The RFamidergic sensory neurons are localized in morphologically homologous neurosecretory systems in both vertebrates and invertebrates, suggesting that the peptides have a cross-phyletically conserved role in the integration of sensory information and the regulation of behaviour (Hartenstein, 2006; Tessmar-Raible et al., 2007).

1.4 Neuropeptides and echinoderm neurobiology and physiology

The phylum Echinodermata (~7,000 species) originated from a Precambrian epibenthic ancestor, attained maximal class diversity during the mid-Palaeozoic era and now encompasses five modern classes, including, in order of species diversity: brittle stars and basket stars (Class Ophiuroidea), starfish (Class Asteroidea), sea cucumbers (Class Holothuroidea), sea urchins (Class Echinoidea), and sea lillies or feather-stars (Class Crinoidea) (Brusca and Brusca, 2003; Southward and Campbell, 2006).

Defining echinoderm characteristics include the absence of cephalization, the adaptation of radial symmetry upon metamorphosis from bilaterally symmetrical larvae, a unique water vascular system, and an endoskeleton composed of calcareous ossicles, which in some classes is fused to form a test. The sessile crinoids diverged earliest from the other echinoderm classes and they are identified by the primitive sessile characteristic of an aboral stalk attached to the substrate. The remaining classes are thought to have derived from a common motile ancestor. Sea urchins, sea cucumbers, and brittle stars are commonly grouped together on the basis of closed interambulacral grooves in opposition to the supposed ancestral characteristics of open grooves retained in extant starfish species (Brusca and Brusca, 2003).

Notwithstanding the absence of a truly closed circulatory system as exemplified in vertebrates, echinoderm neuropeptides acting as neurotransmitters, neuromodulators, and neurohormones are likely to have important roles in the regulation of echinoderm behaviour and physiology. Because of a highly diversified morphology and physiology, the actions and functions of homologous and novel echinoderm neuropeptides will be an important area of investigation for future animal physiologists, marine biologists, and comparative neurobiologists. Moreover, there are a number of reasons why echinoderms are of particular interest for neuropeptide research.

Firstly, adult echinoderms are unique in the animal kingdom in having a pentaradial morphological organization, or *bauplan*, which is both

evolutionarily and developmentally derived from bilateral symmetry (Burke et al., 2006). As mediators of neural signalling, neuropeptides may be involved in the regulation of behavioural and physiological coordination in the context of a pentaradial nervous system, including a novel system of motor coordination that is necessary for a foraging organism lacking cephalization. A peptidergic nervous system is also present in bilaterally organized echinoderm larvae and may be involved in the regulation of behaviour before metamorphosis (Byrne and Cisternas, 2002). Hence localization of neuropeptide expression in echinoderm larvae will be useful in the investigation of neurochemical signalling prior to the morphological transition from bilateral to pentaradial symmetry.

Secondly, echinoderms and vertebrates have been found to possess synapomorphies previously thought to be unique to vertebrates. Together with the phyla Hemichordata and Xenoturbellida, the phylum Echinodermata forms a sister clade to the craniates, urochordates, and cephalochordates (Bromham and Degnan, 1999; Bourlat et al., 2006; Dunn et al., 2008). These taxa comprise all known extant deuterostomes. The genomes of non-vertebrate chordates, including urochordate (*Ciona intestinalis*) and cephalochordate (*Branchiostoma floridae*) representative species, have also been recently sequenced, providing genomic data for organisms that evolved before the evolutionary transition to taxa defined by a vertebral column and a cranium (Delsuc et al., 2006). Due to the proximity of the phylum to chordates, and a shared mode of deuterostomian embryonic development, sea urchins have historically become a valuable model organism for research into developmental biology. Therefore,

comparative analysis of echinoderms and chordates provides a basis for identifying shared deuterostomian traits as well as novel characteristics that define echinoderm evolution, including neuropeptide signalling pathways and their receptors.

Thirdly, neuropeptides may be involved in the remarkable processes of autotomy and regeneration demonstrated by echinoderms (Moss et al., 1994; Thorndyke et al., 2001). Peptides have been found to induce the softening of connective tissue which accompanies evisceration and autotomy in starfish and sea cucumbers prior to tissue regeneration (Smith and Greenberg, 1973; Mladenov et al., 1989; Birenheide et al., 1998; Koob et al., 1999; Byrne, 2001; Thorndyke et al., 2001; Wilkie, 2001).

Pedicellariae attached to the test in sea urchins, appendages of the water vascular system including tube feet, and even entire starfish arms are capable of complete regeneration. Even more remarkably, species from the asteroid genus *Linckia* are capable of complete regeneration of the organism from a single limb (Edmondson, 1935).

Fourthly, echinoderms possess a body wall or test which is unique in being capable of rapidly and reversibly changing its mechanical properties (Wilkie, 2005). The mutable properties of the body wall and its associated appendages have been investigated in detail and neuropeptides, including stichopin identified in the holothurian *Apostichopus japonicas*, are thought to play a role as regulators of the unusual and unique rapid transitions in mechanical state of the dermis (Birenheide et al., 1998; Tamori et al., 2007).

Lastly, neural signalling molecules, including the putative neuropeptides identified in this study, may have a role in synchronization and integration of echinoderm sensory and motor pathways. Although the nervous system is decentralized, dispersed, and segregated, echinoderms are capable of coordinated behavioural and sensory integration (Smith, 1945; Kerkut, 1954; Smith, 1965; Reese, 1966; Dale, 1999). There are three principal nervous systems in the echinoderm; the ectoneural, including both sensory and motor neurons; the hyponeural, which typically runs parallel to the ectoneural system and which may exclusively contain efferent motor neurons; and the typically aboral entoneural system, which is confined to asteroids (starfish) and crinoids (featherstars). A fourth is considered to innervate the enteric system (García-Arrarás et al., 2001).

Studies of echinoderm behaviour have demonstrated coordination of object avoidance and locomotion through tube foot activity (Migita et al., 2005), of escape and directional manoeuvring (Dale, 1999), of spine and pedicellariae activity on the test (Smith, 1965; Peters, 1981; Campbell, 1983) and rhythmic coordination and circadian control of reproductive behaviour (Okada et al., 1984; Coppard and Campbell, 2005). Some studies have provided evidence of echinoderms demonstrating more than just innate behaviour (Valentinčič, 1983; Shul'gina, 2006). In Shul'gina's study starfish demonstrated long-term inhibition of movement in response to a negative stimulus, an adaptation which interestingly disappeared after severance of the nerve ring and two ambulacra, or rows of tube feet, across the madreporite interambulacrum. Further to the observation of an ability to

alter innate behaviour, which in itself implicates a broader concept of behavioural evolution through the adaptation of inhibition, the concept of self-organization has been suggested to explain the flexibility of innate motor behaviour in invertebrates lacking a central organizing locus or "brain" (Haken, 1983; Kelso, 1995). Coordinated movement may be reliant on processes of self-organization, as suggested in a study of stable phases of starfish tube foot and arm locomotion and obstacle avoidance by Migita et al (2005). An alternate explanation for locomotory coordination, observed at the level of both the tube foot and whole organism, may be the presence of rhythm generators sited at the gonopore leading to the gonad which regulate synchronization of gonadal contraction, as suggested in a study by Okada et al. (1984).

1.5 The biochemical identification and characterization of echinoderm neuropeptides

The first echinoderm peptide hormone to be identified was gonad-stimulating substance (GSS), identified in radial nerve extracts of the starfish *Asterias forbesi* (Class Asteroidea, Family Asteriidae) (Chaet and McConnaughy, 1959; Kanatani et al., 1971). Gonad-stimulating substance was found to indirectly induce spawning and oocyte maturation by targeting follicle cells and was the first invertebrate gonadotropin identified (Kanatani et al., 1971; Shirai et al., 1972; Kanatani, 1979; Kanatani, 1983; Shirai, 1986). The structure of gonad-stimulating substance, a peptide heterodimer of approximately 47 residues, has recently been determined following purification from the radial nerve cords of the starfish *Asterina*

pectinifera (Class Asteroidea, Family Asterinidae) and comprises two peptide chains of 24 and 19 residues each, potentially separated by three disulfide bonds and referred to as GSS chains A and B, respectively (Shirai et al., 1987; Kato et al., 2009; Mita et al., 2009) (Table 1).

However, the first echinoderm neuropeptides to be fully sequenced were detected in the starfish *Asterias rubens* and *Asterias forbesi* using antisera raised against molluscan FaRP cardioexcitatory peptide pQDPFLRFamide (Elphick et al., 1991a; Elphick et al., 1991b). The starfish neuropeptides, GFNSALMFamide (S1) and SGPYSFN SGLTFamide (S2), were both denominated SALMFamides after the C terminal motif of S1 (SALMFamide 1) (Table 1). Subsequently two more SALMFamides, GFSKLYFamide and SGYSVLYFamide, were purified from intestinal extracts of the sea cucumber *Holothuria glaberrima* (Class Holothuria) and localized in the intestine using an antiserum raised against the same molluscan FaRP-related peptide (Díaz-Miranda et al., 1992) (Table 2). The peptides identified were subsequently united and defined by the presence of a C-terminal motif, S_x(L/F)_xFamide (where *x* is variable).

The SALMFamide neuropeptides were found to have significant roles in muscle relaxation alongside nitric oxide, a well-documented muscle relaxant and neurotransmitter in vertebrates and invertebrates (Palmer et al., 1987; Elphick et al., 1993; Díaz-Miranda and García-Ararrás, 1995; Elphick et al., 1995; Boeckxstaens and Pelckmans, 1997; Melarange et al., 1999; Elphick and Melarange, 2001; Melarange and Elphick, 2003). The SALMFamide neuropeptides were localized to the radial nerve cord and

tube foot ectoneural nerve plexus of *A. rubens* using monoclonal and polyclonal antibodies (Newman et al., 1995), and an S2-specific antiserum CLII (Elphick et al., 1995). Pharmacological investigation of the action of the peptides on starfish muscle preparations found S2 to be as much as ten times more potent than S1 at inducing relaxation of the starfish cardiac stomach (Elphick et al., 1995).

A different strategy for the isolation and identification of echinoderm neuropeptides was adopted by Iwakoshi et al. (1995). Radial longitudinal muscle and intestinal preparations from the Japanese sea cucumber *Apostichopus japonicus* (Class Holothuroidea, Family Stichopodidae) were used to test for the presence of myoactive peptides in body wall extracts of the same species (Iwakoshi et al., 1995; Ohtani et al., 1999). Amongst the fourteen peptides identified were a further two members of the SALMFamide neuropeptide family, GYSPFMFamide and FKSPFMFamide, and both peptides caused relaxation of sea cucumber muscle preparations consistent with the pharmacological activity of SALMFamides S1 and S2 in the starfish (Ohtani et al., 1999) (Table 2).

Additionally, SALMFamides have been shown to be involved in more diverse roles including the modulation of luminescence in the brittle star *Amphipholis squamata* (Class Ophiuroidea, Family Amphiuridae) and the regulation of GSS secretion in the starfish *Asterina pectinifera* (Bremaeker et al., 1999; Mita et al., 2004).

Several other neuropeptides were identified by Iwakoshi et al. (1995), all of which appear to be unique to the echinoderm phylum in sequence and structure (Table 2). Many of the peptides identified by Iwakoshi et al. (1995) had indirect effects on muscle contractility, either potentiating or inhibiting electrically-evoked contraction of sea cucumber muscle preparations (Iwakoshi et al., 1995; Ohtani et al., 1999). One of the peptides identified, NGIWYamide, was found to cause contraction of sea cucumber longitudinal body wall muscle, tentacles, and intestine (Iwakoshi et al., 1995; Inoue et al., 1999; Ohtani et al., 1999). NGIWYamide immunoreactivity was observed in the body wall dermis, intestine, tentacles, and tube feet, loci supporting pharmacological evidence of the effects of NGIWYamide on contraction of the same tissues (Inoue et al., 1999). Saha et al (2006) investigated the expression and function of NGIWYamide-like peptides in the starfish *Asterina pectinifera*. Similar to the effects observed with sea cucumber preparations, NGIWYamide caused contraction of starfish tube foot preparations. Furthermore, NGIWYamide-like immunoreactivity was found in the starfish radial nerve cords and tube foot nerve plexus (Saha et al., 2006). These combined studies of NGIWYamide neuropeptide function and expression suggest a role for the neuropeptide family in regulating muscle contraction.

More recently NGIWYamide was identified as one of two peptides, the other being QGLFSGVamide, purified from buccal nerve ring and longitudinal radial nerve from the sea cucumber *Apostichopus japonicus* (Kato et al., 2009). The two peptides were isolated on account of their ability to induce oocyte maturation and spawning. Surprisingly, a synthetic

peptide with the sequence NGLWYamide (ie. with the isoleucine residue replaced with a leucine residue) was found to be ten to one hundred times more potent than the endogenous peptide, NGIWYamide (Kato et al., 2009).

Table 1: Asteroid (starfish) neuropeptides.

Name	Sequence
GSS ¹ :	
A Chain	SEYSGIASYCCLHGCTPSELSVVC
B Chain	EKYCDDDFHMAVFRCAVS
S1 ²	GFNSALMF-NH ₂
S2 ²	SGPYSEFNSGLTF-NH ₂

¹ *Asterina pectinifera* (Mita et al., 2009); ² *Asterias rubens* (Elphick et al., 1991a).

Table 2: Holothuroid (sea cucumber) neuropeptides.

Name	Sequence
SWYG-1 ¹	SWYGSLG
SWYG-2 ¹	SWYGTLG
SWYG-3 ¹	SWYGSLA
Holokinin-1 ¹	PLGYMFR
Holokinin-2 ¹	PLGM(O)FR
Holokinin-3 ¹	PLGY(Br)M(O)FR
Holokinin-1(3-7) ¹	GYMFR
NGIWYamide ¹	NGIWY-NH ₂
Stichopin ¹	DRQGWPCYDSKGNKYC
Sticho-MFamide-1 ¹	GYSPEMF-NH ₂
Sticho-MFamide-2 ¹	FKSPEMF-NH ₂
KIamide-9 ¹	KHKTAYTGI-NH ₂
GLRFA ¹	GLRFA
GN-19 ¹	GGRLPNYAGPPRMPWLIHN
-	GYI ²
-	MPMNPADYFSRGTVYIPTRDS ²
-	FDWHALLH ²
-	GPSANSFTYGRCADDR ²
-	KGKQKGAGRGKGN ²
-	APHAIRPPSG ²
QGLFSGVamide ³	QGLFSGV-NH ₂
GFSKLYFamide ⁴	GFSKLYF-NH ₂
SGYSVLYFamide ⁴	SGYSVLYF-NH ₂

¹*Apostichopus japonicus* (Iwakoshi et al., 1995); ²(Ohtani et al., 1999); ³(Kato et al., 2009); ⁴*Holothuria glaberrima* (Díaz-Miranda et al., 1995).

1.6 The sea urchin genome project: an investigation of sea urchin

neural genes

New opportunities to identify and characterize echinoderm neuropeptides have emerged recently with the sequencing of an echinoderm genome from sperm from a single male specimen of the north-east Pacific sea urchin *Strongylocentrotus purpuratus* (Stimpson, 1857) (Order Echinoidea; Family Strongylocentrotidae) (Sodergren et al., 2006). *S. purpuratus* is the first representative echinoderm to have its whole genome sequenced. The sea urchin was chosen for sequencing because of its established role as a model organism for research on embryological development (Briggs and Wessel, 2006). The genome was sequenced using Clone-Array Pooled Shotgun Sequencing (CAPSS) (Cai et al., 2001), a development of the whole-genome shotgun (WGS) sequencing strategy (Weber and Myers, 1997; Green, 2001), using bacterial artificial chromosomes (BACs) as cloning vectors (Sodergren et al., 2006).

Approximately 28,000 genes were identified and assembled into an official gene set from the genomic data using the Glean3 gene prediction algorithm (Sodergren et al., 2006; Elsik et al., 2007). The Glean3 gene predictions have subsequently been labelled using the epithet 'SPU' and are publicly accessible online (<http://annotation.hgsc.bcm.tmc.edu/Urchin/cgi-bin/pubLogin.cgi>). The annotated genome can be accessed via NCBI, the Sea Urchin Genome Project at the Baylor College of Medicine website (<http://www.hgsc.bcm.tmc.edu/project-species-o-Strongylocentrotus%20purpuratus.hgsc?pageLocation=Strongylocentrotus>

%20purpuratus), or SpBase (<http://www.spbase.org/SpBase/>) (Cameron et al., 2009).

The genome was analysed for genes important for neural growth and signalling in both the adult and embryo (Burke et al., 2006), providing data with which to supplement earlier physiological and anatomical studies of the echinoderm nervous system (Hyman, 1955; Bullock and Horridge, 1965; Smith, 1965).

As part of the bioinformatic analysis of the sea urchin nervous system, an investigation was made into the identification of sea urchin voltage-gated ion channels (VGICs), a group of membrane-bound receptors important for neural electrical signalling. Like GPCRs, VGICs have an ancient common ancestor and members have even been identified in prokaryotes (Hille, 2001; Gutman et al., 2005). BLAST searching the sea urchin genome using vertebrate VGIC sequences as query entries resulted in the identification of eleven K⁺ channels, four Ca²⁺ channels and one Na⁺ channel (Appendix B) (Burke et al., 2006). Sodium VGICs are present in both protostomes and deuterostomes, although they are thought to have arisen from T-type calcium VGICs (Burke et al., 2006). However, action potentials in the sea urchin nervous system may be calcium-based (Cobb and Moore, 1989) and ion specificity is determined by small but significant modifications of residue structure in the ion pore. Thus, verification of ion specificity requires further bioinformatic analysis accompanied by physiological investigation.

Although echinoderm morphological studies have shown a lack of synaptic structure in which pre-synaptic terminals are rarely seen in contact with post-synaptic terminals (Cobb and Laverack, 1967), the process of synaptic formation, and its associated genes, is a characteristic of eumetazoans and a defining feature of the nervous system. A study by Sakarya et al. (2007) reported the conservation of synaptic protein scaffold structure in the genome of the poriferan demosponge *Amphimedon queenslandia*, an animal phylum basal to the eumetazoans and characterized by the absence of a nervous system. Investigation of the sea urchin genome also found an absence of ionotropic glutamate receptor subunit genes, important to long term potentiation and synaptic plasticity in vertebrates. In contrast to vertebrate genomes, the sea urchin genome also lacked genes encoding gap junction proteins and receptors for cannabinoid and lysophospholipid neuromodulators, suggesting a predominance of signalling through chemical synapses (Burke et al., 2006).

Previous studies have shown that echinoderm neuromuscular signalling appears to be mediated in part by biogenic amines, acetylcholine, nitric oxide, and GABA (Florey et al., 1975; Elphick and Melarange, 2001), and putative genes encoding proteins homologous to nitric oxide synthase and monoamine neurotransmitter receptors were identified in the genome (Burke et al., 2006).

1.7 The identification of a novel SALMFamide neuropeptide precursor gene in the sea urchin genome

Prior to recent genome sequencing and annotation projects, invertebrate neuropeptides were most commonly identified using more time-consuming biochemical techniques including immunolocalization, mass spectrometry, and chromatography. In the post-genomic era high-throughput bioinformatic techniques offer an alternative approach to the identification of novel and homologous neuropeptides in sequenced genomes.

The first neuropeptide precursor gene to be identified in the sea urchin genome was a gene encoding members of the SALMFamide family of neuropeptides (Elphick and Thorndyke, 2005). The gene was identified by BLAST searching the genome using a search query consisting of the repeated starfish S1 SALMFamide sequence interrupted by basic residues. For example:

...GFNSALMFGKRGFNSALMFGKRGFNSALMFGKRGFNSALMFGKR...

Thus, the query successfully identified the precursor gene by constructing a sequence emulating the structure of a putative precursor coding region. The gene identified encoded a total of seven SALMFamide family peptides, evidence of intragenic duplication, and the peptides were denominated SpurS1 to SpurS7 (Table 3).

1.8 Aims

Building on the strategy employed by Elphick and Thorndyke (2005), the primary aim of this project was to systematically search the sea urchin genome for neuropeptide precursor genes encoding peptides homologous to the widely diverse neuropeptides previously identified and sequenced from both vertebrates and invertebrates, as well as to identify putative sea urchin neuropeptide receptors. The project also aimed to detect the putative peptides identified in *S. purpuratus* tissue extracts using mass spectrometry, and to investigate the myoactive properties of the putative peptides when tested on sea urchin muscle preparations.

Table 3: Echinoid (sea urchin) neuropeptides. All putative peptides identified from the genome of the sea urchin *Strongylocentrotus purpuratus* (Elphick and Thorndyke, 2005).

Name	Sequence
SpurS1	PPVTTRSKFTF-NH ₂
SpurS2	DAYSAFSF-NH ₂
SpurS3	GMSAFSF-NH ₂
SpurS4	AQPSFAF-NH ₂
SpurS5	GLMPSFAF-NH ₂
SpurS6	PHGGSAFVF-NH ₂
SpurS7	GDLAFAF-NH ₂

Chapter 2

THE IDENTIFICATION OF PUTATIVE SEA URCHIN NEUROPEPTIDE RECEPTORS

- 2.1** Introduction: Neuropeptide GPCR structure and diversity
- 2.2** Methods: Bioinformatic identification of putative receptors
- 2.3** Results
 - 2.3.1** Putative sea urchin neuropeptide GPCRs
 - 2.3.2** Putative sea urchin neuropeptide GPCRs identified in cDNA libraries
- 2.4** Discussion

2.1 Introduction: Neuropeptide GPCR structure and diversity

This chapter reports the identification of putative *S. purpuratus* neuropeptide G-protein coupled receptor (GPCR) genes. A bioinformatic investigation of the number of putative sea urchin neuropeptide receptors can be used as an approximate estimate of the number of expressed sea urchin neuropeptide ligands. The structure of the receptor and its homology to sequenced vertebrate receptors can also assist in predicting the structure and identity of a receptor's ligand. This chapter aims to comprehensively identify putative neuropeptide receptors in the sea urchin genome and to correlate their phylogenetic position and molecular structure to putative neuropeptide ligands.

GPCRs are the most common neuropeptide receptors across the animal kingdom. The super-family of receptors are characterized by activating cytoplasmic G proteins upon binding a diverse range of ligands, including peptides, biogenic amines, chemokines, photons, lipids and organic odorants (Lichtarge et al., 1996; Fredriksson et al., 2003; Kamesh et al., 2008). There are many human orphan GPCRs without functional characterisation and for which endogenous ligands and pharmacological agents are yet to be identified. GPCRs are highly important as receptors for pharmacological agents, contributing to up to 50% of the targets for drug design (Gudermann et al., 1997; Flower, 1999; Wilson and Bergsma, 2000).

Due to evolutionarily conserved similarities in sequence structure, phylogenetic tree-building can provide a framework for structural and functional comparison of invertebrate and vertebrate receptors. GPCRs are characterized by seven transmembrane (7TM) domains of 25 to 35 residues that cross the cell membrane seven times (Gudermann et al., 1997). The polypeptide has an extracellular N-terminal and an intracellular C-terminal. Ligand specificity is conferred by sequence variation in the TM domains and/or extracellular inter-domain loops.

Thus, conserved 7TM domain structure and variable inter-domain loop regions enable the construction of receptor phylogenies suggesting divergence from a common ancestral gene (Fredriksson et al., 2003). Smaller genome and organism specific sub-families of receptors can also be identified, which indicate gene duplication and expansion. Therefore, GPCR abundance has previously been used as an estimate of neuropeptide precursor expansion and diversification (Bargmann, 1998; Hewes and Taghert, 2001).

There are more than 800 human GPCRs, including about 342 of which are non-olfactory, making them one of the largest protein families expressed in metazoans, including humans (Fredriksson et al., 2003), insects (Brody and Cravchik, 2000; Hewes and Taghert, 2001; Vanden Broeck, 2001a), cephalochordates (Nordström et al., 2008), and urochordates (Kamesh et al., 2008). Frederiksson's analysis of the complement of human GPCRs (2003) described a GRAFS system of five GPCR families based on phylogenetic analysis; Glutamate, Rhodopsin, Adhesion, Frizzled and

Secretin. Previous methods of GPCR classification involved use of an interphyletic alphabetical/numerical clan system (A-F) (Kolakowski, 1994). In this system clan C corresponds to the glutamate (G) family of receptors, clan A corresponds to rhodopsin (R) family GPCRs, and clan B corresponds to the secretin (S) family GPCRs. Fredriksson (2003) identified two families of GPCRs independent of the former system, adhesion (A) and frizzled (F). Clans D and E exclusively represented fungal pheromone receptors and cAMP receptors and clan F contained archaebacterial opsins. The GRAFS system has subsequently been used to classify GPCRs from 13 sequenced protostome and deuterostome organisms (Kamesh et al., 2008). Therefore, phylogenetic analysis using the receptor amino acid sequences allows for comparative bioinformatic studies of receptors from other phyla (Vanden Broeck, 2001a; Fredriksson et al., 2003; Nordström et al., 2008).

Most neuropeptide receptors are found in the rhodopsin family of GPCRs. The rhodopsin family is the largest of the GPCR families and has been subdivided into four groups (α , β , γ , δ) with 13 main branches (Fredriksson et al., 2003). Neuropeptides are primarily found in the alpha, beta and gamma groups. There are five main branches of vertebrate α subgroup rhodopsin family receptors, including a prostaglandin clade, an amine receptor clade, an opsin receptor clade, a melatonin receptor clade, and a MECA receptor clade, including receptors to melanocortin (Fredriksson et al., 2003). The ligands for beta (β) group rhodopsin receptors are exclusively peptides and the gamma (γ) group includes receptors for chemokines. The delta group of rhodopsin GPCRs includes receptors that

bind neurally secreted glycoproteins. The family of receptors are defined by characteristics including an NSxxNPxxY motif in transmembrane domain VII and a D(E)-R-Y(F) motif at the border of transmembrane domain III (Fredriksson et al., 2003).

The receptors in the glutamate family of GPCRs, corresponding to clan C, include glutamate receptors, GABA receptors, and taste receptors. There are, however, no known vertebrate neuropeptide receptors in this family.

The adhesion family of GPCRs is defined by the presence of a GPS domain (Nordström et al., 2008). There are 33 human adhesion-family GPCRs and 37 amphioxus adhesion-family GPCRs. The frizzled family of GPCRs contains two groups: frizzled receptors and taste receptors. The taste receptors are thought to be expressed in the tongue and palate epithelium (Fredriksson et al., 2003). The frizzled receptors bind glycoproteins and are involved in the regulation of cell growth (Fredriksson et al., 2003). There are also no known vertebrate neuropeptide receptors within either of these families.

In this bioinformatic investigation of sea urchin neuropeptide receptors, vertebrate neuropeptide receptor sequences, which are members of the families and groups of neuropeptide receptors defined by Fredriksson (2003), were used to identify homologous putative sea urchin GPCR genes.

2.2 Methods: Bioinformatic identification of putative receptors

As part of the sea urchin genome annotation project, sea urchin GPCRs had been identified using Pfam domain analysis and compiled into secretin and rhodopsin families according to Fredriksson's system of classification (Bateman et al., 2000). The complement of *S. purpuratus* GPCRs corresponding to each sub-family of vertebrate neuropeptide receptors was identified using the full amino acid sequence of known human neuropeptide GPCR sequences as BLAST query entries against the set of Glean3 gene predictions assembled from the sequenced sea urchin genome (<http://www.spbase.org/SpBase/wwwblast/blast.php>). Putative receptors were identified on the basis of strong homology to vertebrate neuropeptide GPCRs. For example, a putative sea urchin neuropeptide CCK-like receptor (Sp_115741932) was 33% homologous to the *Ciona* CCK-like receptor (GI:74136025) and 36% homologous to human CCK B receptor (GI:119589141).

Human neuropeptide GPCRs were also used to identify homologous GPCRs in six other organisms using BLAST searches of the NCBI protein database and phylogenetic analysis using ClustalX and NJ Plot (European Bioinformatics Institute) (Higgins and Sharp, 1988; Thompson et al., 1997). The representative organisms from which GPCR gene sequences were collated included *Homo sapiens* (Phylum Chordata, Subphylum Vertebrata), the lancelet *Branchiostoma floridae* (Phylum Chordata, Subphylum Cephalochordata), the sea squirt *Ciona intestinalis* (Phylum Chordata, Subphylum Urochordata), the roundworm *Caenorhabditis*

elegans (Phylum Nematoda), the opisthobranch mollusc *Aplysia californica* (Phylum Mollusca), the fruitfly *Drosophila melanogaster* (Phylum Arthropoda), and the sea anemone *Nematostella vectensis* (Phylum Cnidaria). Phylogenetic trees were rooted using closely-related but phylogenetically separate GPCRs supported using bootstrap analysis and described in the study by Fredriksson et al (2003).

The same strategy of using vertebrate GPCR sequences as BLAST search queries was used to identify GPCRs in an EST database derived from embryonic and larval stages, and adult tube foot and radial nerve tissues to refine the set to genes expressed within those tissues (<http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=7668>).

2.3 Results

The majority of sea urchin GPCRs identified by the presence of a 7TM are putative orphan receptors without strong homology to known vertebrate or invertebrate neuropeptide receptors. Ascribing a putative ligand for orphan sea urchin GPCRs is unachievable via bioinformatic analysis alone.

However the genome contains at least 38 putative neuropeptide GPCRs identified on the basis of strong homology and sequence alignment to vertebrate rhodopsin family beta group GPCRs (Table 4). These results were summarized and included as part of a published bioinformatic analysis of the sea urchin nervous system (Burke et al., 2006). The results presented in this chapter are discussed according to GPCR group as defined

by the GRAFS system of classification, including putative sea urchin neuropeptide receptors of the rhodopsin and secretin family GPCRs.

2.3.1 Putative sea urchin neuropeptide GPCRs

There are 874 putative sea urchin rhodopsin-family GPCRs predicted by Pfam protein domain analysis. A total of fifty-four putative sea urchin neuropeptide GPCRs were identified. Thirty-eight putative sea urchin neuropeptide receptors showed homology to GPCRs of the beta (β) group of Rhodopsin GPCRs defined by Fredriksson (2003) (Figures 1-9). Putative neuropeptide receptors were also identified showing homology to vertebrate neuropeptide receptors within the secretin GPCR family and the gamma (γ) and delta (δ) groups of rhodopsin GPCR family. The melanocortin receptors, belonging to the ' α ' subfamily MECA receptor clade, as well as the cluster of human GPCRs containing GPR3, GPR6, and GPR12, appear not to have homologous invertebrate receptors.

The complement of sea urchin beta group rhodopsin receptors is illustrated in comparison to 32 human beta (β) rhodopsin family receptors (Figures 1-4). Putative sea urchin receptor-encoding genes showing sequence homology accompanied by strong bootstrap support include two tachykinin-like receptors (Figures 1 & 5), two vasotocin-like receptors (Figures 4 & 6), three gonadotropin-releasing hormone receptors (Figures 4 & 6), one cholecystokinin-like receptor (Figures 1 & 7), two bombesin/neuromedin B-like receptors (Figure 3 & 8), and two orexin/hypocretin-like receptors (Figures 1 & 9). Twenty-six other putative

rhodopsin beta group receptors were identified, three of which grouped with a clade of vertebrate receptors including neuromedin U (Figure 3).

Fredricksson (2003) distinguished two separate subgroups within the gamma (γ) group: the SOG subgroup, which includes receptors for galanin-like peptides, the RFamide receptor GPR54, somatostatin receptors and opioid receptors, and the MCH (melanin-concentrating hormone) subgroup. The sea urchin GPCRs identified included homologues to vertebrate receptors in both of these sub-clades, including a cluster of six GPR54-like receptors (Figure 10), a galanin-like receptor (Figure 11), two receptors grouped with the orphan vertebrate receptors GPR7 and GPR8 (Figure 12), and a melanin-concentrating hormone-like receptor (Figure 13).

The delta (δ) group of rhodopsin receptors includes receptors for the vertebrate glycoproteins luteinizing hormone, follicle-stimulating hormone, and thyroid stimulating hormone (Hsu et al., 2002; Fredriksson et al., 2003). One putative sea urchin GPCR was found to show homology to glycoprotein hormone receptors and two putative sea urchin GPCRs were found to show homology to relaxin-like receptors LGR7 and LGR8 (Figure 14).

From the secretin family of GPCRs two putative sea urchin neuropeptide receptor homologues were identified; these showed homology to corticotropin-releasing hormone receptor and calcitonin receptor (Figure 14).

2.3.2 Putative sea urchin neuropeptide GPCRs identified in EST libraries

Seventy-four GPCRs were identified in EST libraries derived from embryonic and adult tissues (Table 5). The majority were found in EST libraries produced from echinoderm larvae at early developmental stages, however the number identified is reflective of the size of the EST dataset. Three ESTs encoding putative neuropeptide GPCRs were identified. The putative cholecystinin-like receptor (SPU_026458 or Sp_115741932) was found in the cDNA library assembled from adult radial nerve tissue (Table 5). Two other putative neuropeptide receptors are also found in the radial nerve tissue cDNA library (SPU_012974 and SPU_026973), which both have reciprocal BLAST matches to the human galanin receptor. Putative non-peptidergic receptors identified in the EST libraries included an α -1 adrenergic-like receptor (SPU_001020), two muscarinic cholinergic-like receptors (SPU_011401 and SPU_021092), and a histamine H₁-like receptor (SPU_022857). A putative receptor homologue of *Drosophila* ecdysis-triggering hormone receptor was also identified in a sea urchin testis cDNA library (SPU_027212).

Table 4: Putative neuropeptide G-protein coupled receptors in *S. purpuratus* and their vertebrate homologues, identified by BLAST search, sequence alignment, and bootstrap analysis. The twenty-eight were identified on the basis of strong sequence homology.

Rhodopsin Family Beta Group-like Receptors

Tachykinin-like	Sp_115665194, SPU_011338
Vasotocin-like	SPU_021291, Sp_115899360
Gonadotropin-releasing hormone-like	Sp_185134947, Sp_185134985, Sp_185134933
Cholecystokinin-like	Sp_115741932
Bombesin/neuromedin B-like	Sp_72172891, Sp_115928214
Orexin-like	SPU_023223, Sp_72146578

Rhodopsin Family Gamma Group-like Receptors

GPR54-like	SPU_021291, SPU_012318, SPU_007422, SPU_025226, SPU_022468, SPU_022376, SPU_015139
Galanin-like	Sp_115841642
Neuropeptide W	Sp_115923312, SPU_026626
Melanin-concentrating hormone-like	Sp_72011294

Rhodopsin Family Delta Group-like Receptors

Gonadotropin-like	Sp_115678957
Relaxin/insulin-like	SPU_015740, SPU_003527

Secretin Family Neuropeptide Receptors

Corticotropin-releasing hormone-like	Sp_115720160
Calcitonin/CGRP-like	SPU_018314

The genes labelled by four digits preceded by zero (SPU_0XXXX) are identical to the four number Glean3 gene predictions (XXXX) assigned as part of the sea urchin genome annotation project. Where available the identifier has been with the gene index number (Sp_X).

Table 5: Putative *S. purpuratus* neurotransmitter and neuropeptide receptors identified in EST libraries assembled from larval and adult sea urchin tissues.

EST Library	GPCR gene prediction (SPU_0XXXXX)
Blastula-stage	000498, 000703, 002312, 003105, 003365, 004925, 005097, 005162, 005239, 008205, 011108, 012382, 012383, 013918, 015161, 015968, 018339, 021363, 021507, 021683, 022683, 023047, 024079, 024655, 024811, 025314, 025436, 025453, 025821, 027584, 027884, 028801, 028801, 006250
Mesenchyme	000703, 003101, 004805, 004925, 007574, 008205, 012382, 012610, 014765, 014768, 015263, 019158, 020348, 020361, 022219, 022220, 023251, 006919, 006920
7 hour embryo	000703, 004925, 023253
20 hour embryo	002875, 002876, 009880, 011953, 013248, 027884, 028801
40 hour embryo	002876, 011953, 013248
Larval	002875, 002876, 003698, 010167, 011320, 011953, 012979, 013248, 013760, 014448, 015217, 015263, 016400, 017624, 021176, 021683, 024079, 028056
Coelomocyte	001020 (alpha-1 adrenergic receptor), 013595, 015593, 028824
Radial nerve	002312, 012974, 021092 (cholinergic muscarinic 3-like receptor), 022857 (histamine H1-like receptor), 026458 (CCK-like receptor), 026973
Lantern	011401 (cholinergic muscarinic 1-like receptor)
Testis	027212

The genes labelled by four digits preceded by zero (SPU_0XXXXX) are identical to the four number Glean3 gene predictions (XXXX) assigned as part of the sea urchin genome annotation project.

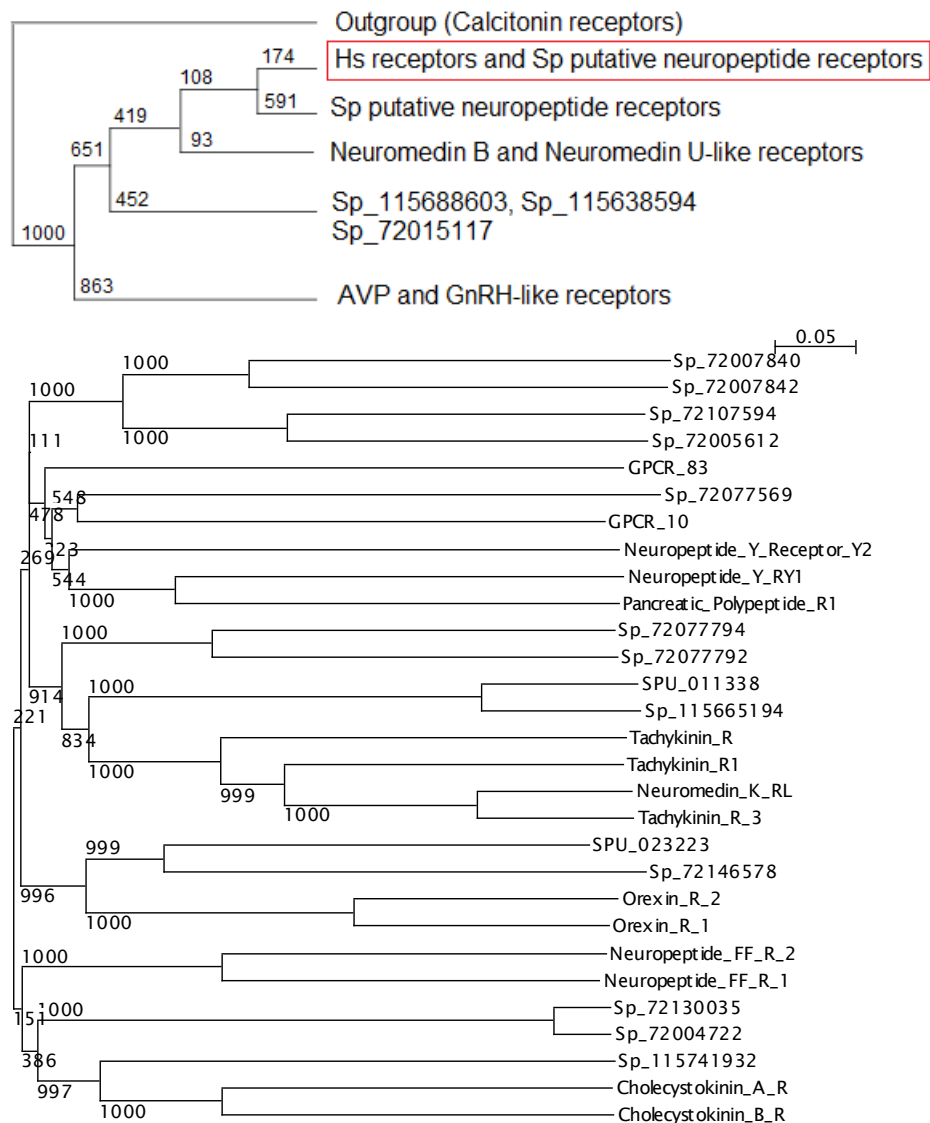


Figure 1. Human rhodopsin family beta group receptors and putative sea urchin homologues. The key above the phylogeny illustrates the position of the subclade within a clade of rhodopsin family-like GPCRs assembled from sea urchin and human GPCRs using ClustalX and NJPlot. Calcitonin receptors, which belong to the secretin family of human GPCRs, were used as an outgroup to the rhodopsin family-like GPCR clade. Bootstrap values are shown adjacent to tree branches. Putative sea urchin genes Sp_X are labelled by gene index number and putative sea urchin genes SPU_X are labelled according to the Glean3 gene prediction numbers assembled as part of the sea urchin genome annotation project. Hs = *Homo sapiens*. Sp = *Strongylocentrotus purpuratus*.

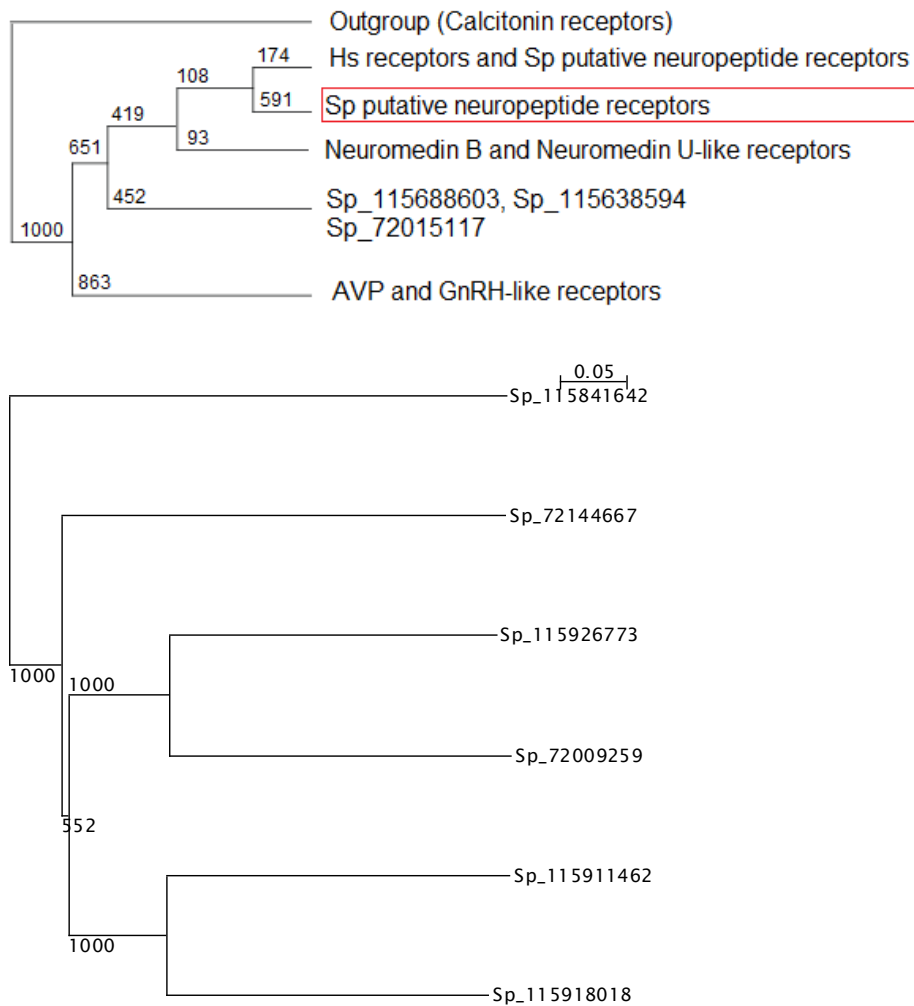


Figure 2. Putative sea urchin peptide receptors homologous to human rhodopsin family beta group receptors. The key above the phylogeny illustrates the position of the subclade within a clade of rhodopsin family-like GPCRs assembled from sea urchin and human GPCRs using ClustalX and NJPlot. Calcitonin receptors, which belong to the secretin family of human GPCRs, were used as an outgroup to the rhodopsin family-like GPCR clade. Bootstrap values are shown adjacent to tree branches. Putative sea urchin genes Sp_X are labelled by gene index number. Hs = *Homo sapiens*. Sp = *Strongylocentrotus purpuratus*.

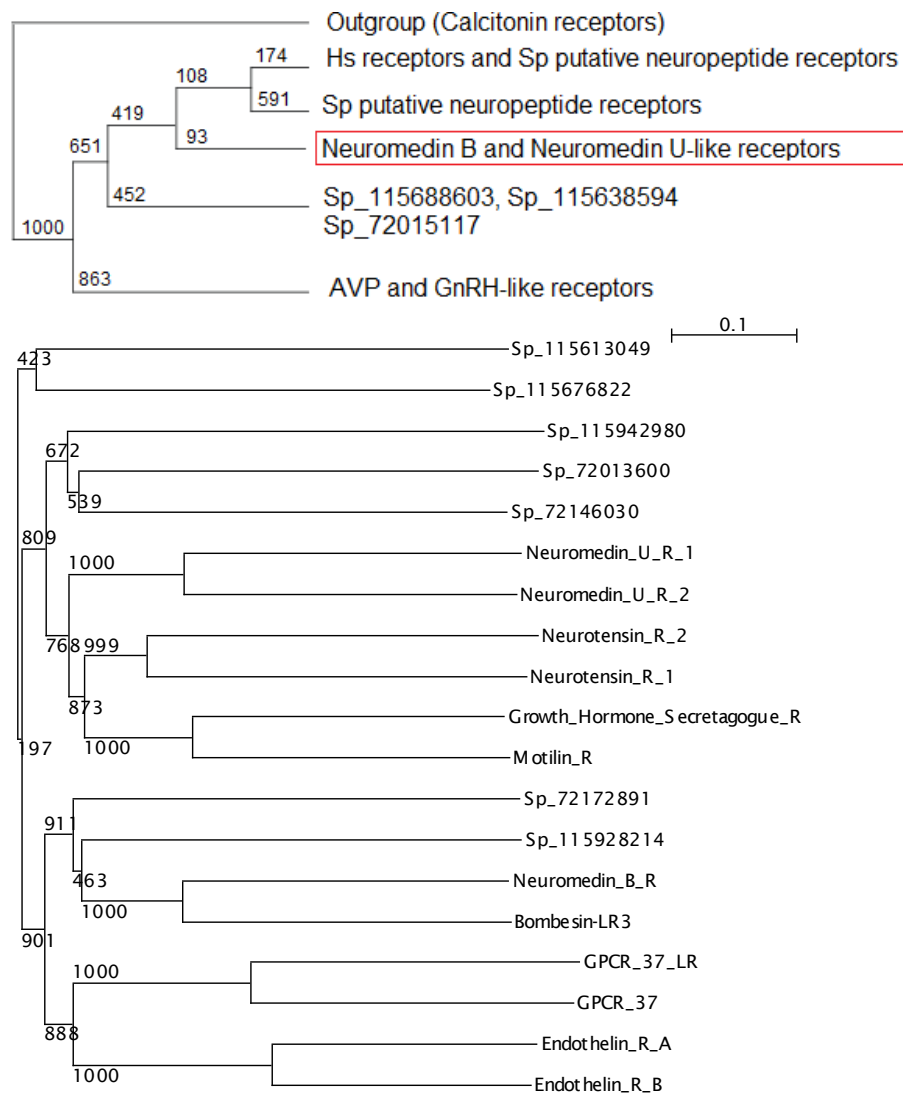


Figure 3. Human and putative sea urchin neuromedin B/bombesin, neuromedin U, neurotensin, and growth hormone secretagogue-like receptors. The key above the phylogeny illustrates the position of the subclade within a clade of rhodopsin family-like GPCRs assembled from sea urchin and human GPCRs using ClustalX and NJPlot. Calcitonin receptors, which belong to the secretin family of human GPCRs, were used as an outgroup to the rhodopsin family-like GPCR clade. Bootstrap values are shown adjacent to tree branches. Putative sea urchin genes Sp_X are labelled by gene index number. Hs = *Homo sapiens*. Sp = *Strongylocentrotus purpuratus*.

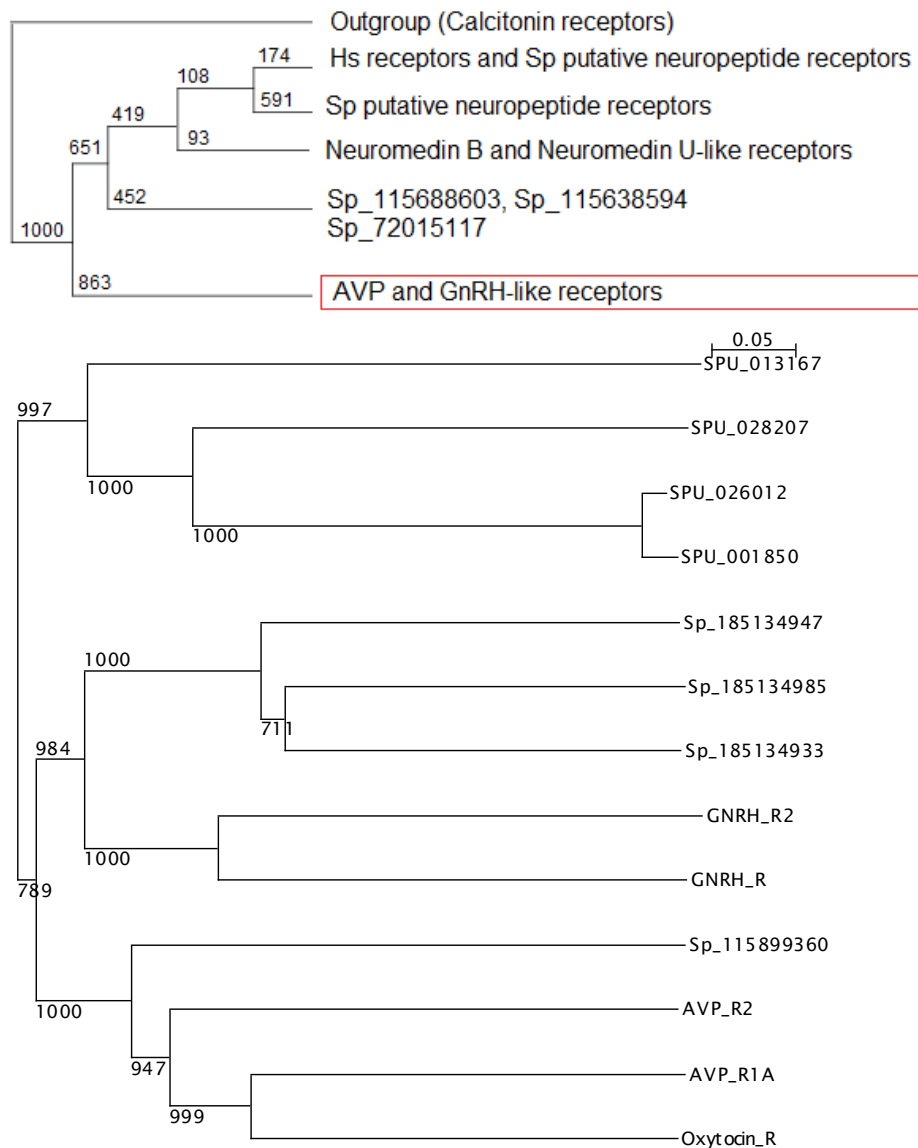


Figure 4. Human and putative sea urchin AVP and GnRH-like receptors. The key above the phylogeny illustrates the position of the subclade within a clade of rhodopsin family-like GPCRs assembled from sea urchin and human GPCRs using ClustalX and NJPlot. Calcitonin receptors, which belong to the secretin family of human GPCRs, were used as an outgroup to the rhodopsin family-like GPCR clade. Bootstrap values are shown adjacent to tree branches. Putative sea urchin genes Sp_X are labelled by gene index number and putative sea urchin genes SPU_X are labelled according to the Glean3 gene prediction numbers assembled as part of the sea urchin genome annotation project. Hs = *Homo sapiens*. Sp = *Strongylocentrotus purpuratus*.

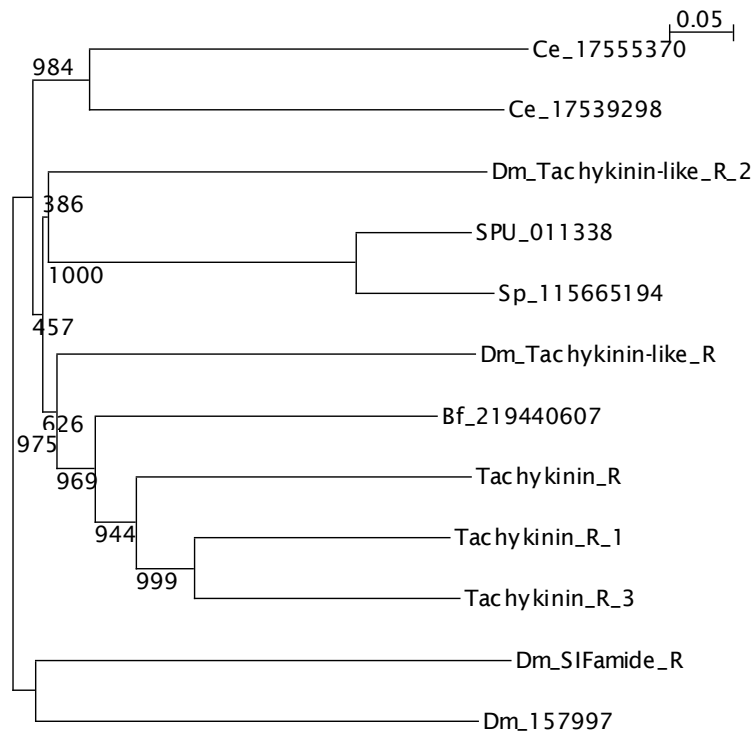


Figure 5. Human and putative non-vertebrate tachykinin-like receptors. Dm_SIFamide_R (*Drosophila* SIFamide receptor) and Dm_157997 (*Drosophila* receptor GI:157997) were used as outgroups. Putative sea urchin genes are labelled by gene index number (Sp_X). Sp = *Strongylocentrotus purpuratus*. Dm = *Drosophila melanogaster*. Bf = *Branchiostoma floridae*. Ce = *Caenorhabditis elegans*.

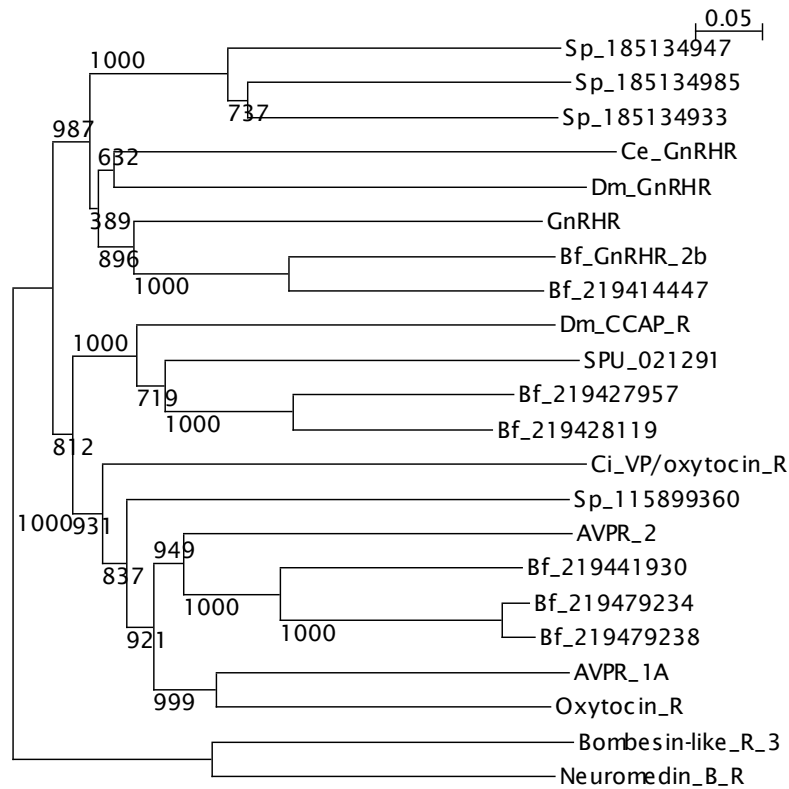


Figure 6. Human and putative non-vertebrate gonadotropin-releasing hormone and vasotocin-like receptors. Human neuromedin B and bombesin-like receptors were used as outgroups. Putative sea urchin genes are either labelled by gene index number (Sp_X) or sea urchin genome annotation index number (SPU_X). Sp = *Strongylocentrotus purpuratus*. Dm = *Drosophila melanogaster*. Bf = *Branchiostoma floridae*. Ce = *Caenorhabditis elegans*.

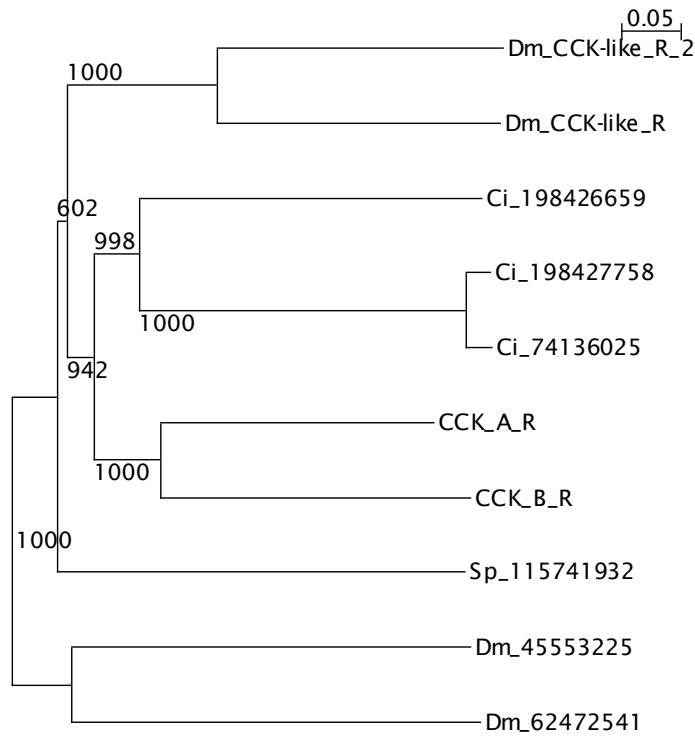


Figure 7. Human and putative non-vertebrate cholecystokinin-like receptors. Two putative *Drosophila* neuropeptide receptors (gene index numbers 4553225 and 62472541) were used as outgroups. Putative sea urchin genes are labelled by gene index number (Sp_X). Sp = *Strongylocentrotus purpuratus*. Dm = *Drosophila melanogaster*. Bf = *Branchiostoma floridae*. Ce = *Caenorhabditis elegans*. Ci = *Ciona intestinalis*.

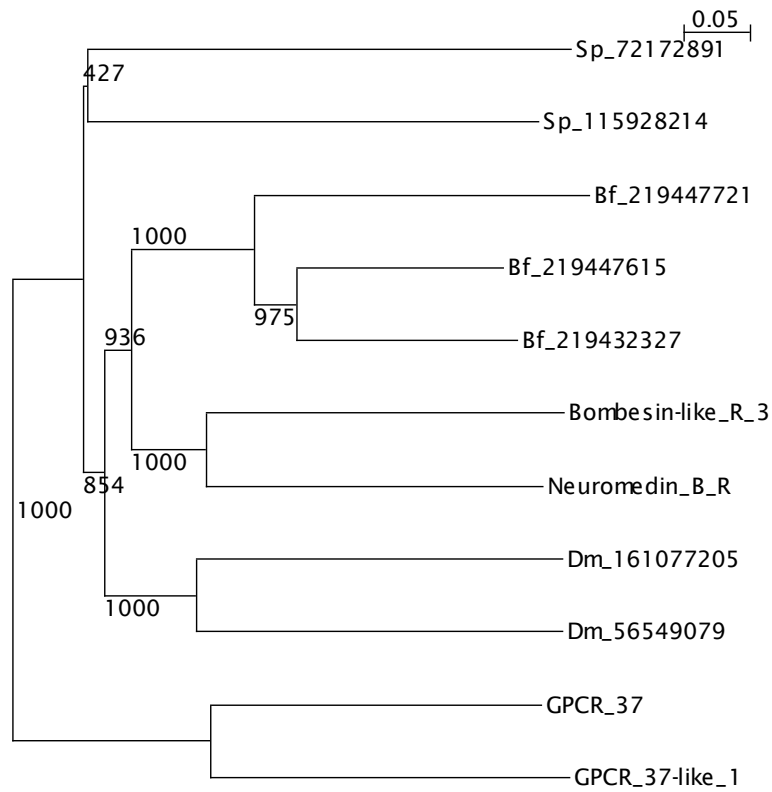


Figure 8. Human and putative non-vertebrate bombesin and neuromedin B-like receptors. Human GPR37 and GPR37-like receptors were used as outgroups. Putative sea urchin genes are labelled by gene index number (Sp_X). Sp = *Strongylocentrotus purpuratus*. Dm = *Drosophila melanogaster*. Bf = *Branchiostoma floridae*.

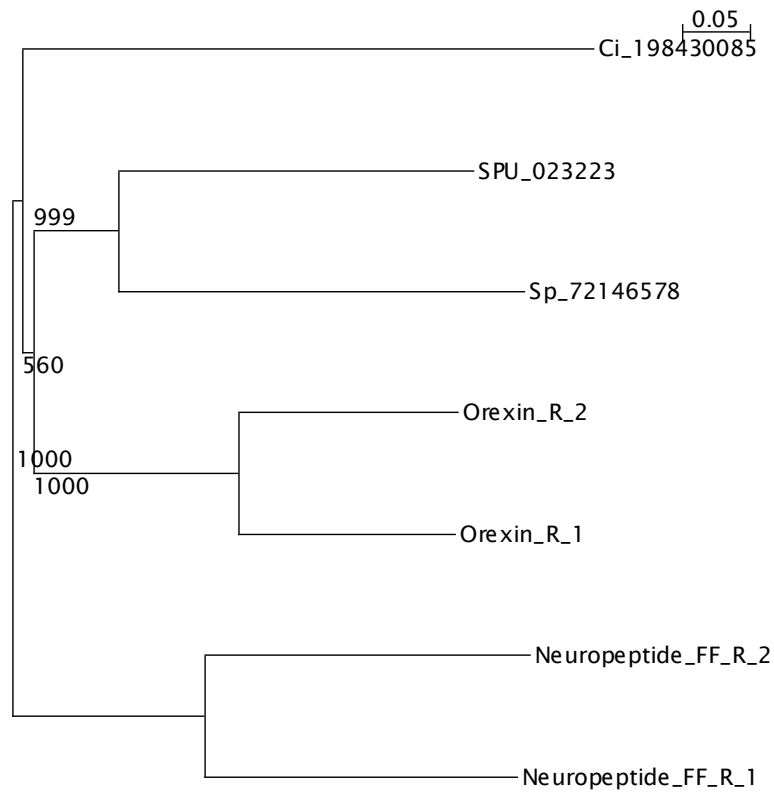


Figure 9. Human and putative non-vertebrate orexin/hypocretin-like receptors. Putative sea urchin genes are either labelled by gene index number (Sp_X) or sea urchin genome annotation index number (SPU_X). Human neuropeptide FF receptors were used as a outgroups. Sp = *Strongylocentrotus purpuratus*. Ci = *Ciona intestinalis*.

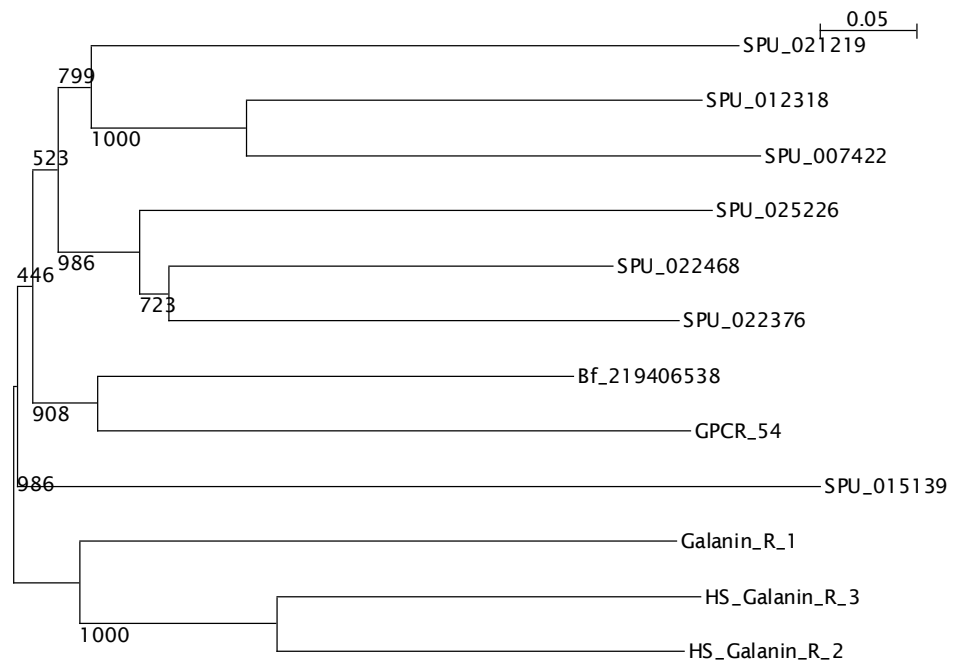


Figure 10. Human and putative non-vertebrate GPR54-like receptors. Human galanin receptors were used as outgroups. Putative sea urchin genes are labelled by sea urchin genome annotation index number (SPU_X). Sp = *Strongylocentrotus purpuratus*. Dm = *Drosophila melanogaster*. Bf = *Branchiostoma floridae*. Ci = *Ciona intestinalis*.

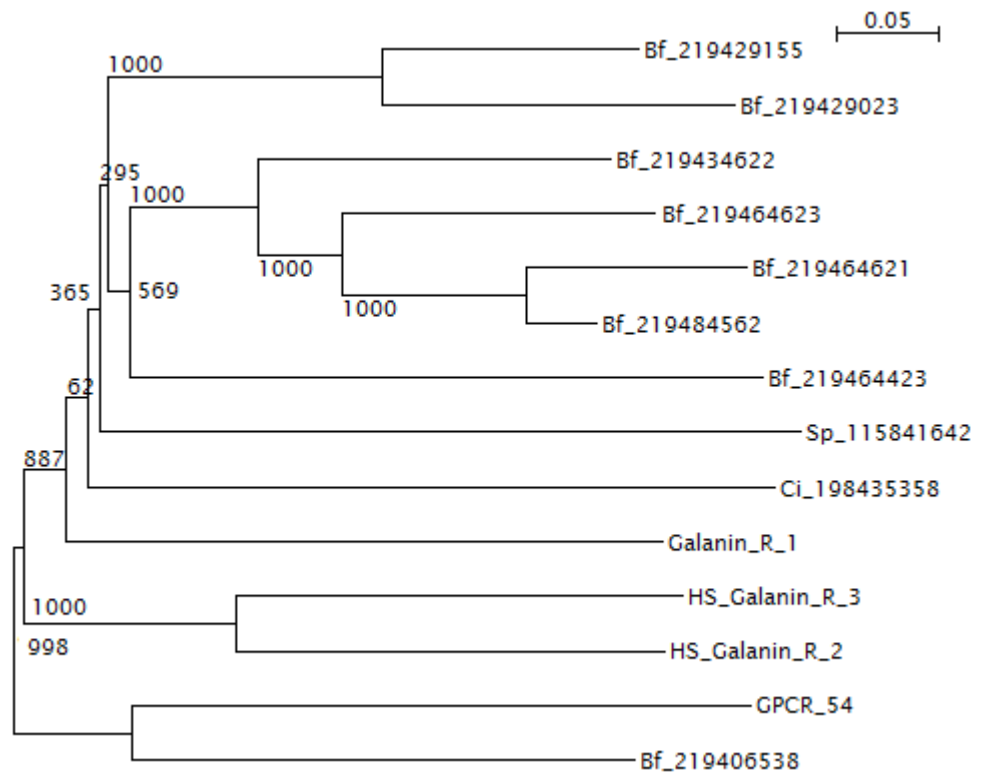


Figure 11. Human and putative non-vertebrate galanin-like receptors. Kisspeptin receptor GPR 54 and a putative GPR 54-like *Branchiostoma* receptor were used as outgroups. Putative sea urchin genes are labelled by gene index number (Sp_X). Sp = *Strongylocentrotus purpuratus*. Dm = *Drosophila melanogaster*. Bf = *Branchiostoma floridae*. Ci = *Ciona intestinalis*.

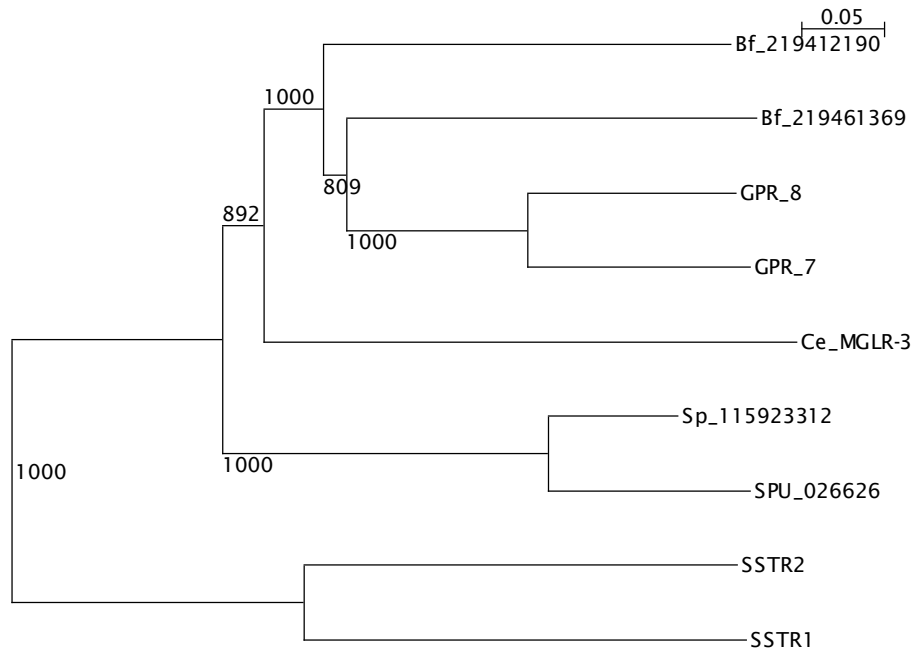


Figure 12. Human and putative non-vertebrate GPR7 and GPR8-like receptors. Human somatostatin receptors were used as outgroups. Putative sea urchin genes are either labelled by gene index number (Sp_X) or sea urchin genome annotation index number (SPU_X). Sp = *Strongylocentrotus purpuratus*. Dm = *Drosophila melanogaster*. Bf = *Branchiostoma floridae*.

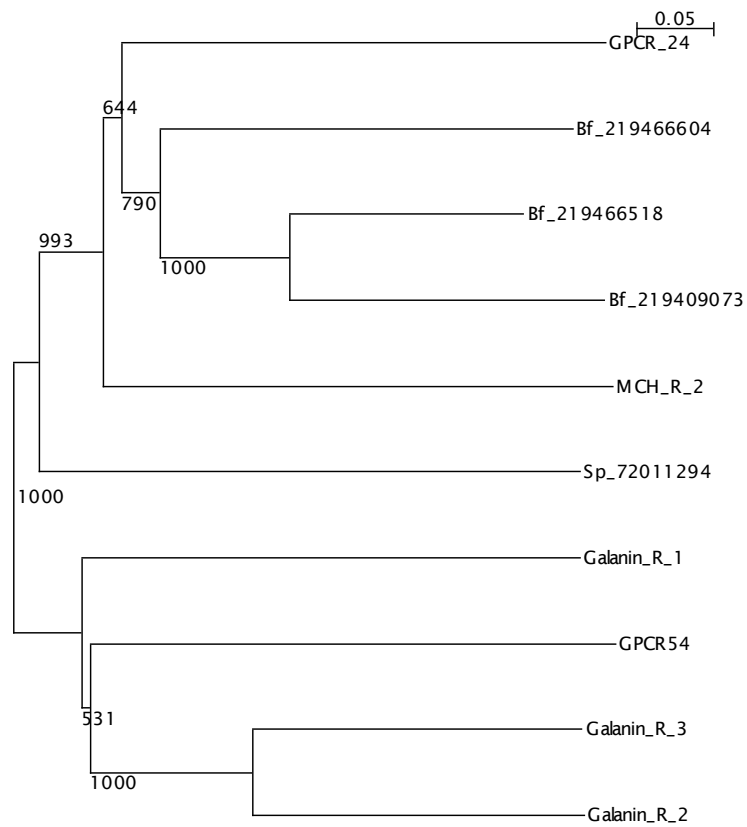


Figure 13. Human and putative non-vertebrate melanin-concentrating hormone (MCH)-like receptors. Human galanin receptors and GPCR 54 were used as outgroups. Putative sea urchin genes are either labelled by gene index number (Sp_X). Sp = *Strongylocentrotus purpuratus*. Dm = *Drosophila melanogaster*. Bf = *Branchiostoma floridae*.

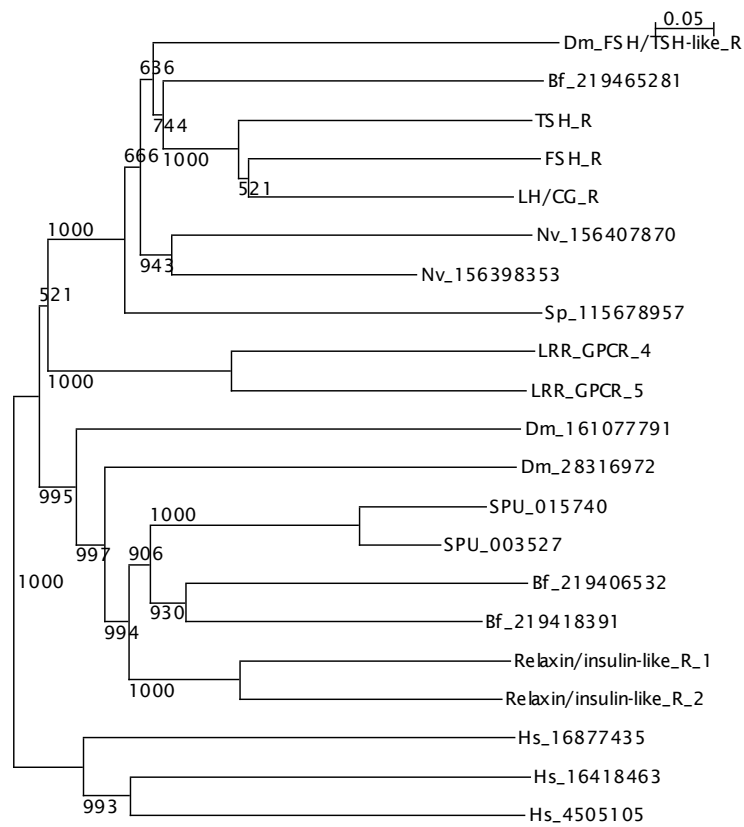


Figure 14. Human and putative non-vertebrate gonadotropin and relaxin-like receptors. Human MAS1 oncogene-like receptors were used as outgroups. Putative sea urchin genes are either labelled by gene index number (Sp_X) or sea urchin genome annotation index number (SPU_X). Sp = *Strongylocentrotus purpuratus*. Dm = *Drosophila melanogaster*. Bf = *Branchiostoma floridae*. Nv = *Nematostella vectensis*.

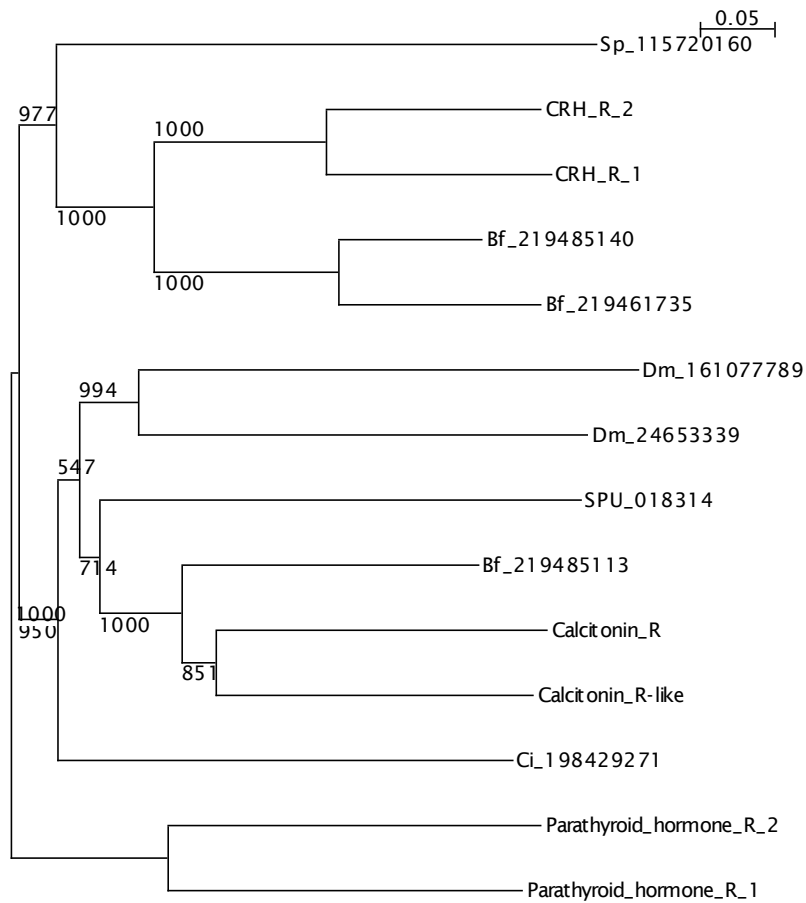


Figure 15. Human and putative non-vertebrate corticotropin-releasing hormone (CRH)-like and calcitonin/CGRP-like receptors. Human parathyroid hormone receptors were used as outgroups. Putative sea urchin genes are either labelled by gene index number (Sp_X) or sea urchin genome annotation index number (SPU_X). Sp = *Strongylocentrotus purpuratus*. Dm = *Drosophila melanogaster*. Bf = *Branchiostoma floridae*. Ci = *Ciona intestinalis*.

2.4 Discussion

The number of putative neuropeptide GPCRs identified (54, including 38 putative rhodopsin family beta subfamily neuropeptide GPCRs) is comparable to the thirty-two putative GPCRs found in a similar analysis of the *Drosophila melanogaster* genome (Cazzamali and Grimmelikhuijzen, 2002). It can be suggested that a similar number of neuropeptide ligands may be expressed in the sea urchin. Structurally homologous peptides encoded by the same precursor may bind a single target with variable affinities and efficacies, and, conversely, homologous peptides may bind more than receptor. Thus, a strict 1:1 peptide precursor to receptor ratio cannot be assumed.

Putative urchin neuropeptide receptors were homologous to receptors in both of the neuropeptide receptor-containing families described by Frederiksson (2003), the rhodopsin and secretin families. There were no putative receptors homologous to neuropeptide receptors in the alpha group of rhodopsin receptors, however there were 39 putative sea urchin receptors homologous to human receptors in the beta group of rhodopsin family receptors. Twelve of these were highly homologous to receptors for human neuropeptides, including tachykinin, arginine vasopressin/oxytocin, gonadotropin-releasing hormone, cholecystokinin, and orexin/hypocretin (Table 4).

Human tachykinins include substance P and neurokinin A, encoded by preprotachykinin A, and neurokinin B. Each of the tachykinins binds

selectively to one of the human tachykinin receptors. The putative sea urchin tachykinin-like receptors are basal to the vertebrate receptors, alongside the two *Drosophila* tachykinin-like receptors identified by Poels et al (2009). *Drosophila* tachykinin-like neuropeptides, referred to as drosotachykinins, have also been identified from a precursor bearing similarity to human preprotachykinin A (Siviter et al., 2000). Therefore, there may be one or more similar neuropeptide precursors in the sea urchin genome.

The single vasotocin-like receptor identified (Sp_115899360) may bind a vasotocin-like peptide and the identification of a single receptor is consistent with the hypothesis that oxytocin and vasopressin originated by gene duplication of a common ancestral precursor encoding a vasotocin-like neuropeptide during early chordate evolution (Hoyle, 1999).

Vasotocin-like peptides have been identified in both protostome and deuterostome phyla and the phylogeny illustrated includes the receptor for the *Drosophila* crustacean cardioactive peptide neuropeptide (Dm_CCAP_R; Figure 6), which also bears sequence similarity to vasotocin-like peptides in having a structure incorporating an amidated C-terminal and a disulfide bridge between two cysteines (PFCNAFTGC-NH₂). Three putative sea urchin GnRH-like receptors have been identified compared to the single human GnRH receptor expressed in the human pituitary gland, which when activated by GnRH induces the release of gonadotropins.

The single putative CCK-like receptor was also identified in the limited cDNA library of adult sea urchin radial nerve tissue, suggesting that the receptor is highly expressed in the sea urchin nervous system and that a hypothetical sea urchin CCK-like peptide may act as either a neurotransmitter or neuromodulator. However, the receptor is basal to protostome CCK-like receptors, which bind protostome CCK-like neuropeptides, suggesting significant structural divergence or a structurally unrelated ligand.

The two orexin/hypocretin receptors bind the two orexin/hypocretin neuropeptides expressed in the vertebrate brain and which are encoded by a single neuropeptide precursor (De Lecea et al., 1998; Sakurai et al., 1998). Furthermore, the identification of an orexin/hypocretin-like receptor in the sea squirt *Ciona intestinalis* genome suggests orexin/hypocretin-like peptides may be expressed in a non-vertebrate deuterostome.

Other putative sea urchin receptors were homologous to rhodopsin family gamma group receptors for human neuropeptides, including kisspeptin, galanin, neuropeptide W, and melanin-concentrating hormone (Table 4). Although GPCRs 7 and 8 are about 33% similar in sequence to human opioid receptors (Fredriksson et al., 2003), homologs of which may be absent from non-vertebrate genomes, both receptors bind neuropeptide W and neuropeptide B and are expressed in the vertebrate central nervous system (Singh and Davenport, 2006). Interestingly, a GPR7/8-like receptor was also identified in the genome of the nematode *C. elegans*, although a peptide ligand has yet to be identified.

Vertebrate RFamide receptors are distributed across several GPCR families. Seven sea urchin neuropeptide receptors were found to show homology to vertebrate rhodopsin family gamma group receptor GPR 54, which binds the vertebrate RFamide kisspeptin/metastin, originally named for its action as a metastasis suppressor. The other vertebrate RFamide receptors, for which homologous putative sea urchin receptors were not identified, are found in the beta group (neuropeptide FF and prolactin-releasing peptide receptors) and the delta group (26 RFamide receptor).

One putative sea urchin neuropeptide receptor (Sp_115678957) was homologous to the receptors for human gonadotropin glycoprotein neuropeptides follicle-stimulating hormone (FSH), thyroid-stimulating hormone (TSH), and luteinizing hormone (LH). Insect gonadotropins are more diverse and unrelated (Kuczer et al., 2007), however a putative *Drosophila* peptide receptor with sequence homology to human gonadotropins was identified in this study and two putative insect gonadotropin receptors have previously been identified by De Loof et al (2001).

Lastly, two putative sea urchin receptors were found to be homologous to corticotropin-releasing hormone and calcitonin receptors. Human corticotropin-releasing hormone is a 41 amino acid peptide releasing factor and neurotransmitter which induces the release of corticotropin from the anterior pituitary. Although two CRH-like receptors are expressed in both the human and cephalochordate genomes, a single putative CRH-like

receptor was identified using phylogenetic analysis and bootstrap values support a hypothesis of independent duplication in each deuterostome lineage.

The phylogenetic trees illustrated conform to accepted models of animal evolution. The results of phylogenetic analysis place many of the vertebrate GPCRs and putative sea urchin GPCRs into a model which conforms to current understanding of deuterostomian evolution (Figures 1-15). The results of sequence alignment and bootstrap analysis support the hypothesis of a super-family of GPCRs arising from a eukaryote ancestor. Putative cnidarian GPCRs are located at basal branches of clades representing structurally distinct sub-families of human GPCRs (e.g. Figure 13).

Evidence of monophyletic vertebrate receptors in paraphyletic relationship to putative sea urchin neuropeptide receptors may correlate with the emergence of vertebrate paralogs accompanying whole genome duplication events in early chordate evolution (Lundin, 1993; Holland et al., 1994; Panopoulou et al., 2003). Particular receptors may only be expressed within the chordate phylum (e.g. endogenous opiate receptors), the deuterostomes (e.g. hypocretin and melanin-concentrating hormone receptors), or may be expressed across a broad range of bilaterian phyla (e.g. receptors for calcitonin, tachykinin, gonadotropin-releasing hormone, cholecystokinin, vasotocin-like peptides, and FaRPs).

There were 161 sea urchin secretin-family GPCRs predicted by Pfam genomic analysis, representing a large expansion in comparison to the 15

secretin-family receptors in the human genome (Fredriksson et al., 2003) and the 16 secretin-family receptors identified in the cephalochordate genome (Nordström et al., 2008). Twenty-eight of these are predicted to be metabotropic glutamate receptors.

In conclusion, the identification of putative sea urchin neuropeptide receptors reported here will assist in future studies of neuropeptide function in echinoderm larval development and adult physiology and behaviour.

Chapter 3

NGFFFamide AND ECHINOTOCIN: NOVEL MYOACTIVE NEUROPEPTIDES DERIVED FROM NEUROPHYSIN-CONTAINING PRECURSORS

- 3.1** Introduction: Identification of neuropeptide precursor genes in *Strongylocentrotus*
- 3.2** Methods
 - 3.2.1** Bioinformatics: Identification of putative sea urchin neuropeptide genes through BLAST homology searching
 - 3.2.2** Mass Spectrometry: Development of protocols for the identification of starfish SALMFamides
 - 3.2.3** Methods: Immunohistochemistry protocols for the localization of echinotocin
 - 3.2.3.1** Immunoblot Assay Protocol
 - 3.2.3.2** Immunohistochemistry: Tissue fixation, sectioning and immunostaining
 - 3.2.4** Methods: Pharmacological investigation of echinotocin and NGFFFamide myoactivity
- 3.3** Results
 - 3.3.1** Identification of NGFFFamide
 - 3.3.2** Identification of the neurophysin domain
 - 3.3.3** Identification of echinotocin
 - 3.3.4** Mass spectrometry: detection of starfish SALMFamides and sea urchin NGFFFamide
 - 3.3.5** Results: Immunohistochemistry
 - 3.3.6** Results: Characterization of NGFFFamide and echinotocin myoactivity
 - 3.3.6.1** Acetylcholine and carbachol
 - 3.3.6.2** Echintocin and NGFFFamide
- 3.4** Discussion
 - 3.4.1** Identification of sea urchin glycoprotein hormones
 - 3.4.2** Neurophysin-encoding precursors in the sea urchin genome
 - 3.4.3** Mass spectrometric analysis: Detection of a putative NGFFFamide peptide
 - 3.4.4** Discussion: Immunohistochemical investigation of echinotocin expression
 - 3.4.5** Discussion: Contractile myoactivity of echinotocin and NGFFFamide

3.1 Introduction: Identification of putative neuropeptide precursor genes in *Strongylocentrotus*

In the studies reported in this chapter bioinformatic methods were used to identify genes encoding the precursors of two putative sea urchin neuropeptides, NGFFFamide and echinotocin. The precursors were identified by BLAST homology searching using deuterostome neuropeptide and neuropeptide precursor sequences as queries. Biochemical methods, including mass spectrometry and immunohistochemistry, were also used to detect the two peptides and to localize echinotocin expression.

Pharmacological methods, specifically *in vitro* bioassays of echinoderm tissue, were used to test the myoactivities of the identified putative neuropeptides on sea urchin tissue preparations as an indication of peptide bioactivity.

In the post-genomic era there has been an increase in the number of studies using bioinformatic methods to identify invertebrate neuropeptides (Liu et al., 2006; Husson et al., 2007). However there is no current high-throughput method for the identification of neuropeptide genes in newly sequenced genomes. One reason for the difficulty involved in identifying echinoderm neuropeptides through conventional BLAST homology searching is neuropeptide sequence divergence, which accompanies functional diversification, and the consequent lack of sequence homology with neuropeptides identified in other phyla.

Despite having acquired significant structural variation, neuropeptides derived from the same ancestral prohormone gene may have retained many of the ancestral protein's functional characteristics, including its expression patterns, receptor specificity, signal transduction pathway and even physiological role. However, non-coding residues in the preproprotein sequence are unlikely to be conserved and neuropeptide coding regions are characteristically short, sometimes as few as three residues (for example the vertebrate neuropeptide thyrotropin-releasing hormone with the sequence pyroGlu-His-Pro-NH₂). Because of the short length of most neuropeptides, BLAST search queries of translated genomic data using known neuropeptide sequences can generate an unmanageable number of search results (Liu et al., 2006).

Bioinformatic techniques for the identification of novel and homologous neuropeptides have previously included the use of consensus patterns from other neuropeptides for BLAST algorithm searching (Riehle et al., 2002; Hummon et al., 2006; Liu et al., 2006). Pattern searching for neuropeptides within a sequenced genome relies on the repetition of peptide coding sequences, a structural characteristic of neuropeptide precursor sequences. A sea urchin gene encoding seven SALMFamide neuropeptide family members has already been identified in the genome using this method of BLAST homology searching (Elphick and Thorndyke, 2005) (see Chapter One).

The identification of putative novel neuropeptide genes in the sea urchin genome does not in itself provide evidence of cellular expression or

bioactivity. For this reason sea urchin body wall and visceral tissue extracts were analysed using High Performance Liquid Chromatography Mass Spectrometry (HPLC-MS) for the purpose of detecting the expression of the identified neuropeptides.

Mass spectrometry has been an essential biochemical technique for the identification of peptides and proteins. Before the current 'post-genomic' era the technique was often used to determine the sequences of novel peptides from bioactive or immunoreactive fractions. Mass spectrometry continues to be used for the identification of novel neuropeptides from either solution (e.g. LC-MS) or matrix (e.g. Matrix-assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry or MALDI-TOF MS) incorporating sequencing techniques using tandem MS and peptide fragmentation (Nachman et al., 2002; Husson et al., 2005; Husson et al., 2007).

Pharmacological bioassays for the investigation of bioactivity and immunohistochemical investigations for the purpose of localizing echinotocin were made possible for two reasons. Firstly, specimens of *Echinus esculentus* (Class Echinoidea, Family Echinidae) were readily available for pharmacological *in vitro* bioassays and for tissue sectioning from a source in Scotland.

Secondly, immunohistochemical investigation of echinotocin expression was feasible due to the commercial availability of antibodies prepared against vasopressin. Vasopressin (CYFQNCPRG-NH₂) shares sequence

similarity to the putative *S. purpuratus* echinotocin neuropeptide (CFISNCPKG-NH₂), and is likely to be structurally homologous to unidentified echinotocin-like peptides expressed in other echinoderm species. The two peptides are structurally differentiated by a Phe³ residue in vasopressin compared to an Ile³ in echinotocin. Although a hydrophobic residue in the third position (Ile³) is typically associated with oxytocin-like peptides, echinotocin structurally resembles vasopressin-like peptides in sharing a basic residue in the eighth position (Lys⁸). Furthermore, a polar Ser⁴ residue in echinotocin takes the place of a polar Gln⁴ in vasopressin. Overall, echinotocin and vasopressin share about 56% sequence identity. Therefore, anti-vasopressin antibodies may cross-react with echinotocin and could potentially be used for the localization of echinotocin-like peptides in *Echinus esculentus* tissue sections.

3.2 Methods

3.2.1 Bioinformatics: Identification of putative sea urchin neuropeptide genes through BLAST homology searching

The search for genes encoding homologues of known neuropeptides in the sea urchin genome was initiated using the Basic Local Alignment Search Tool (BLAST) facility available on the Baylor College of Medicine Human Genome Sequencing Center website (<http://www.hgsc.bcm.tmc.edu/projects/seaurchin>). Neuropeptide consensus sequences and vertebrate and invertebrate neuropeptide sequences were used to search for homologues encoded in the sea urchin

genome (see Appendix C for vertebrate neuropeptide sequences). The strategy used was similar to that used to identify the novel SALMFamide-encoding neuropeptide precursor, whereby the BLAST query consisted of an artificial sequence of tandem peptide repeats (Elphick and Thorndyke, 2005). When the neuropeptide sequence was long enough to generate manageable number of BLAST search results the full sequence was entered as a direct search query. Thus, the query sequence, submitted as a tBLASTn search (a protein query against the six nucleotide reading frames of the genomic data), comprised multiple copies of the neuropeptide sequence separated by the residues GKR. A glycine (G) terminal residue often provides a substrate for neuropeptide C-terminal amidation and the lysine-arginine (KR) dipeptide sequence is a dibasic cleavage site for endopeptidases. The search query used to identify NGFFFamide used the three repeats of the sea cucumber NGIWYamide sequence, originally isolated by Iwakoshi et al (1995). For example,

...NGIWYGKRNGIWYGKRNGIWYG...

The presence and location of an N-terminal signal peptide, essential for directing the transport of a neuropeptide prohormone through the endoplasmic reticulum, was predicted using the online program Signal P (www.cbs.dtu.dk/services/SignalP) (Emanuelsson et al., 2007). Two sets of genes predicted by the gene prediction programs Gnomon, developed by Souvorov (unpublished), and the Glean3 algorithm, as well as GenScan (<http://genes.mit.edu/GENSCAN.html>), developed by MIT (Burge and Karlin, 1997), which provides a model of gene intron and exon structures

derived from the input of a nucleotide search query, were used to support the putative gene structures, including the organization and structure of exons and introns. Additionally, the *S. purpuratus* EST libraries were searched, using the putative precursor amino acid sequences as queries, for evidence of the expression of the putative precursor gene, as well as for evidence of intron and exon expression. Search results were also used as query sequences for BLAST searching of the sea urchin genome to identify similar genes, including other putative neuropeptide precursors. The tBLASTn search algorithm was used for precursor sequence queries and the BLASTx was used sea urchin EST queries. Lastly, online homology modelling software was used for structural investigation, including DeepView Swiss PDB Viewer (<http://spdbv.vital-it.ch/index.html>) as software for protein modelling and SWISS-MODEL (<http://swissmodel.expasy.org>) for the generation of homology models based on 3-D protein structural data.

3.2.2 Mass Spectrometry: Development of protocols for the identification of starfish SALMFamides

As random mutations may contribute to variation in peptide residue identities across species and class, the detection of peptides with masses accurately corresponding to the predicted molecular weights of the identified peptides would require samples from *S. purpuratus* specimens, which were not readily available. Therefore, the two starfish neuropeptides, SALMFamides S1 and S2, respectively an octapeptide and a dodecapeptide, originally isolated from the starfish *Asterias rubens* using

HPLC (Elphick et al., 1991a) were used to develop a protocol for mass spectrometric analysis of echinoderm tissue samples. The starfish SALMFamides were selected for protocol development for two reasons. Specimens could be purchased from a commercial fishery at Southend-on-Sea, operating in the Thames Estuary and North Sea, and secondly because both neuropeptides were originally identified and sequenced by Elphick et al. (1991a) using an extraction protocol that could be replicated.

However, replicating the acetone extraction procedure as used by Elphick et al. (1991a) failed to identify SALMFamide S1 (Gly-Phe-Asn-Ser-Ala-Leu-Met-Phe-NH₂). An alternate extraction protocol derived from a mass spectrometric study of nematode neuropeptides by Husson et al. (2005), and which used methanol as a solvent for peptide extraction, was trialled and proved successful for the detection of both SALMFamide neuropeptides, S1 (Gly-Phe-Asn-Ser-Ala-Leu-Met-Phe-NH₂) and S2 (Ser-Gly-Pro-Tyr-Ser-Phe-Asn-Ser-Gly-Leu-Thr-Phe-NH₂). Furthermore, rather than dissecting only *Asterias rubens* radial nerve cords for peptide extraction as in the investigation made by Elphick et al. (1991a), starfish tube feet were also included and encompassed the bulk of the tissue dissected for extraction. The extraction medium was a solution of methanol, water, and acetic acid (90:9:1) (Husson et al., 2005). To delipidate the extract, it was mixed in a 1:1 ratio with ethyl acetate and *n*-hexane, before removal of the insoluble suspension by pipette.

The C18 SepPak cartridges were prepared by washing with 40% acetonitrile and water, both with 0.1% formic acid, prior to loading

approximately 5 ml of extract (0.1% formic acid). The loaded cartridges were then washed in water (0.1% formic acid), and the pigment removed with 30% acetonitrile (0.1% formic acid). The peptides were eluted with 80% acetonitrile (ACN). The first nine drops eluted from the cartridge were discarded, and the proceeding five to ten drops captured for analysis. After evaporation of the organic solvent from the purified extract overnight the final volume was approximately 0.5 ml.

The HPLC-MS system used for peptide detection incorporated an Agilent 1100 Series pumping system, a Vydac C8 chromatography column (Cat # 208TP54), an Atmospheric Pressure Interface (API) electrospray source, an ion trap mass analyzer, and a dynode-based system ion detector (Daly detector). Mass spectrometry parameters were set at default values: the nebulizer pressure was set to 40 psi, the flow rate was set to 0.5 mL/min, and the injection volume was set to 20 μ L.

After establishing a protocol for the identification of SALMFamide peptides, a protocol for the detection of the putative novel sea urchin neuropeptides was developed. Echinotocin and NGFFFamide were synthesized for use as standards in the development of an identification protocol. The peptides were synthesized by the Advanced Biotechnology Centre at Imperial College London. Both peptides were synthesized with an amidated C-terminal. Echinotocin was constructed with a disulfide bond between the first and sixth residues.

NGFFFamide (MW=629.7162)	NGFFF-NH ₂
Echinotocin (MW=965.1654)	CFISNCPKG-NH ₂

Synthetic peptide solutions were firstly subjected to the same extraction protocol that proved effective for detection of the SALMFamide neuropeptides. Mass spectrometric detection of each of the synthetic peptides using HPLC-MC was standardized using a 20 μ l injection of a 10 μ M peptide solution and 0.5%, 1%, 1.67%, 2% and 8.3% per minute gradients of organic solvent (see Table 8, Chapter 4). The full gradient for the identification of SALMFamide S1 ran at 8.3%/min from 20% to 45%, at 2%/min ACN from 45% to 65% in 10 minutes, and at 8.3%/min ACN from 65% to 90%. The full gradient for the identification of SALMFamide S2 ran at 2%/min from 30% to 90% in 30 minutes.

Capillary, skimmer, octopole, trap drive, and lens voltages, or 'expert parameter settings' were optimized for each synthetic peptide by injecting the mass spectrometer directly with a 1 μ M peptide solution, bypassing separation on the HPLC column (Appendix D). These parameters were then used to construct segmented gradients for the elution and detection of more than one peptide on a single run to minimise the amount of extract expended.

Ten specimens of *S. purpuratus* were dissected into body wall and viscera by Robert Burke's laboratory at the University of Victoria, British Columbia. The viscera and body wall were separately homogenized in a Waring blender, centrifuged (15000 x g for 10 minutes), the supernatant decanted and the pellet discarded, lyophilized overnight, and stored at -80°C prior to transport in separate vials on dry ice to Queen Mary,

University of London, via Fedex. The vials were combined for each tissue type, re-homogenized (Ultra-Turrax T25, max. 24000 rpm) and filtered and purified using a C18 SepPak column according to the same procedure used for the extraction of SALMFamide neuropeptides from starfish tube foot and radial nerve tissue.

The identification of putative peptides on detection by mass spectrometry was guided by correlating the putative peptide's elution point to the elution point of the synthetic peptide run on an identical gradient immediately after the extract. As a negative control, pure water was run on the same gradient immediately prior to the extract.

3.2.3 Methods: Immunohistochemistry protocols for the localization of echinotocin

Two different commercially available rabbit anti-vasopressin polyclonal antibodies were purchased for the localization and detection of echinotocin: one from Millipore (Cat. AB1565), another from Calbiochem (Cat. PC234L). The antibodies were stored at -20°C in 50% glycerol aliquots. Synthetic vasopressin, [Arg⁸]-vasopressin, was also purchased for use as a positive control (Cat. 2935, Tocris Bioscience).

Firstly, immunoblot assays were used to test the antibody's affinity for the peptide on a nitrocellulose membrane. Secondly, frozen tissue sections and sections prepared from paraffin-embedded tissues of *Echinus esculentus* tube foot and oesophagus were prepared to test for immunoreactivity with anti-vasopressin antibodies.

3.2.3.1 Immunoblot Assay Protocol

Immunoblot assays were used to test the immunoreactivity of both antibodies against synthetic peptide. The synthetic peptides were blotted onto a nitrocellulose membrane, washed in phosphate-buffered saline (PBS; pH 7.4) and blocked in 5% w/v Marvel™ milk powder for two hours. Unrelated synthetic peptides (Sp-GKamide and Sp-GnRH) were used as negative controls. Primary antibody, a dilution of 1:1000, was left to incubate overnight at 4°C in a blocking solution of MPBST (PBS; 1% Marvel™, 0.05% Tween 20). The membrane was given three ten minute washes in MPBST before incubation with the secondary antibody, alkaline phosphatase-labelled goat anti-rabbit immunoglobulin (AP-1000, Vector Laboratories), at a concentration of 1:1000 in MPBST (1% Marvel™, 0.05% Tween 20). The secondary antibody was incubated for two hours at room temperature. The membrane was given three ten minute washes in PBST (PBS; 0.05% Tween 20) and one ten minute wash in 100 mM Tris buffer (pH 9.5). To detect immunoreactivity, the membrane was submerged in BCIP/NBT substrate (12761220, Roche), diluted 1:100 in buffer (3.3% 3M NaCl; 10% 1M Tris pH 9.5). Immediately after staining the membrane strips were washed in water and air-dried.

3.2.3.2 Immunohistochemistry: Tissue fixation, sectioning and immunostaining

Sea urchin (*E. esculentus*) tube foot, oesophagus, and intestine tissues were fixed in water (4% w/v PFA), sea water (4% w/v PFA), and Bouin's fixative (saturated picric acid, 24% formalin, 5% acetic acid) and either embedded in paraffin or prepared for freezing (see below). The tissues fixed in 4% w/v PFA solutions were incubated overnight at 4°C, while the tissues fixed in Bouin's fixative were incubated in the fixative for several days at 4°C. The tissues were rinsed of fixative before stepped dehydration in ethanol (EtOH) (50%, 70%, 90%, and 2 x 100%) with gentle shaking for one hour per stage at room temperature. The dehydrated tissues were submerged twice in xylene, to displace the alcohol, at a maximum duration of five minutes each, before embedding in paraffin at 56°C for six to twelve hours depending on the size of the fixed tissue.

Sections of seven to twelve micrometres were prepared from the paraffin-embedded tissues using a rotary microtome (RM2145, Leica). The sections were left to adhere to SuperFrost® slides on a work bench heater set at a warm temperature for several hours before storage. Prior to staining, sections prepared were treated in xylene (three separate immersions for ten minutes each) before stepped rehydration in EtOH (2 x 100%, 90%, 70%, 50%, water) for ten minutes per step. Before immersion in 90% EtOH, the sections were treated with hydrogen peroxide in methanol (1% H₂O₂) for 30 minutes to eliminate endogenous peroxidase activity. The sections were

washed in normal PBS buffer, then PBS-X (PBS; 0.2% Triton X-100), before blocking, incubation with antibody, and staining.

To prepare fixed samples for freezing, the tissues were gradually infused with a 30% sucrose solution overnight, embedded in Tissue-Tek O.C.T. Compound 4583 tissue-embedding medium and frozen in liquid nitrogen. Sections of approximately 20 μm in thickness were prepared using a cryostat (Leica CM3050 S) and mounted on gelatin-coated SuperFrost® slides. The sections were dried for half an hour before storing at -20°C for no longer than 24 hours prior to three ten minutes washes with PBS, one ten minute PBS-X, and staining.

Prior to antibody incubation, slides were blocked in 5% w/v Normal Goat Serum (G9023, Sigma) PBS-X at room temperature for two hours. Primary antibodies (AB1565, Millipore; PC234L, Calbiochem) were diluted at concentrations of 1:500 and 1:1000 in 5% NGS PBS-X blocking buffer, and incubated with the prepared sections overnight. After three washes with PBS-X (0.05% Triton X-100) for 10 minutes each wash, the sections were incubated in secondary antibody, horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin (A-0545, Sigma), diluted 1:500 in 1.5% w/v NGS PBS-X (0.05% Triton X-100) for four hours at room temperature. The sections were washed again in PBST (PBS; 0.05% Tween 20) three times before using a 3,3'-diaminobenzidine/nickel staining kit according to the manufacturer's instructions (solution ratios: 1 substrate buffer: 2 DAB: 1 H_2O_2 : 1 Ni) (SK-4100, Vector Laboratories). Tissue section staining was carried out in the dark at room temperature for approximately ten minutes.

The staining reaction was stopped with PBS buffer before the sections were dehydrated in alcohol (10 minute stages of 50%, 70%, 90%, 2 x 100% EtOH), treated with twice with xylene (10 min) to remove the alcohol, covered with a coverslip mounted on DPX, and viewed under a light microscope.

Preabsorption controls, prepared in parallel to primary antibody incubation and negative controls without primary antibody, involved incubating the primary antibody at a 1:100 dilution in a 20 μ M echinotocin PBS buffered solution. The vasopressin antibody was incubated with synthetic echinotocin for two hours with gentle rocking at room temperature, before centrifugation and dilution to 1:1000. Positive controls of mouse brain sections containing vasopressinergic neurons, the paraventricular nucleus (PVN) and supraoptic nucleus (SON), were used to test the ability of the antibodies to recognise vasopressin in fixed tissues.

3.2.4 Methods: Pharmacological investigation of echinotocin and NGFFFamide myoactivity

Live specimens of the sea urchin *Echinus esculentus* Linnaeus (Class Echinoidea, Family Echinidae) were maintained in a sea water aquarium at approximately 4°C. Synthetic echinotocin and NGFFFamide were tested for myoactivity on *E. esculentus* oesophagus and tube foot preparations. All recordings of *E. esculentus* oesophagus and tube foot myoactivity were made using an isometric force transducer (Harvard Apparatus). When relaxed and fully extended, tube feet were severed from approximately

halfway along a sea urchin radial canal. The tube foot epithelial layer was either sheared using a blunt seeker probe or left intact. Because sea urchin tube feet which have been stripped of their epithelial layer are very fragile, for some of the preparations only half of the epithelium was removed (subsequently referred to in this study as 'semi-stripped' preparations). Oesophagi were removed and cut into 1.5 cm long segments. Although tube foot and oesophagus tissues were the easiest to dissect, intestine sections and muscles involved in lantern movement, including the compass depressor muscle and the lantern protractor and retractor muscles, were also dissected for investigation.

The preparations were tied using a silk ligature, suspended in an 11°C, 20 mL bath of aerated sea water supplied from the aquarium, before being connected to the force transducer. All tissues were maintained in the bath at a constant temperature for several hours until attaining a state of stable tension and before applying the synthetic peptide. Peptide solutions were freshly prepared at concentrations from 10 nM to 10 µM. The solutions were pipetted into the sea water bath at a series of concentration increments typically increasing by a factor of ten. The water bath was washed out with sea water (11°C) approximately ten minutes after applying each peptide solution. Individual applications of the synthetic peptides at high concentrations (3 µM) were recorded on separate tube foot and oesophagus preparations to observe the magnitude of responses independent of the effects of desensitization. Because individual responses were recorded when contraction of the preparation were clearly stronger than background variation, single observations of contractile activity were recorded and

included the response data. The transducers were calibrated with weights so that the force of contraction (mN) could be calculated from the number of units displaced on the recorder as a result of contractile activity. As a negative control, distilled water was pipetted into the bath to determine whether mechanical stimulation also induced contraction. Acetylcholine (ACh), which has previously demonstrated myo-excitatory activity on sea urchin tissue preparations (Florey and Cahill, 1980), and carbachol, a nicotinic and muscarinic cholinergic agonist, were also applied to the preparations. Both were used as positive controls and for pharmacological comparisons of efficacy, or strength of response for a given contraction, and potency, or the concentration of agent which induces 50% of the maximal response observed.

3.3 Results

Four neuropeptide homologues were identified by BLAST searching the sea urchin genome. The putative neuropeptides identified included homologues of vertebrate hormone vasotocin and the sea cucumber neuropeptide NGIWYamide. Genes encoding homologues of both the alpha and beta chains of the insect cysteine knot cuticle-hardening hormone bursicon were also identified in the genome by BLAST homology searching (SPU_003984 and SPU_017707) (Figure 16).

3.3.1 Identification of NGFFFamide

Analysis of the *S. purpuratus* genome using tBLASTn with a query constructed from three repeats of the sea cucumber NGIWYamide

neuropeptide resulted in the identification of a sequence ('KRNGFFFGKRNGFFFGKR') encoding two copies of a putative NGIWYamide-like peptide, Asn-Gly-Phe-Phe-Phe-NH₂ (or NGFFFamide, pronounced "Neg-famide"). An N-terminal signal peptide sequence was identified in the Gnomon prediction model (hmm5205), upstream of the peptide coding region. However, the putative NGFFFamide precursor gene was not contained within the set of genes predicted by the Glean3 algorithm. Therefore, the putative NGFFFamide precursor gene was included as one of a number of 'novel genes' identified as part of the sea urchin genome annotation project (SPU_030074) (Burke et al., 2006; Sodergren et al., 2006).

The amino acid sequence of the NGFFFamide precursor was then submitted as a tBLASTn query against a database of the adult radial nerve EST sequences. Corresponding 5-prime and 3-prime sequences of a cDNA clone (RNSP-5L15) from the *S. purpuratus* radial nerve cord library were identified, providing evidence of expression of the putative NGFFFamide precursor in neural tissue. A five prime EST corresponding to the three prime region of the precursor was also identified in an EST library of *S. purpuratus* 20 hour blastula stage embryos (GI:34799242).

3.3.2 Identification of the neurophysin domain

On submitting the EST nucleotide sequences as a BLASTx query against the NCBI protein database the 3-prime EST sequence strongly aligned to neurophysin sequences of vasopressin/oxytocin-like neuropeptide precursors. The presence of the neurophysin-encoding exon was also

supported by the Gnomon gene prediction model (hmm5205). The complete NGFFFamide encoding preproprotein was found to contain two copies of the putative neuropeptide NGFFFamide and a C-terminal neurophysin sequence of up to 108 amino acids (Figure 17).

3.3.3 Identification of echinotocin

A novel neuropeptide also containing a C-terminal neurophysin domain, denominated echinotocin (Cys-Phe-Ile-Ser-Asn-Cys-Pro-Lys-NH₂), was identified by BLAST searching the genome on the basis of sequence similarity to the vertebrate vasopressin neuropeptide amino acid sequence (Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly-NH₂). Although the precursor gene was not predicted by Gnomon gene prediction heuristics, the precursor was predicted by both GenScan and Glean3 gene prediction programs and was subsequently annotated in the sea urchin annotation database (SPU_006899). However, the echinotocin sequence was not found to be amongst the nucleotide sequences represented in adult or embryonic EST libraries. The putative echinotocin encoding preproprotein contains one copy of the neuropeptide, a neurophysin sequence of up to 95 amino acids, and a C-terminal glycoprotein or putative copeptin molecule of 30 amino acids (Figure 18).


```

Sp_beta  MPEATSNEMSTCRLAKVIILFTPKPPGIWLYTMLLLLLVLLC
Dm_beta  -----MHVQELLFVAAILVP
Dm_alpha -----MLRHLLRHENNKVFVLILILLYCVLVSIL
Sp_alpha -----MATLTTTQSSLLMLALILL

Sp_beta  GAGSVSRVSSLADGETCELSHGETTIDVEVVDDNLQRTLRCRK
Dm_beta  QCLRALRYSQGTGDENCETLKSEIHLIKEEFDELGRMQRTCNA
Dm_alpha KLCTAQPDSSVAATDN-DITHLGDDCQVTPVIHVLQYPGCVPK
Sp_alpha SLTTSS-----LVLADCRKLGLQYKLARPGCRPV

Sp_beta  RVQVNQCEGKCISQVSPTVLQ-HGFDKKCHCCREHGMVHKKVV
Dm_beta  DVIVNKCEGLCNSQVQPSVITPTGFLKECYCCRESFLKEKVIT
Dm_alpha PIPSFACVGRCASYIQVSGSKIWQMERSCMCCQESGEREAAVS
Sp_alpha TLDSVGCRGTCSGYTRISPNNYLEVERSCTCCQEMGFLERTQR

Sp_beta  MTNCYD-HALGITDPDF-THQVTLKQPEACRCQICTF-----
Dm_beta  LTHCYDPDGTRLTSPEMGSMDIRLREPTECKCFKCGDFTR---
Dm_alpha LF-----CPKVKPGERKFKKVLTKAPLECMCRPCTSIEESGI
Sp_alpha LQ-----CPTLNPPFR---DVTYRIPRRCSCRPCRSVASVSR

Sp_beta  -----
Dm_beta  -----
Dm_alpha PQEIAGYSDEGPLNNHFRRIALQ
Sp_alpha VQTLEDLRLG-----

```

Figure 16. ClustalX alignment of two putative alpha and beta sea urchin bursicon-like preprohormone sequences with the two *Drosophila* alpha and beta bursicon preprohormone sequences. N-terminal signal peptide sequences are identified in bold italic and the eleven cysteines common to each of the four sequences are underlined.

```

1          aaacttga
9   ttcgatcgcggggagcataacaccaggagagatttttgatgttaagagagtggacaaaac
69  acagcatttgatcattggttagtggaactaggtggaagagaattcaatggttcgcctcgc
129 tggaaacaacagtggagaataggcacgagctatactgtaggagaaaaggttttcgttagatt
189 taggtgcctcgagaaaagatcgaggtcgattggtttaaggaaggtaccgaaggaaaag
249 atgggctacgagagacggatactgcgagcgttattgtcgcatttaatagtactagcatca
    M G Y E R R I L R T L L S I L I V L A S 20
309 ttcgtcacagtctatggagaaagagatagcaactttatgcaacagaagcaattcaggaat
    F V T V Y G E R D S N F M Q Q K Q F R N 40
369 atagtaccgtcaccgcttatccaaaaatggcgggaaaatcgaatgggtcccgcggaggag
    I V P S P L I Q K W R E N R M G P A E E 60
429 aagactagcaacgaacagtggagggacgagctccttagtaaccttagaaaactgcttaga
    K T S N E Q W R D E L L S N L R N V L R 80
489 aaacacaacgcatcaccatcatcacgaagcagagacagaaacgacatcacagcctacggc
    K H N A S P S S R S R D R T D I T A Y G 100
549 ctccaagaacctatgcagcagcttcctgcagacgtaacggccgatcagttgttcatacta
    L Q E P M Q Q L P A D V T A D Q L F I L 120
609 gagggcgctgtaaaactcaccagggaaaactacgaggaggaaacgccattgacgaggat
    E G A V N S P R E N Y E E E T P I D E D 140
669 aagagaaacggatttttcttcggtaaacgtaacggatttttcttcgggaagaggtcggat
    K R N G F F F G K R N G F F F G K R S D 160
729 agcgatgcctcatcaaccaagatggacgatgacagactacccaaatcgaatcatcagga
    S D A S S T K M D D D R L P K Y E S S G 180
789 tcatttgataagtgcagaccatgcggtccaggccggcagggacgatgcgtaatggtgggt
    S F D K C R P C G P G R Q G R C V M V G 200
849 acatggtgtagtcacctatctcggtgctacctattcacaccgaagccgcagcgtgtatg
    T C C S P L F G C Y L F T P E A A A C M 220
909 acagaagatgtgtcccgtgtcaactcaatgacgcttcctgtggccttgacggaaggtgt
    T E D V S P C Q L N A P S C G L A G K C 240
969 gtaccgatgggatttgctgttctgcccagagggcgcctgccaccttgaccgcacctgt
    V A D G I C C S A A E G A C H L D P T C 260
1029 acgtcgatgcatcattaaattaatttttgacgcattggtgtatttttcaaatcaatatcatg
    T S M S L N * 266
1089 gtgtacagtttttctgaccctgtaatttctctagctccgaaaataatcttcttactcg
1149 accgttccatttaccttttctttggtcgaattcattaggaacgtgcttacaaaatgt
1209 tgtgtaattgtactatcaactcctggttattatgagcatttattcgcatttgagattgtt
1269 gtatttccgtactcaaaactgaaatccaacgcgct

```

Figure 17. NGFFFamide precursor. The putative precursor gene contains 1309 nucleotides and the preprohormone sequence contains 266 residues. There are four exons separated at splice sites indicated in bold underlined font. The signal peptide sequence is indicated in blue font, the two peptides in red font, dibasic cleavage site ('KR') in green font, and the neurophysin sequence in purple. Neurophysin cysteine residues are underlined. The translation stop codon is indicated by an asterisk.

```

1   atgatgtccgtcaagagtatagtcacctggttgtttctgtcattgggtctggctctatgg
   M M S V K S I V T C L F L S L V L A L W      20
61  ataggggggagcttcgcctgttttatctccaactgcccaaaggggggtaaaagatcaaat
   I G G S F A C F I S N C P K G G K R S N      40
121 tctcgtccacttagacagtgcctcgaatgctgaccaggcggttaggaagggtgcatgggc
   S R P L R Q C L E C G P G G V G R C M G      60
181 ccagggatctgtgctgctgacccaacgattggctgtcacatcaacacacacacacactgtcc
   P G I C C G P T I G C H I N T Q H T L S      80
241 tgtatgtagagaaaacgagatctcaacgccatgtgaaactcccaggaacccttgtcagact
   C M R E N E I S T P C E L P G N P C Q T      100
301 gtccaagtgtacatgtggagcaatgggtgtatgctgcaatagtaattcttgttcagaa
   V P S G T C G A M G V C C N S N S C S E      120
361 gatgcatcctgtctgatgattgaagaggatgactcctcaaaagatttgagcagatgagt
   D A S C L M I E E D D S L K R F E Q M S      140
421 cgagaggaaaacggttcgacgaggaaagacctccgggttaaacttctcgatttacttttg
   R E E N G S T R K D L R V K L L D L L L      160
481 aacatgcaggatcaataa
   N M Q D Q *      165

```

Figure 18. Echinotocin precursor. The putative precursor gene contains 499 nucleotides and the preprohormone sequence contains 165 residues. There are three exons separated at splice sites indicated in bold underlined font. The signal peptide sequence is indicated in blue font, echinotocin in red font, a single dibasic cleavage site ('KR') in green font, and the neurophysin sequence in purple. Neurophysin cysteine residues are underlined. A putative copeptin-like sequence is underlined in purple. The translation stop codon is indicated by an asterisk.

3.3.4 Mass spectrometry: detection of starfish SALMFamides and sea urchin NGFFFamide

Both SALMFamide peptides were successfully detected in the 0.1% formic acid starfish tube foot and radial nerve extract when injected into the HPLC-MS system (Figures 19 & 20). SALMFamide S1 was detected at approximately 70% ACN from a gradient of 8.33%/min ACN, 0.1% formic acid, at a 0.5 mL/min flow rate (Figure 19). SALMFamide S2 was detected at approximately 50-51% ACN from a gradient of 2%/min ACN, 0.1% formic acid, at a 0.5 mL/min flow rate (Figure 20).

Synthetic NGFFFamide eluted at between 39% and 46% on a 0.5% ACN/min gradient, at between 44% and 46% on a 1% ACN/min gradient, and at between 44% and 47% on a 2% ACN/min gradient. Synthetic echinotocin eluted at between 32% and 39% on a 1%/min ACN gradient, at between 37% and 39% on a 1.67%/min ACN gradient, and at approximately 42% on a 2%/min ACN gradient. The peptide masses identified by mass spectrometry, in protonated form, are compared to the calculated masses in Table 6.

A molecule with a mass corresponding to that of NGFFFamide was identified in extracts of *S. purpuratus* visceral tissue, coinciding with the elution point of the synthetic peptide. The molecule (630.3 u) appeared in all mass spectrometry traces and was absent from the blank standard (water, 0.1% formic acid) run prior to the extract and confirmed on running synthetic NGFFFamide on an identical gradient after the extract. On

running the body wall extract, a peak corresponding to NGFFFamide was identified again, at a stronger intensity than the peak identified from the visceral extract (2×10^5 cps) (at 32 to 32.9 min) (Figure 21). There were no mass spectrometric peaks corresponding to the mass of echinotocin, coincident with the elution point of synthetic echinotocin, detected in either the sea urchin body wall or visceral tissue purified extracts.

Table 6: Predicated and observed synthetic peptide masses of echinotocin and NGFFFamide detected using HPLC-MS. The difference observed between observed and calculated mass gave There was an error range of 0.082% between the observed and calculated masses.

Peptide	Predicted Mass	Observed Mass
NGFFFamide	630.7162 (H ⁺)	630.3 (H ⁺)
Echinotocin	966.1654 (H ⁺)	965.4 (H ⁺), 483.2 (2H ⁺)

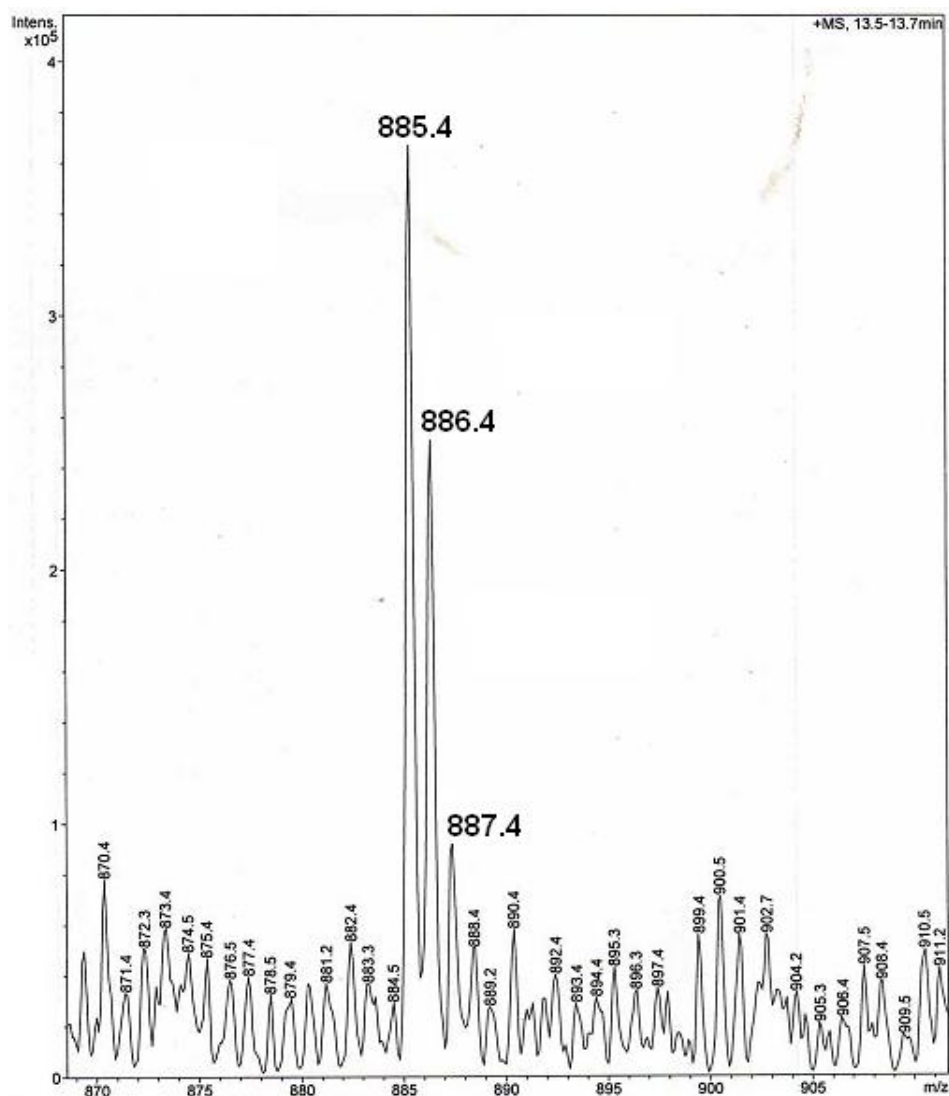


Figure 19. SALMFamide S1 (MW=885.4, predicted MW=885.1) detected in *Asterias rubens* tube foot and radial nerve tissue extract. The peptide was eluted at approximately 70% acetonitrile (ACN), 0.1% formic acid. The smaller peaks (886.4, 887.4) correspond to larger masses representing isotopic variation in atomic carbon mass.

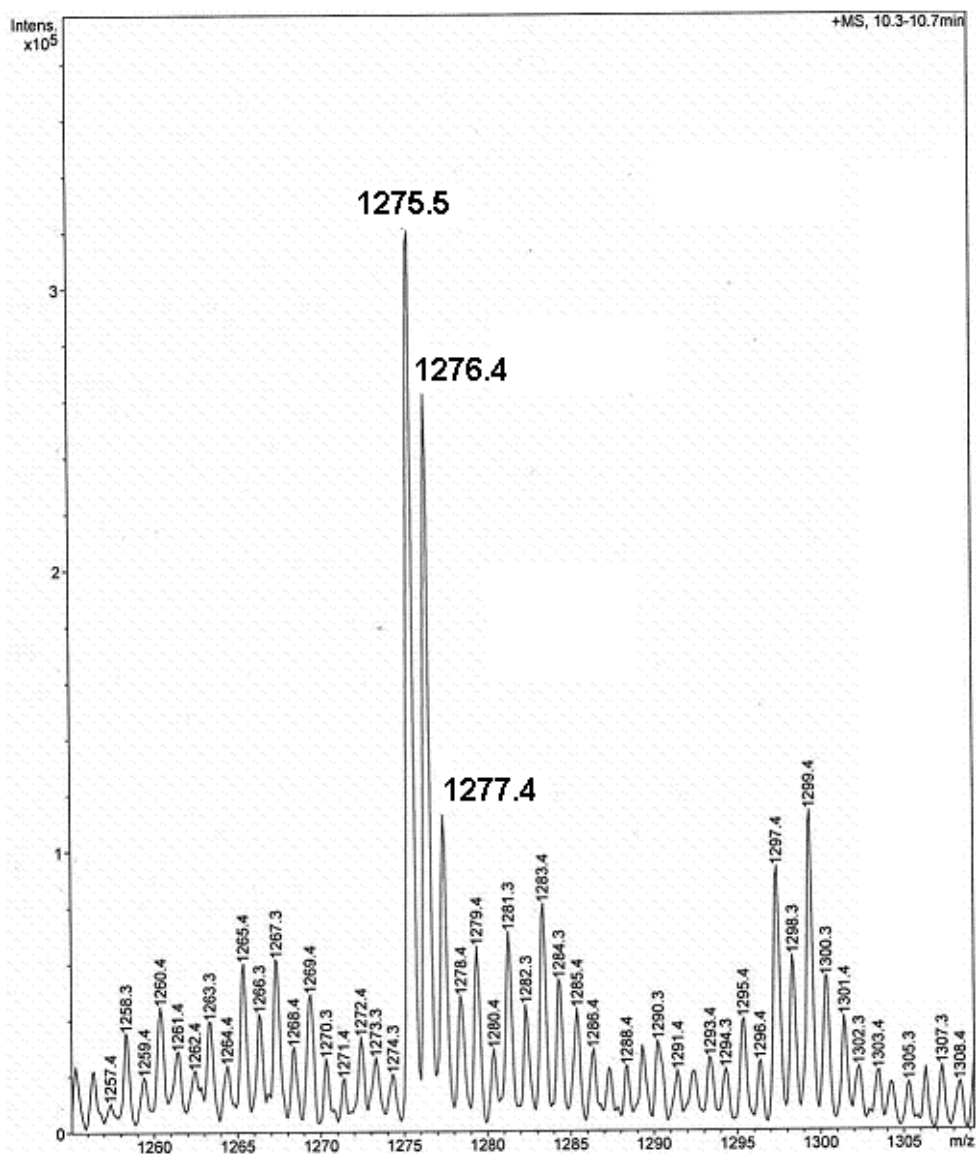


Figure 20. SALMFamide S2 (MW=1275.5, predicted MW=1275.4) detected in *Asterias rubens* tube foot and radial nerve tissue extract. The peptide was eluted at approximately 50-51% acetonitrile (ACN), 0.1% formic acid. The smaller peaks (1276.4, 1277.4) correspond to larger masses representing isotopic variation in atomic carbon mass.

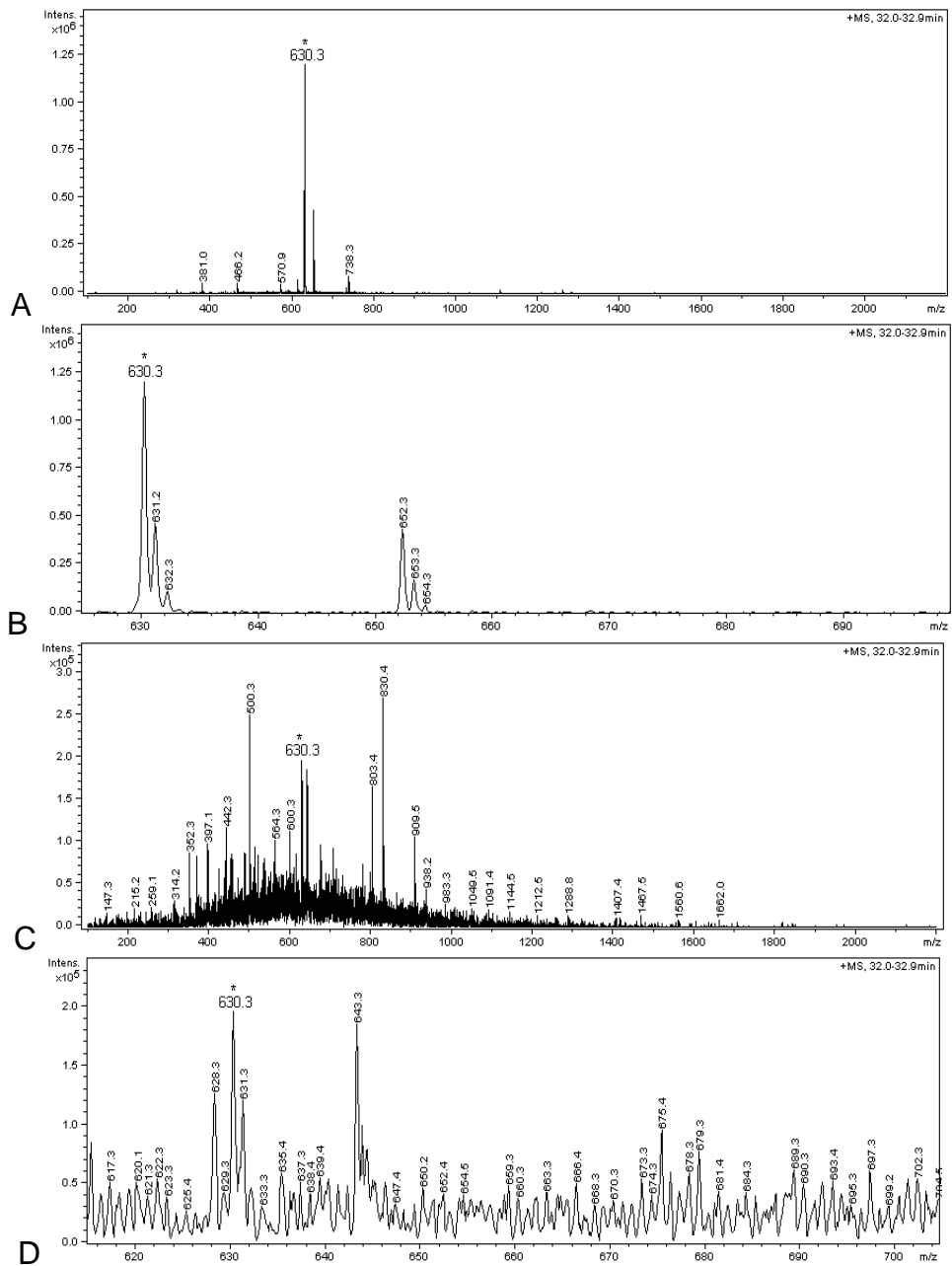


Figure 21. Detection of NGFFamide using HPLC-MS.

A) Full mass spectrum showing detection of synthetic NGFFamide peptide (630.3) (*).

B) Magnified mass spectrum view of 17A, showing NGFFamide synthetic peptide (630.3) (*) detected at 32.0 to 32.9 minutes.

C) Full mass spectrum showing detection of NGFFamide (630.3) (*). Spectrum obtained from sea urchin *S. purpuratus* body wall extract.

D) Magnified mass spectrum view of 17C. Spectrum obtained from sea urchin *S. purpuratus* body wall extract. Putative NGFFamide neuropeptide (630.3) (*) was detected at 32 to 32.9 minutes.

3.3.5 Results: Immunohistochemistry

The anti-vasopressin antibody AB1565 recognized 1ng to 1µg of echinotocin blotted on to a nitrocellulose membrane. Both synthetic vasopressin and echinotocin were also found to be immunoreactive to the anti-vasopressin antibody PC234L when bound to a nitrocellulose membrane. As an additional positive control, antibody PC234L successfully recognized synthetic vasopressin on an immunoblot assay (results not shown). However, both primary antibodies preabsorbed with vasopressin failed to eliminate immunoreactivity to neuronal staining in the paraventricular and supraoptic nuclei in mouse brain sections, where vasopressin is synthesized (Figures 22, 23 & 24).

Sea urchin tube foot and oesophagus tissues sections demonstrated immunoreactivity to anti-vasopressin antibodies; however, the areas of positive staining were not eliminated in preabsorption controls. In comparison to negative controls, the areas of primary staining were strongest in sections stained with the PC234L anti-vasopressin antibody. Positive staining was observed in sections of the oesophagus mucosal endocrine plexus and cells bordering the basiepithelial nerve plexus (Figure 25A-E). These sections were obtained from paraffin-embedded tissue fixed with Bouin's fixative. Staining was also stronger in the mucosal endocrine plexus and longitudinal muscle layer in frozen sections of oesophagus tissue (not shown). The staining was absent from negative controls, except for the preabsorption control. Staining in the mucosal endocrine plexus of the intestine was also observed in both frozen sections and sections

prepared from paraffin-embedded tissue (not shown). Although the staining was not eliminated in preabsorption controls, tube foot tissues prepared by cryostat showed significantly stronger staining than observed in negative controls without primary or secondary in both the inner and outer epithelial layer (not shown).

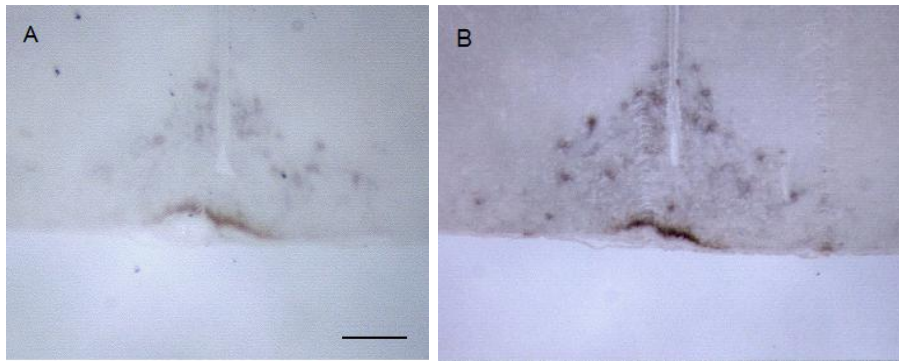


Figure 22. Positive control tests for the AB1565 vasopressin antibody. Scale bar=200 μ m
 A) Anti-vasopressin AB1565 antibody immunoreactivity in the mouse brain, transverse section, at the PVN and SON.
 B) Preabsorption control of anti-vasopressin AB1565 antibody immunoreactivity in the mouse brain, transverse section, at the PVN and SON.

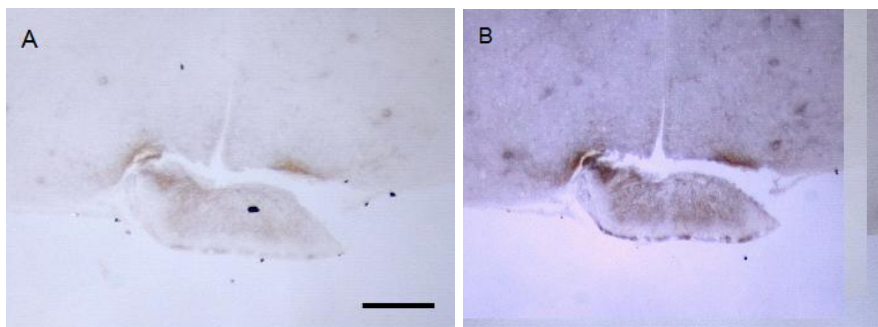


Figure 23. Positive control tests for the AB1565 vasopressin antibody. Scale bar=100 μ m
 A) Anti-vasopressin AB1565 antibody immunoreactivity in the mouse brain, transverse section, at the PVN and SON (transverse section).
 B) Preabsorption control of anti-vasopressin AB1565 antibody immunoreactivity in the mouse brain at the PVN and SON (transverse section).

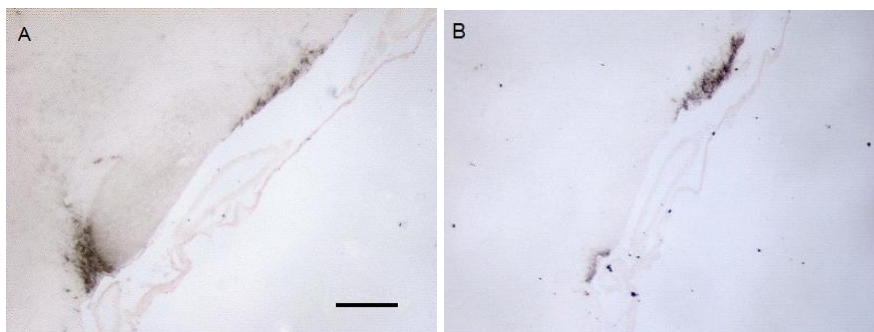


Figure 24. Positive control tests for the PC234L vasopressin antibody. Scale bar=100 μ m
 A) Anti-vasopressin PC234L antibody immunoreactivity in the mouse brain at the PVN and SON (longitudinal section).
 B) Preabsorption control of anti-vasopressin PC234L antibody immunoreactivity in the mouse brain at the PVN and SON (longitudinal section).

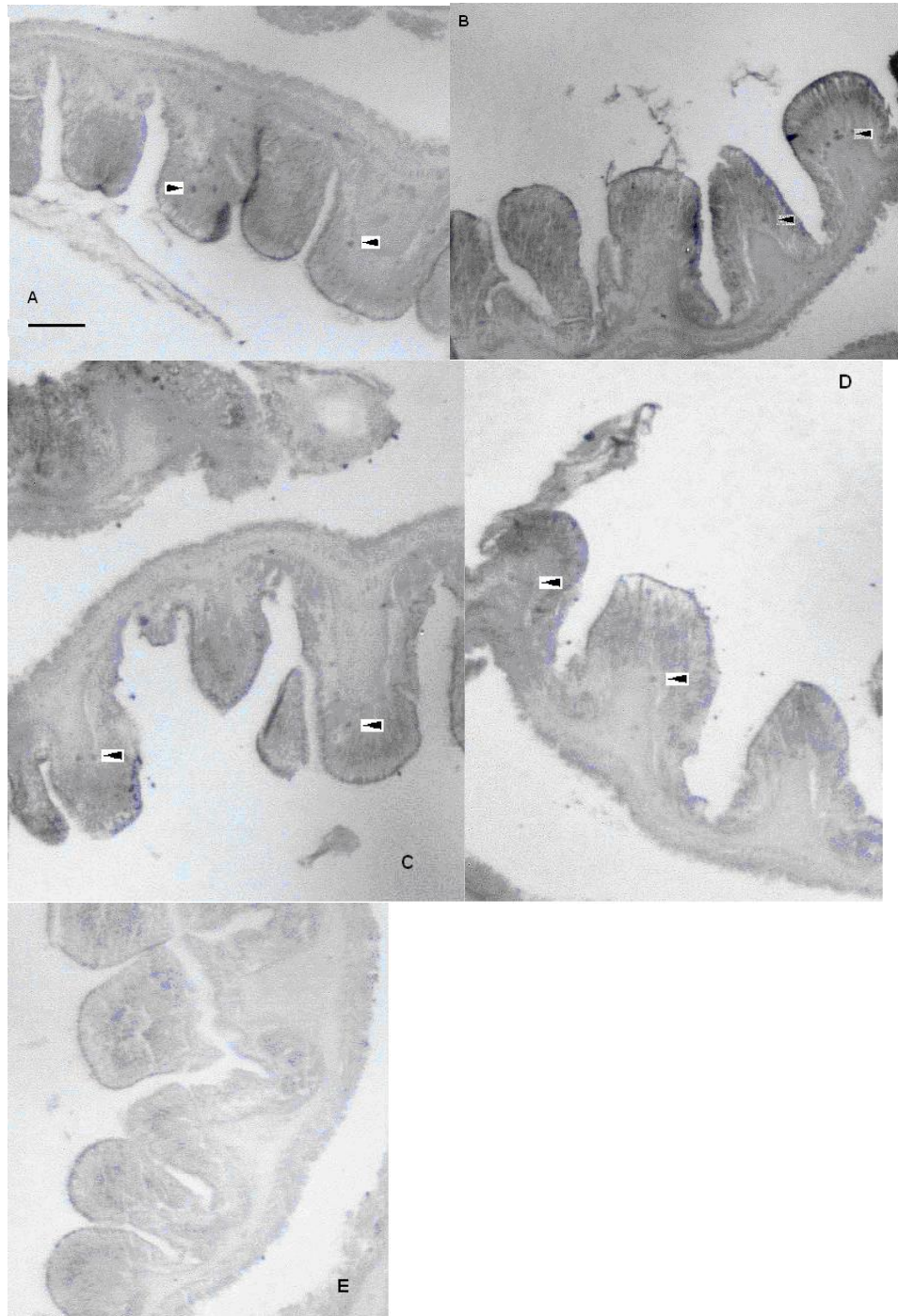


Figure 25. Sea urchin oesophagus transverse section prepared from paraffin-embedded tissue stained with rabbit anti-vasopressin. Oesophagus fixed with Bouin's solution. Arrows indicate areas of staining. Scale bar=200 μ m

A-D) Staining in the mucosal endocrine plexus and the region of the basiepithelial plexus in sections incubated with anti-vasopressin PC234L antibody (1:1000).

E) Negative control without primary anti-vasopressin PC234L antibody.

3.3.6 Results: Characterization of NGFFFamide and echinotocin

myoactivity

The two novel putative neuropeptides echinotocin and NGFFFamide both caused contraction of sea urchin and oesophagus *in vitro* tissue preparations, demonstrating that both peptides have myoactive properties. Of the two peptides, NGFFFamide was more potent in causing contraction of both preparations than echinotocin, and therefore had the greater efficacy. Acetylcholine and carbachol also induced contraction of sea urchin tube foot and oesophagus tissues.

3.3.6.1 Acetylcholine and carbachol

Acetylcholine caused contraction of sea urchin oesophagus, intestine, lantern protractor muscle, lantern retractor muscle, compass depressor muscle, and 'semi-stripped', 'stripped', and 'unstripped' tube foot preparations (Table 7). 'Semi-stripped' tube feet were more responsive to acetylcholine and carbachol than unstripped preparations (not shown). 'Semi-stripped' tube foot responses to acetylcholine increased in contractile strength with concentration (Figure 26). Contractions were observed at concentrations ranging from 100 nM to 300 μ M. The highest average force was recorded from applications of 100 μ M ACh (\bar{x} = 0.719 mN, s.e.m. = 0.281, n = 8). 'Semi-stripped' tube foot responses to carbachol also increased in contractile strength with concentration (Figure 27). Contractions in response to carbachol were observed at concentrations

ranging from 10 nM to 30 μ M. The highest force was recorded from an application of 30 μ M carbachol (0.300 mN, n=1).

Starfish tissues were also responsive to acetylcholine. *A. rubens* stripped tube foot preparations also contracted in response to acetylcholine concentrations from 1 nM to 300 μ M recorded on an isometric transducer. Subsidiary experiments also demonstrated that apical muscle preparations were responsive to acetylcholine concentrations from 1 μ M to 1 mM (Gill, unpublished).

3.3.6.2 Echinotocin and NGFFFamide

Both NGFFFamide and echinotocin induced contraction of *E. esculentus* oesophagus and unstripped tube foot preparations (Figures 28 & 29). NGFFFamide, at a concentration of 3 μ M, induced contraction of tube feet (\bar{x} = 1.350 mN, s.e.m. = 0.210, n = 3) and at a greater magnitude than echinotocin (\bar{x} = 0.810 mN, s.e.m. = 0.42, n = 3) (representative traces illustrated in Figure 28B & 29B). Similar differences in efficacy, or strength of contraction at a fixed concentration, were observed when testing both peptides on oesophagus preparations. NGFFFamide at a concentration of 3 μ M induced contractions (\bar{x} = 1.466 mN, s.e.m. = 0.230, n = 3) of a greater magnitude than echinotocin (\bar{x} = 0.714 mN, s.e.m. = 0.01, n = 3) (representative traces illustrated in Figure 28A & 29A) .

NGFFFamide induced dose-dependent contractions of *E. esculentus* oesophagus preparations at concentrations from 10 pM to 1 μ M and of

unstripped tube foot preparations at a higher range of concentrations from 10 nM to 1 μ M (Figure 28C). NGFFFamide induced contraction of 'semi-stripped' tube foot preparations at a similar range of concentrations from 10 nM to 10 μ M (Figure 30), with a highest mean contraction of 0.230 mN (s.e.m. = 0.129, n = 15) at a 3 μ M peptide concentration. Sea urchin intestine also contracted in response to NGFFFamide at a concentration of 3 μ M (0.308 mN, n = 1). In subsidiary experiments, NGFFFamide was also found to induce contraction of starfish cardiac stomach at a concentration of 1 μ M.

Echinotocin induced dose-dependent contractions of *E. esculentus* oesophagus preparations at concentrations from 10 nM to 100 nM, before desensitization of the preparation, and of unstripped tube foot preparations at a similar range of concentrations from 10 nM to 1 μ M (Figure 29C). In subsidiary experiments on preparations from the starfish *Asterias rubens*, echinotocin induced contraction of starfish cardiac stomach at a concentration of 1 μ M (Navaratnam, unpublished) and stripped tube foot (1.39 mN, n = 1) at a concentration of 1 μ M.

Table 7: Acetylcholine-induced contraction of sea urchin tissue preparations.

Tissue	Concentration (μM)	Force (mN)	n
Tube foot;			
- Semi-stripped	100	0.719	8
- Stripped	200	0.173	1
- Unstripped	100	2.622	2
Oesophagus	3	1.653	3
Intestine	3	1.169	1
Lantern protractor muscle	5	0.437	1
Lantern Retractor Muscle	3	1.209	1
Compass Depressor Muscle	6	5.127	1

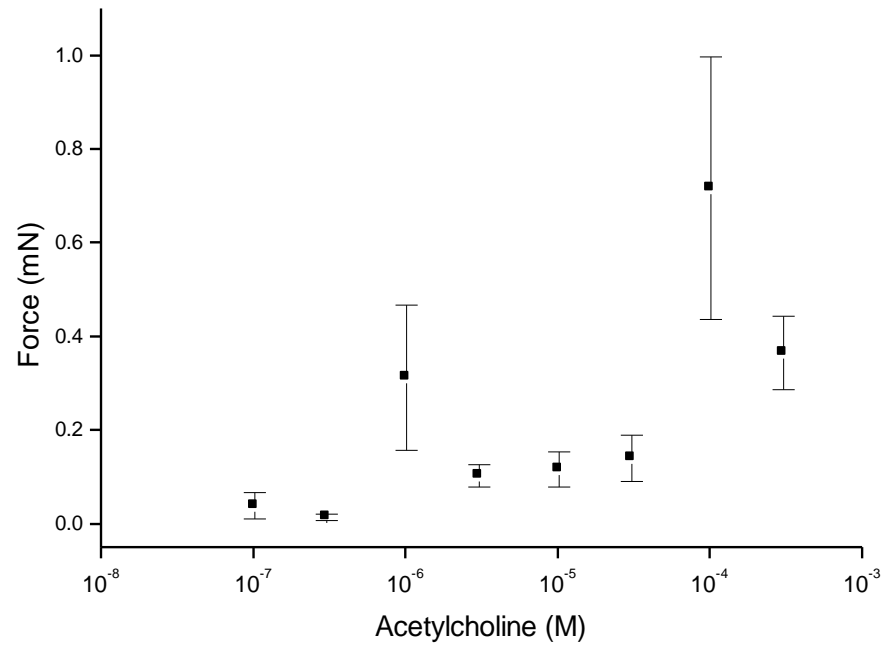


Figure 26. Acetylcholine-induced contraction of *Echinus esculentus* 'semi-stripped' tube foot preparations measured using an isometric transducer. Tube foot preparations contracted in response to 100 nM (n = 5), 300 nM (n = 5), 1 μ M (n = 14), 3 μ M (n = 15), 10 μ M (n = 15), 30 μ M (n = 5), 100 μ M (n = 8), and 300 μ M (n = 4).

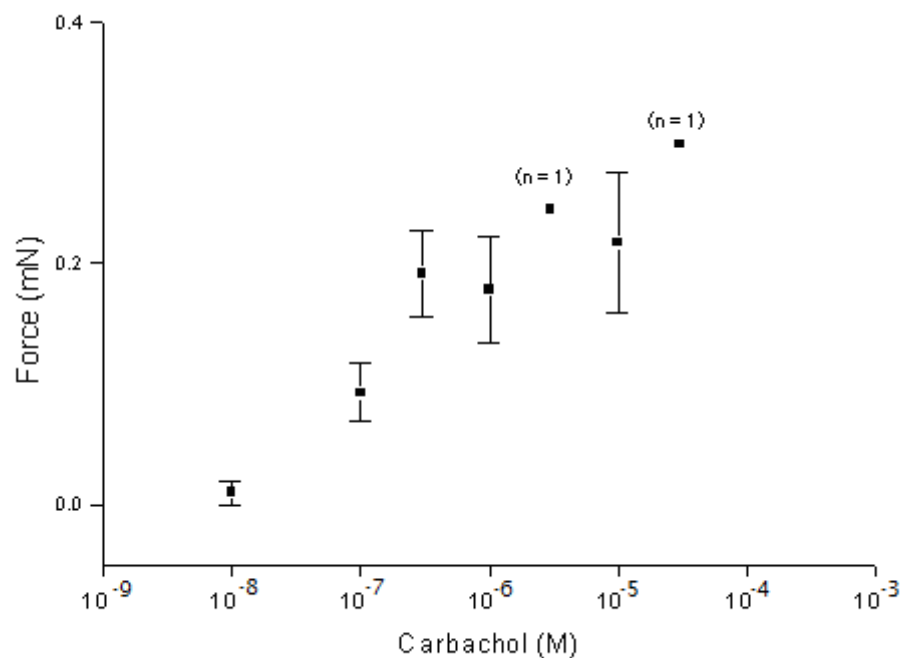


Figure 27. Carbachol induced contraction of *Echinus esculentus* 'semi-stripped' tube foot preparations measured using an isometric transducer. Tube foot preparations contracted in response to 10 nM (n = 3), 100 nM (n = 7), 300 nM (n = 3), 1 μ M (n = 8), 3 μ M (n = 5), 10 μ M (n = 7), and 30 μ M (n = 1).

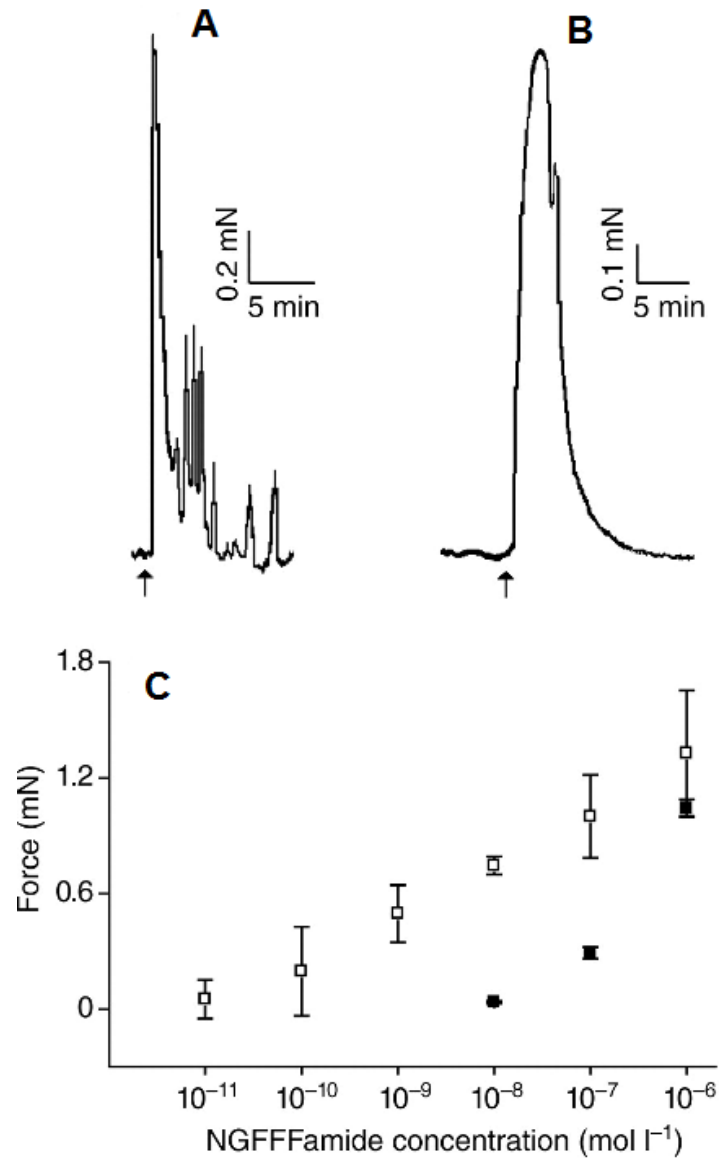


Figure 28. NGFFFamide causes contraction of oesophagus and 'unstripped' tube foot preparations from the sea urchin *Echinus esculentus*. Traces (A-D) show representative responses to peptide applications at $3 \times 10^{-6} \text{ mol l}^{-1}$ (arrows).

A) NGFFFamide induced contraction of oesophagus.

B) NGFFFamide induced contraction of unstripped tube foot.

C) Dose-dependent effect of NGFFFamide on tube foot (filled squares) and oesophagus (open squares) preparations. Peptide concentrations were increased by a factor of ten. Data points are mean values ($n = 4$) with bars showing s.e.m.

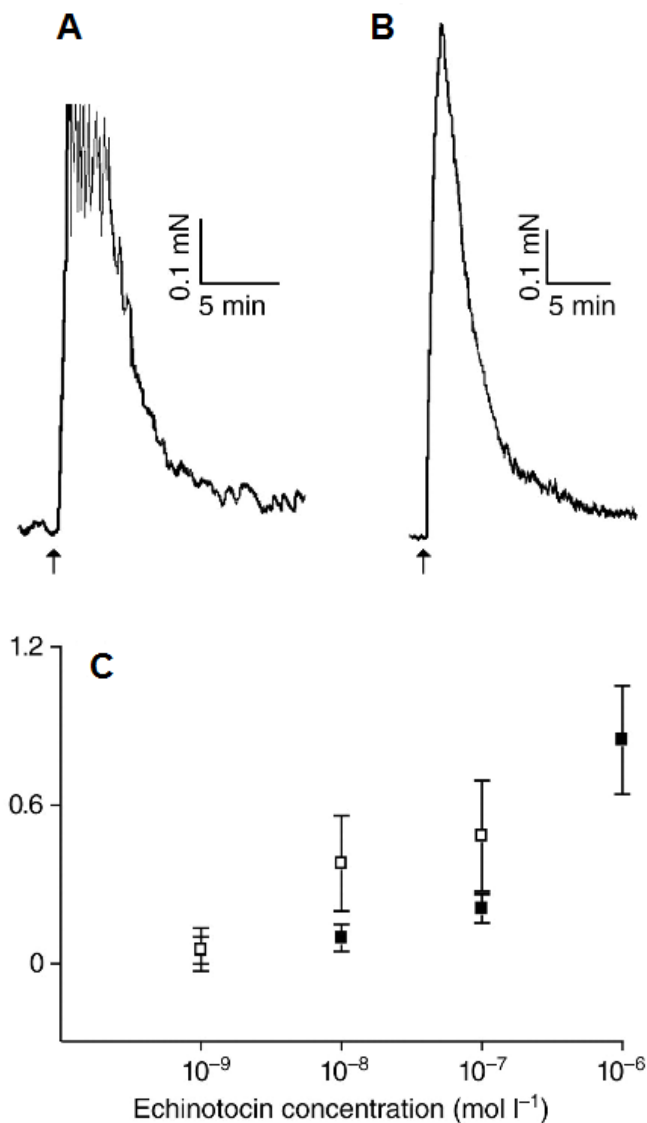


Figure 29. Echinotocin causes contraction of oesophagus and 'unstripped' tube foot preparations from the sea urchin *Echinus esculentus*. Traces (A-D) show representative responses to peptide applications at $3 \times 10^{-6} \text{ mol l}^{-1}$ (arrows).
 A) Echinotocin induced contraction of oesophagus.
 B) Echinotocin induced contraction of unstripped tube foot.
 C) Dose-dependent effect of echinotocin on tube foot (filled squares) and oesophagus (open squares) preparations. Oesophagus contractions after 100 nM were not included due to tissue desensitization. Peptide concentrations were increased by a factor of ten. Data points are mean values ($n = 4$) with bars showing s.e.m.

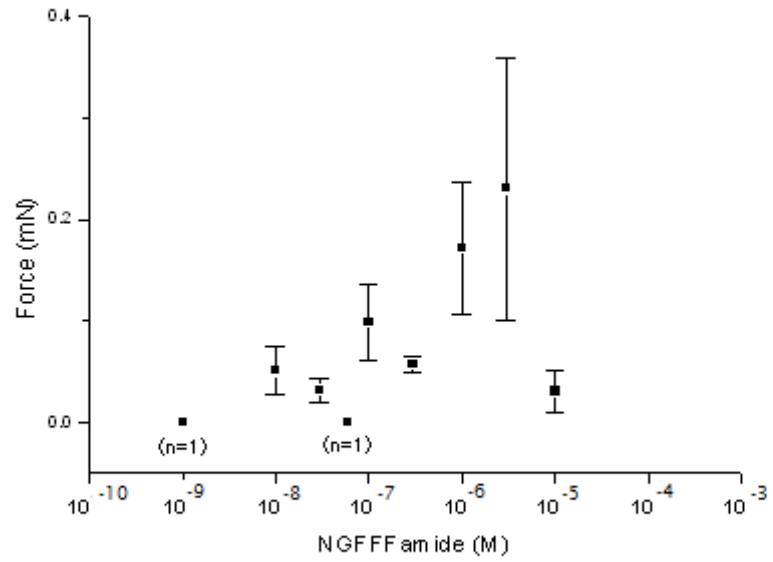


Figure 30. NGFFFamide induced contraction of *Echinus esculentus* 'semi-stripped' tube foot preparations measured using an isometric transducer. Tube foot preparations contracted in response to 1 nM (n = 1), 10 nM (n = 6), 30 nM (n = 9), 60 nM (n = 1), 100 nM (n = 17), 300 nM (n = 19), 1 μ M (n = 24), 3 μ M (n = 15), and 10 μ M (n = 9).

3.4 Discussion

Two novel putative neuropeptide precursor genes containing neurophysin domains have been identified in the genome of the sea urchin *S. purpuratus*. The genes identified encode the putative neuropeptides NGFFFamide and echinotocin, homologues of the sea cucumber neuropeptide NGIWYamide and vasopressin/oxytocin-type neuropeptides, respectively. Furthermore, two precursors of bursicon-like peptides have been identified. NGFFFamide is structurally-unrelated to vasopressin/oxytocin-like peptides and a sequence encoding the NGFFFamide precursor was detected in a neural EST library. Furthermore, a peptide with the same molecular mass as synthetic NGFFFamide was detected in a mass spectrometric analysis of *S. purpuratus* tissue extracts, demonstrating that the amidated pentapeptide is also expressed in sea urchins. Prior to this study, only vasopressin/oxytocin-like family neuropeptide genes have been known to encode neurophysin sequences. Although echinotocin expression was not detected in the neural EST library or through mass spectrometric analysis, the peptide may still be expressed in neural or endocrine sea urchin tissues.

3.4.1 Sea urchin glycoprotein hormones

The subunits of the insect dimeric hormone bursicon are encoded by separate precursors (Luo et al., 2005). Consistent with the structure of bursicon, the putative sea urchin bursicon-like dimer subunits appear to be also encoded by two separate precursors (see Figure 16). Bursicon targets

GPCRs characterized by leucine-rich repeat regions, which as a subgroup of the GPCR superfamily includes the glycoprotein gonadotropins and relaxin/insulin-like growth factor hormone receptors found in vertebrates (Luo et al., 2005).

There may be other putative ligands to sea urchin glycoprotein-like and relaxin/insulin-like receptors. Mita et al (2009) purified gonad-stimulating substance (GSS) from radial nerves of the starfish *Asterina pectinifera* and identified the dimeric peptide as a member of the relaxin-like hormone family. Therefore, like bursicon, GSS may bind a sea urchin receptor with homology to vertebrate leucine-rich repeat containing GPCRs.

Furthermore, Burke et al. (2006) reported the identification of two sea urchin insulin-family peptides; *Sp-IGF1* and *Sp-IGF2*, which may also bind the receptors within the same subgroup. Three putative receptors were identified in the sea urchin genome by homology searching; Sp_115678957 showing homology to vertebrate glycoprotein receptors, and both SPU_015740 and SPU_003527 showing homology to relaxin/insulin-like receptors) (see Figure 14, Chapter Two). However, the insulin-like growth factors identified may bind sea urchin tyrosine kinase receptors (Lapraz et al., 2006).

3.4.2 Neurophysin-encoding precursors in the sea urchin genome

The putative echinotocin neuropeptide is the first vasopressin/oxytocin-like neuropeptide to be identified in the genome of an echinoderm and the putative peptide may share a common ancestor with other members of the

vasopressin/oxytocin neuropeptide family (Figure 31). Oxytocin and vasopressin are neuropeptides encoded by two separate genes which are thought to have originated by gene duplication at least 450 mya in primitive jawed vertebrates. The ancestral neuropeptide precursor to both oxytocin and vasopressin may have encoded a peptide similar to vasotocin, a vasopressin-like neuropeptide expressed in non-mammalian vertebrates (Urano et al., 1992; Van Kesteren et al., 1992; Hoyle, 1999). The hypothesis is supported by the presence of only a single vasotocin-encoding gene in the lamprey *Lampetra japonica* (Suzuki et al., 1995).

All described vasopressin/oxytocin-like neuropeptides have a ring structure formed by a disulfide bridge between two cysteine residues (C1-C6), a structural characteristic likely to be shared by echinotocin. With the exception of vasopressin/oxytocin family neuropeptides identified in the nematode *C. elegans* and urochordates *Styela plicata* and *Ciona intestinalis*, all vasopressin/oxytocin-like neuropeptides, including the echinotocin neuropeptide, have nine residues (Kawada et al., 2008; Ukena et al., 2008) (Figure 31). Amino acid substitutions in positions three and eight of the two nonapeptides affect their activity and specificity. A basic residue in the eighth position is common to vasopressin-like hormones, whereas a neutral amino acid in the same position is shared by oxytocin-like hormones. In vertebrates a hydrophobic isoleucine residue in position three is essential for the stimulation of oxytocin receptors, while a basic arginine or lysine residue in position eight is essential for the stimulation of vasopressin receptors (Chini et al., 1995). Echinotocin is similar to

oxytocin in having a hydrophobic residue (Ile) in position three and is similar to vasopressin in having a basic residue (Lys) in position eight.

In chordates the vasopressin/oxytocin preprohormone gene consists of three exons, the first encoding a signal peptide sequence, vasopressin/oxytocin, and the first nine residues of neurophysin, the second encoding residues 10-76 of neurophysin and the third containing residues 77-95 of neurophysin. The echinotocin neurophysin domain is encoded by three exons whereas the NGFFFamide-associated neurophysin domain is only encoded by two. Alignment of the echinotocin and NGFFFamide precursors with eight other vasopressin/oxytocin-like peptide precursors from four different phyla using ClustalX is shown in Figure 32. Furthermore, NGFFFamide neurophysin did not cluster with any of the representative neurophysin sequences, suggesting that the NGFFFamide-encoding gene arose either through gene duplication before protostome and deuterostome divergence or through recent sequence divergence within the echinoderm phylum (Figure 33). The phylogenetic tree illustrated shows that the NGFFFamide neurophysin domain is no more similar to other chordate neurophysin domain sequences, including echinotocin neurophysin, than to neurophysin domains encoded by other protostome precursors. Sequence alignment analysis alone may be an insufficient means for generating hypotheses about the NGFFFamide precursor's evolutionary origin.

Figure 32 also illustrates the common residues shared by the eleven neurophysin sequences. Putative models of NGFFFamide and echinotocin neurophysins were generated using the Swiss PDB Viewer (DeepView) and

SWISS-MODEL from the crystal structure of peptide-bound oxytocin (Figure 34). Neurophysins contain 14 cysteine residues forming as many as seven disulfide bridges (Nicolas et al., 1980; de Bree and Burbach, 1998). All 14 cysteine residues involved in the formation of disulfide bonds are crucial for tertiary structure and are structurally conserved in both echinotocin and NGFFFamide neurophysins. Neurophysins form stable dimers, a characteristic thought to be important in storing and transporting the bound hormone. Furthermore, neurophysin domain mutations affecting tertiary structure may disrupt endopeptidase-mediated cleavage of vasopressin/oxytocin-like peptide (Legros and Geenen, 1996; De Bree, 2000). The ability of neurophysins to form stable dimers is greatly enhanced when bound to the neuropeptide (Nicolas et al., 1978). Two homologous regions within the neurophysin sequence are required for dimerization. The two regions may be linked by a single α -helix and a disulfide bond between C⁹ and C⁵³ for echinotocin neurophysin and C¹⁰ and C⁵² for NGFFFamide neurophysin (de Bree and Burbach, 1998).

The hormone-binding pocket of the neurophysin molecule is formed by the interaction between specific neurophysin residues (de Bree and Burbach, 1998). The strongest interaction with the hormone occurs in a salt bridge between the γCOO^- of a glutamate residue (Glu⁷⁸ in oxytocin neurophysin), which is conserved in both echinotocin and NGFFFamide neurophysins, and the αNH_3 of the peptide (de Bree and Burbach, 1998). The γCOO^- group of the glutamate residue also interacts with the side chain of Arg³⁹ in neurophysin, a residue conserved in echinotocin neurophysin but not in NGFFFamide neurophysin, where the position is taken by an aspartic acid

residue (Figures 32 & 34). Only the N-terminal three amino acids of the hormone are required for neurophysin binding. Most important in formation of the hormone-neurophysin complex is a hydrogen bond between a cysteine residue (C⁴⁴ in vasotocin) and the tyrosine residue (Y²) of the neuropeptide. Echinotocin has an aromatic phenylalanine residue (F²) in position two, a characteristic shared with the molluscan vasopressin/oxytocin-like peptide and 'nematocin' (CFLNSCPY), a putative nematode vasopressin/oxytocin-like peptide from *C. elegans*. The putative 'nematocin' precursor was identified in the NCBI protein database (GI:86564869) and the putative peptide differs from other vasopressin/oxytocin-like peptides in consisting of only eight residues and by the absence of a C-terminal glycine residue for peptide amidation. NGFFFamide has a remarkably different residue structure to vasopressin/oxytocin-like peptides, and lacks the C1 and C6 residues necessary to form a disulfide bridge. However, the presence of aromatic phenylalanine residues in NGFFFamide may be of structural importance to the binding of NGFFFamide to neurophysin.

Interestingly, there is remarkable conservation in the distance between consecutive cysteine residues in all neurophysins (Figure 32). The glutamic acid residue in the fourth position from the seventh cysteine in mammalian vasopressin/oxytocin neurophysins, which is essential for binding the hormone, is conserved in both the NGFFFamide and echinotocin neurophysins, as well as a proline residue N-terminal to the eighth cysteine residue (Figures 32 & 34). Interestingly an asparagine residue (N⁴⁸ in vasopressin neurophysin) is also retained in echinotocin, vasopressin, Lys-

conopressin, cephalotocin and annetocin neurophysins, but is absent from NGFFFamide neurophysin (Figures 32 & 34). NGFFFamide neurophysin deviates from the vertebrate neurophysins in two places, perhaps most importantly close to the residues involved in binding the hormone, and this is noticeable in the neurophysin homology model (Figure 34B). Between the seventh and eighth cysteine residues there has been a deletion of two residues. The second deviation of NGFFFamide neurophysin from vertebrate neurophysin occurs in the insertion (or deletion) of two residues between the twelfth and thirteenth cysteine residues. Echinotocin deviates from vertebrate neurophysin with a two residue insertion between the ninth and tenth cysteine residues.

Although absent from the NGFFFamide precursor, the C-terminal region of the echinotocin precursor shows homology to the vertebrate glycoprotein copeptin. Copeptin, which is encoded by the C-terminal region of vertebrate neurophysin precursor genes and typically consists of 22-39 residues, is characterized in chordates by a conserved Leu-Leu-Arg-Leu hydrophobic amino acid sequence (de Bree and Burbach, 1998). It is separately cleaved and potentially acts synergistically in neurotransmission and neuromodulation (van den Hooff et al., 1990). Although this short conserved sequence is absent from the putative echinotocin copeptin, there are six strongly hydrophobic residues (LLLRLVQL) in the C-terminal sequence. The *N*-glycosylation site (N⁶-X-T⁸), conserved in chordates, but absent from molluscan models, is also present in the echinotocin putative copeptin/glycoprotein. Although the cleaved neurophysin and glycoprotein molecules have previously only been identified in their processed state in

mammalian vertebrates, a dibasic cleavage site is present between the echinotocin neurophysin and copeptin/glycoprotein sequences. It is still uncertain at what point in evolutionary history the C-terminal region of the precursors attained the function of a glycoprotein.

Furthermore, a putative vasotocin-like G-protein coupled receptor (GPCR) was identified in the sea urchin genome (GI:115899360) (Chapter Two). Phylogenetic analysis places the putative receptor in a GPCR clade shared by the vasopressin and oxytocin receptors of vertebrates, suggesting the receptors arose through duplication during chordate evolution (Burke et al., 2006).

3.4.3 Mass spectrometric analysis: Detection of a putative NGFFFamide peptide

The analysis was the first mass spectrometric study of echinoderm neuropeptides since Elphick et al. (1991a) identified the two SALMFamide neuropeptides from the starfish *Asterias rubens* and *Asterias forbesi*. The results of that experiment were successfully replicated in this study using the starfish *Asterias rubens* with a modified peptide extraction protocol.

Analysis of sea urchin tissue extract using LC-MS provides evidence of the expression of the putative sea urchin neuropeptide NGFFFamide. The peptide was expressed at a greater intensity in the extract derived from the sea urchin body wall than in the extract derived from visceral tissues. The sea urchin test, or body wall, includes the radial nerve cords and axons

innervating the tube feet, while visceral tissue includes neurons innervating the enteric system, referred to as the enteric nervous system. Therefore, the difference in intensity may be reflective of higher levels of expression in the neurons of the nerve cords, test, and tube feet.

The absence of echinotocin in the purified sea urchin tissue extracts may reflect expression at concentrations not discernible above background intensities when analysed using mass spectrometry, which be a consequence of low levels of expression outside the reproductive season when the tissues were dissected. It is also noteworthy that the echinotocin precursor cDNA was absent from both the EST neural expression library derived from adult radial nerve cords and a transcriptome database assembled as part of the sea urchin genome project

(http://www.systemix.org/Data/Strongylocentrotus_purpuratus/Genome_tiling/).

3.4.4 Discussion: Immunohistochemical investigation of echinotocin expression

The immunoblot assays performed indicated that both antibodies (AB1565 and PC234L) recognize echinotocin but not unrelated and structurally dissimilar peptides (Sp-GKamide and Sp-GnRH). Furthermore, antibody PC234L recognized synthetic vasopressin. Where positive immunoreactivity was evident in sections of sea urchin tissue, the staining was absent from negative controls but present in preabsorption controls. Both preabsorbed polyclonal antibodies retained immunoreactivity when

tested on PVN and SON sections of the mouse brain (Figures 22-24), suggesting that the immunoreactivity observed may be a non-specific marker of vasopressin-expressing neurons. The results indicate that expression of echinotocin may not be detectable by immunohistochemistry or that any specific vasopressin-like immunoreactivity detected in the *E. esculentus* sea urchin and tube foot tissues may have been masked by non-specific staining. The results also support the possibility that the putative neuropeptide echinotocin may not be expressed in the sea urchin, or the tissues investigated, or that it may be rhythmically expressed, potentially coinciding with the sea urchin spawning period in the summer. Hence, further investigation is required to provide evidence of the expression of a vasopressin/oxytocin--like peptide in an echinoderm.

3.4.5 Discussion: Contractile myoactivity of echinotocin and NGFFFamide

Acetylcholine, carbachol, echinotocin, and NGFFFamide all induced contraction of sea urchin tube foot and oesophagus tissue preparations. The novel peptides NGFFFamide and echinotocin may be acting as natural ligands, targeting specific receptors expressed by the muscle cells of the tissue preparation. However, they may also be acting as full or partial non-specific agonists, activating receptors specific for alternate ligands, or as physiological agonists, inducing contractile responses similar to those observed after applying acetylcholine by activating different receptors.

Myoactivity has previously proven to be a useful bioassay for the determination of bioactivity for several peptides for which a secondary or alternative primary physiological role has later been identified. For instance, allatotropins have physiological roles as hormone mediators important in arthropod development and reproduction, as well as functions related to myoactivity and gut contraction (Elekonich and Horodyski, 2003). Cholecystokinin provides another example of a neuropeptide having a myo-regulatory role in addition to roles in the regulation of feeding behaviour (Dufresne et al., 2006). Furthermore, cholecystokinin is processed differentially in the CNS neurons of the brain to yield an active eight residue C-terminally truncated form (Dockray, 1980). Other brain-gut peptides include glucagon, gastrin, and the tachykinins. Thus, while the peptides predicted in this study may have myoactive properties and roles in the physiological regulation of echinoderm muscular tissue, the peptides may not be limited to such a role and may have equally important functions in the regulation of behaviour, reproduction, and neural integration in sea urchins and other echinoderms.

Acetylcholine, the major neuromuscular transmitter in vertebrates, has previously been found to have a similar role in both protostome and deuterostome species where the molecule also acts as a neuromuscular transmitter (Twarog, 1954; Mendes et al., 1970), including echinoderms (Bolt and Ewer, 1963; Welsh, 1966; Pentreath and Cottrell, 1968; Florey et al., 1975; Florey and Cahill, 1980). Although both nicotinic and muscarinic acetylcholine receptor-like proteins have been identified in the sea urchin genome their tissue and cell-specific expression patterns are unknown

(Burke et al., 2006). The neurotransmitter is synthesized from choline and acetyl-CoA by choline acetyltransferase and degraded by acetylcholinesterase. Homologues of both of these proteins have also been identified in the sea urchin genome (Burke et al., 2006) and a cholinesterase has been localized on the muscle fibres of sea urchin tube feet (Florey et al., 1975). Florey et al. (1975) detected the neurotransmitter in echinoderm radial nerves and demonstrated contraction of echinoderm tube feet in response to acetylcholine at concentrations as low as 550 nM but responses were observed more frequently at micromolar concentrations. Carbachol was found by Florey et al (1975) to be a more effective than acetylcholine at contracting sea urchin tube feet. Consistent with Florey's observations in this study carbachol was found to be ten-fold more potent than acetylcholine (Figures 26 & 27).

Aside from the myoactive actions of acetylcholine as an echinoderm neurotransmitter, other hormones and neurotransmitters have induced responses in *in vitro* bioassays prepared from echinoderm tissues. Noradrenaline and adrenaline induce variable responses (Welsh, 1966). GABA, despite being an inhibitory neurotransmitter in both vertebrate and arthropod neuromuscular systems, can induce either relaxation or contraction depending on the tissue investigated (Florey et al., 1975; Devlin, 2001). Alongside SALMFamide neuropeptides, nitric oxide (NO) has also been shown to be a powerful muscle relaxant of starfish muscle preparations (Elphick and Melarange, 2001).

The pharmacological study provides evidence that both NGFFFamide and echinotocin evoke contraction of sea urchin muscle preparations.

NGFFFamide has a greater potency and efficacy than echinotocin on both tube foot and oesophagus tissues. NGFFFamide also induced contractions of 'semi-stripped' tube foot preparations. However, high variability in the responses of the preparations was likely to be attributable to the procedure involved in stripping the tube feet during tissue preparation. Starfish tube feet have been observed to contract on application of the sea cucumber neuropeptide NGIWYamide, a holothurian homologue of NGFFFamide (Saha et al., 2006). In that study, tube foot contractile responses at NGIWYamide concentrations of 100 μM were also stronger than those induced by acetylcholine at the same concentration, suggesting that NGIWYamide-like peptide may have a more important role than acetylcholine in the contraction of echinoderm muscle. NGIWYamide also caused contraction of sea urchin stripped tube foot preparations (Gill, unpublished), and contraction of 'semi-stripped' sea urchin tube foot at a concentration of 10 μM (0.215 mN, n = 1) and oesophagus at a concentration of 3 μM (0.111 mN, n = 1).

Although NGFFFamide was found to induce contraction of starfish cardiac stomach, the peptide surprisingly induced relaxation of starfish stripped tube foot tissue at concentrations from 100 nM to 10 μM , but failed to induce a response from unstripped starfish tube foot (Gill & Navaratnam, unpublished).

Echinoderm tube feet have a number of behavioural and physiological functions, most importantly securing the organism to the substrate, but they may also have a role in chemosensory perception, photoreception and photoperiodicity (Yoshida, 1966; Blevins and Johnsen, 2004; Sodergren et al., 2006). To understand the physiological action of putative neurotransmitters or neuromodulators in causing contraction or relaxation of tube foot tissues a brief description of tube foot anatomy is needed. The layers from the lumen to the external epithelium include an inner epithelial layer, the circular and longitudinal muscle layers, a connective tissue layer, the nerve plexus, and an outer epithelial layer. Kawaguti (1964) observed an absence of neuromuscular synapses and, in response, Florey and Cahill (1977) have suggested that nerve terminals from the plexus terminate on or within the layer of connective tissue. Florey and Cahill hypothesized that acetylcholine acts on the muscle layer of echinoderm tube feet by traversing or diffusing across the connective tissue layer. If both echinotocin and NGFFFamide are secreted by neurons they may act in an analogous way, targeting the muscle layer from neurohormonal terminals located across a connective tissue barrier. Alternatively, the peptides may be released from neurons within the luminal epithelium or the connective tissue layer itself.

The sea cucumber NGIWYamide neuropeptide has also been observed to stiffen sea cucumber dermal tissue (Birenheide et al., 1998), suggesting a role for the peptide in the biological phenomenon of 'mutable collagenous tissue'. A number of hypotheses have been suggested to explain the biochemical processes responsible for the phenomenon, including the

action of a 'tensilin' protein on collagenous tissue (Wilkie, 2002; Tipper et al., 2003) and neuromuscular transmission (Elphick and Melarange, 2001). Previous studies have suggested a neurohormonal mechanism for echinoderm muscle contraction through the innervation of muscle fibres from the luminal epithelium (Kawaguti, 1964), from isolated neurosecretory granule-containing cells within the muscle layer (Coleman, 1969; Cobb and Pentreath, 1977; Florey and Cahill, 1977; Cobb, 1978) or even from isolated juxtaligamental cells within a connective tissue layer separating the nerve plexus from the muscle layer in tube feet (Santos et al., 2005). Localization studies of NGFFFamide expression in tube feet and oesophagus tissues will contribute to understanding the mechanism of the peptide's possible role in collagenous stiffening and muscle activity in all classes of echinoderm.

NGIWIamide has also been identified, functionally characterized, and denominated 'cubifrin' in a study by Kato et al (2009). Cubifrin was found to induce spawning behaviour and germinal vesicle breakdown in the presence of ovarian wall. The effect is comparable to that of the gonad-stimulating substance (GSS) peptide heterodimer (Mita et al., 2009).

Moreover, expression of the NGFFFamide precursor gene in the radial nerve provides evidence that the peptide is neurosecretory. The NGFFFamide coding sequence is found in both radial nerve and embryonic EST data and has been detected by mass spectrometric analysis of *S. purpuratus* body wall and visceral tissue. Therefore, the precursor gene may be expressed early in sea urchin development prior to development of

the adult nervous system. The lack of expression data for echinotocin, by contrast, may be reflective of comparably low expression and of the EST dataset's quality, an absence of neural expression, or even of an absence of expression in the sea urchin. Previous studies have successfully investigated the expression of NGIWAYamide, the possible sea cucumber NGFFFamide orthologue, in different classes of the echinoderm phylum. Saha et al., (2006) also investigated NGIWAYamide expression in the starfish and observed NGIWAYamide-like immunoreactivity in the radial nerves, marginal nerves, and tube feet. If NGFFFamide is functionally homologous, in having a similar physiological role, or even orthologous, in originating from a common gene, to NGIWAYamide, it may have a similar expression distribution throughout the sea urchin nervous system.

Vasopressin/oxytocin family neuropeptides appear to show functional homology in regulating muscle activity. The role in muscle regulation is often linked to the regulation of reproductive behaviour (Van Kesteren et al., 1995; Fujino et al., 1999; Goodson and Bass, 2000; Tessmar-Raible et al., 2007). Oxytocin has sex-specific roles in vertebrates, notably contraction of uterine smooth muscle and the mammary gland and in social behaviour, including pair bonding (Donaldson and Young, 2008).

Vasopressin/oxytocin-like peptides are known to have myoactive properties across both vertebrate and invertebrate phyla. Vasopressin is an osmoregulator and anti-diuretic hormone and echinotocin may demonstrate functional homology in the echinoderm phylum. A vasotocin-like peptide, *Styela* SOP, identified in *Styela plicata* of the deuterostome urochordate

subphylum, may have a homologous osmoregulatory role (Ukena et al., 2008). SOP has been shown in both *in vivo* and *in vitro* preparations to induce contraction of the inhalant and exhalant siphons responsible for maintaining favourable osmotic conditions. Hypothetically, sea urchins may react to unfavourable hypo-osmotic conditions by contracting their tube feet and reducing water-flux. Hence, it is possible that echinotocin may have a functionally homologous role in echinoderms as an osmoregulator.

Genes encoding neuropeptides showing structural and functional homology to vasopressin/oxytocin have also been identified in protostomian invertebrates. In molluscs the vasotocin-like neuropeptide conopressin causes contraction of the vas deferens (Van Kesteren et al., 1992). In annelids the peptide annetocin induces egg-laying behaviour (Fujino et al., 1999). Tessmar-Raible et al. (2007) have postulated an evolutionary linkage between the regulation of reproductive behaviour by vasopressin/oxytocin-like neuropeptides observed in protostomes and deuterostomes and an involvement of the same secretory neurons in photoreception and photoperiodic behaviour and physiology. The hypothesis is supported by the observation of cell-specific expression of microRNA and transcription factors by vasopressin/oxytocin-like neuropeptide secreting neurons in the annelid *Platynereis dimurilii* and the zebrafish *Danio rerio* (Tessmar-Raible et al., 2007). Other studies have also suggested a link between vasopressin/oxytocin family neuropeptide neurosecretory cells and photoreception in organisms as distantly related as killifish and insects (Taylor, 1986; Thompson and Bacon, 1991; Ota et al.,

1999; Gozdowska et al., 2006). It would be interesting to investigate whether a similar relationship between photoreception and neurosecretory neurons exists in the echinoderm, where the neurosecretory centres are located, and whether vasopressin/oxytocin-like peptide secretory pathways have acquired a modified role. Opsins have been identified in the sea urchin genome and it is possible that tube feet may act as photoreceptive organs. Therefore, photoreceptors targeting unidentified neurohormonal secretory centres may be located on echinoderm tube feet and, consequently, neurohormones including echinotocin may be involved in the regulation of echinoderm reproductive cycles.

The full diversity of neurophysin-associated neuropeptide precursors in invertebrates is unrealised and the identification of two novel neurophysin-encoding neuropeptide precursors will advance investigation into evolution of the gene family and into the functional roles of the associated peptides. Furthermore, their identification will instigate investigation of their function and expression in the unique echinoderm pentaradial nervous system.

Peptide	Sequence
Echinotocin ¹	Cys-Phe-Ile-Ser-Asn-Cys-Pro-Lys-Gly-NH ₂
Vasopressin ²	Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly-NH ₂
Oxytocin ²	Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly-NH ₂
Arg-vasotocin ³	Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Arg-Gly-NH ₂
Ciona-VP ⁴	Cys-Phe-Phe-Arg-Asp-Cys-Ser-Asn-Met-Asp-Trp-Tyr-Arg
Styela-VP ⁵	Cys-Tyr-Ile-Ser-Asp-Cys-Pro-Asn-Ser-Arg-Phe-Trp-Ser-Thr-NH ₂
Conopressin ⁶	Cys-Phe-Ile-Arg-Asn-Cys-Pro-Arg-Gly-NH ₂
Cephalotocin ⁷	Cys-Tyr-Phe-Arg-Asn-Cys-Pro-Ile-Gly-NH ₂
Annetocin ⁸	Cys-Phe-Val-Arg-Asn-Cys-Pro-Thr-Gly-NH ₂
Inotocin ⁹	Cys-Leu-Ile-Thr-Asn-Cys-Pro-Arg-Gly-NH ₂
Nematocin ¹⁰	Cys-Phe-Leu-Asn-Ser-Cys-Pro-Tyr

Figure 31. Alignment of vasopressin/oxytocin-like peptides including sea urchin echinotocin and nematode 'nematocin'. Basic residues are labelled in blue, acidic residues in purple, polar residue in green, nonpolar residues in orange, and aromatic residues in red. The sequences were obtained from peptide precursors isolated from ¹*Strongylocentrotus purpuratus* (this study), ²*Homo sapiens* (Light and Du Vigneaud, 1958), ³*Lethenteron japonicum* (Suzuki et al., 1995), ⁴*Ciona intestinalis* (Kawada et al., 2008), ⁵*Styela plicata* (Ukena et al., 2008), ⁶*Lymnaea stagnalis* (Van Kesteren et al., 1992), ⁷*Octopus vulgaris* (Reich, 1992), ⁸*Eisenia foetida* (Oumi et al., 1994), ⁹*Locusta migratoria* (Proux et al., 1987; Aikins et al., 2008; Stafflinger et al., 2008), and ¹⁰*Caenorhabditis elegans* (GenBank NP_001033548, GI:86564869).



Figure 32. ClustalX multiple alignment of the putative precursor sequences of the *S. purpuratus* NGFFFamide precursor, the *S. purpuratus* echinotocin precursor and precursors of vasopressin/oxytocin-like peptides in other species. Signal peptides are shown in blue, neuropeptides are shown in red, dibasic cleavage sites are shown in green and neurophysin-like domains, including co-peptin-like sequence, are shown in purple. The conserved cysteine residues in the neurophysin-like domains are underlined and numbered 1–14. The precursors of vasopressin/oxytocin-like peptides from other species include precursors of human vasopressin and oxytocin (Mohr et al., 1985), vasotocin from the lamprey *Lethenteron japonicum* (Suzuki et al., 1995), an oxytocin-like peptide (SOP) from the urochordate *Styela plicata* (Ukena et al., 2008), Lys-conopressin from the mollusc *Lymnaea stagnalis* (Van Kesteren et al., 1992), cephalotocin from the mollusc *Octopus vulgaris* (Reich, 1992), annetocin from the annelid *Eisenia foetida* (Oumi et al., 1994) and inotocin from the arthropod (insect) *Tribolium castaneum* (Aikins et al., 2008; Stafflinger et al., 2008). Also included is a precursor from the nematode *Caenorhabditis elegans* (GenBank: NP_001033548, GI:86564869) that has not been reported previously in the literature. The nematode precursor contains an unusual putative oxytocin-like peptide comprising just eight residues (CFLNSCPY), which has been named 'nematocin'.

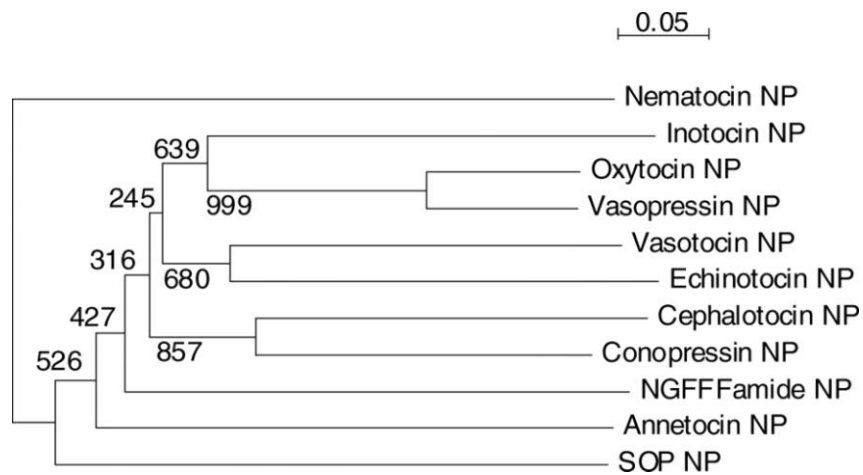


Figure 33. Neighbour-joining tree (with bootstrap values) based on a ClustalX multiple alignment of neurophysin (NP) sequences, incorporating residues from the first to the fourteenth conserved cysteines. The tree shows that the neurophysin domain of the *S. purpuratus* NGFFFamide precursor does not have a higher level of overall sequence similarity with the neurophysin domain of the *S. purpuratus* echinotocin precursor than with the neurophysins from other species.

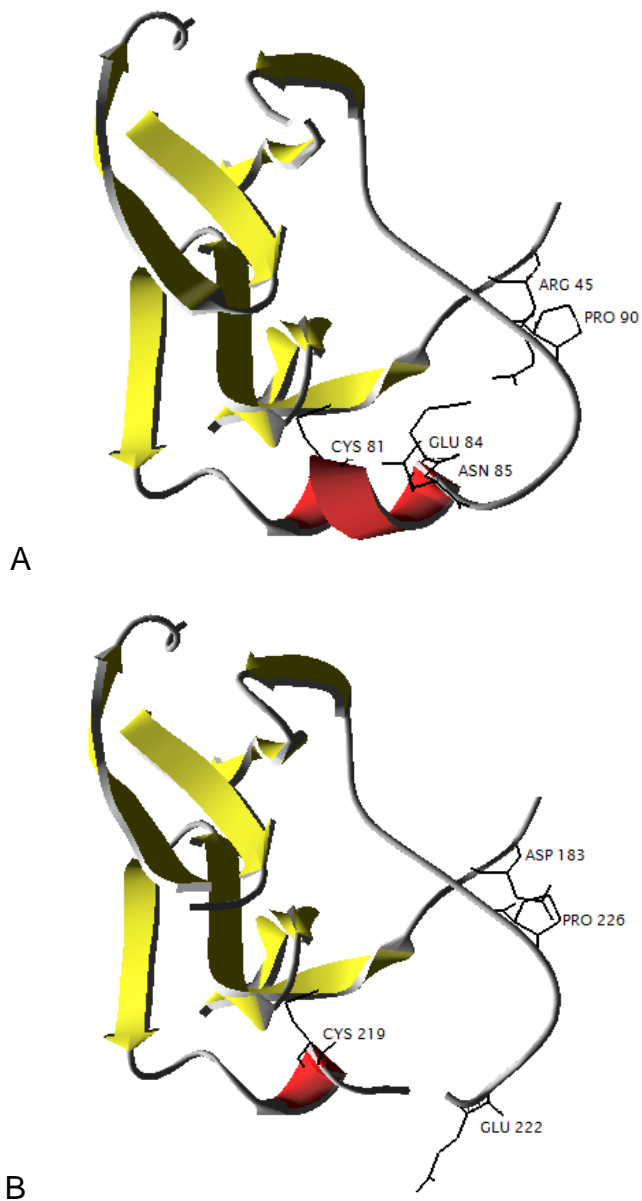


Figure 34. Homology models of echinotocin (A) and NGFFFamide (B) neurophysin molecules. The models were produced using DeepView and structural alignment to the crystal structure of peptide bound oxytocin-neurophysin. In the NGFFFamide neurophysin model there has been a deletion of two residues. The structures are highly conserved and residues structurally equivalent to those in the oxytocin neurophysin sequence may be important for peptide binding. Peptide-neurophysin interactions include a salt bridge between a glutamate carboxylate side-chain (GLU84 in A and GLU222 in B) and the alpha amino group of the peptide and an interaction with an arginine residue (conserved ARG45 in A, but substituted with an aspartic acid residue ASP183 in B). A hydrogen bond is located between the second tyrosine residue of the peptide (corresponding to the phenylalanine residue in echinotocin, and possibly one of the phenylalanine residues in NGFFFamide) and a neurophysin cysteine residue (CYS81 in A and CYS219 in B). Conserved residues between the seventh and eighth cysteine residues include; proline (PRO90 in A, PRO226 in B), asparagine (ASN85 in A, but missing in B).

Chapter 4

IDENTIFICATION OF PUTATIVE SEA URCHIN NEUROPEPTIDE PRECURSORS FROM A RADIAL NERVE EXPRESSED SEQUENCE TAG (EST) LIBRARY

- 4.1** Introduction: Identification of sea urchin neuropeptides using a cDNA library

- 4.2** Methods
 - 4.2.1** Identification of novel neuropeptide precursors

 - 4.2.2** Peptide detection using mass spectrometry

- 4.3** Results
 - 4.3.1** Results: Identification of putative sea urchin neuropeptide precursors

 - 4.3.2** Results: MS analysis of sea urchin body wall and visceral tissue

 - 4.3.3** Results: Sp-TRH and Sp-GnRH myoactivity

- 4.4** Discussion
 - 4.4.1** SpurS8 & SpurS9: two peptides encoded by a second sea urchin SALMFamide neuropeptide precursor

 - 4.4.2** Sea urchin gonadotropin-releasing hormone-like peptide

 - 4.4.3** Sea urchin thyrotropin-releasing hormone-like peptide

 - 4.4.4** Sea urchin calcitonin/CGRP-like peptide

 - 4.4.5** Two novel pedal peptide-like peptides

 - 4.4.6** GKamide: a novel sea urchin peptide

4.1 Introduction: Identification of sea urchin neuropeptides using a cDNA library

In the studies reported in this chapter a sea urchin neural expression EST library was used to identify echinoderm neuropeptide precursor genes and seven putative precursors were identified. The use of a smaller dataset representative of neural expression has circumvented some of the problems associated with searching larger genomic databases, such as an unmanageable set of sequence matches resulting from a short BLAST search query. Crucial to the analysis, however, was the use of both neuropeptide precursor structural criteria and homology searching against the NCBI protein database for annotated neuropeptide sequences from other phyla homologous to the putative peptides identified.

The characteristic criteria used in this study are generally applicable to neuropeptide precursors. Because the characteristics can be recognized from the full precursor amino acid sequence by itself, they lend themselves to bioinformatic or 'in silico' investigative techniques. Neuropeptide preproprotein sequences tend to be of short length, generally less than 500 residues (Liu et al., 2006) and encode repeated variable copies of the bioactive peptide. An N-terminal signal peptide sequence is essential for directing the translated preproprotein to the endoplasmic reticulum prior to processing by endopeptidases and structural modification of the peptide. Endopeptidases cleave bioactive peptides from the proprotein at monobasic and dibasic residue cleavage sites (e.g. Lys-Lys or Lys-Arg) found within the precursor. Lastly, many neuropeptides are amidated, which requires a

glycine residue at the C-terminus of the processed peptide (Bradbury et al., 1982).

Using all of these characteristics to manually search for neuropeptide precursors in an *S. purpuratus* neural Expressed Sequence Tag (EST) library, seven putative neuropeptide precursors have been identified and designated *spp-1* to *spp-7*. Four of the putative neuropeptides encoded by the precursors show clear homology to known neuropeptides in the animal kingdom, including SALMFamide neuropeptides, calcitonin/CGRP, gonadotropin-releasing hormone, and thyrotropin-releasing hormone. The three others may represent members of previously undescribed echinoderm neuropeptide families.

4.2 Methods

4.2.1 Identification of novel neuropeptide precursors

Analysis involved the separation and identification of putative neuropeptide precursors from a set of ESTs (Figure 35). ESTs are full or partial reads of copy DNA sequences derived from the mRNA of a particular cell or tissue. The ESTs sequences used in this study were derived from sea urchin adult radial nerve cord cDNA reads by Cameron et al. (2000) and consisted of 999 five prime reads and 1026 three prime reads, all approximately 1,000 nucleotides in length. Analysis was restricted to the set of 999 five prime EST because the signal peptide used to identify a potential precursor sequence is located five prime to the proprotein coding sequence. Five

prime ESTs are also more likely to contain coding nucleotide sequences as five prime untranslated regions (UTRs) are typically shorter than the three prime UTRs (Pesole et al., 1997). Peptide cleavage sites were identified by monobasic or dibasic motifs with a glycine residue as a substrate for C-terminal amidation, including GKR, GRR, GR, GRK, GKK, KR, KR, RR, GK, RK, KK, and R (Liu et al., 2006).

EST sequences with homology to larger proteins from the NCBI database were eliminated as an initial stage of the analysis. These included matches to proteins forming a set of 'hypothetical proteins' determined by the Gnomon gene prediction programme and used as part of the automated genome annotation process for *S. purpuratus* (Sodergren et al., 2006) (Figure 35). Therefore, Gnomon gene predictions of less than 500 residues represented within the EST set were included in the analysis. The online programme, Signal P (DTU, Denmark) (<http://www.cbs.dtu.dk/services/SignalP>) was used to identify likely signal peptide regions in the N-terminal region of each gene prediction.

The second stage of analysis involved grouping together the remaining 185 five prime ESTs sharing a coding sequence with a corresponding three prime EST. A second gene prediction programme, GenScan (MIT) (<http://genes.mit.edu/GENSCAN.html>) was used to identify Open Reading Frames (ORFs) within genomic DNA covering the two overlapping reads.

The third stage of the analysis involved BLAST searching the set of 302 remaining five prime ESTs against the set of official gene predictions

determined by the Glean3 gene prediction algorithm prior to searching the sequences for neuropeptide precursor characteristics. The online BLAST search facilities were accessed at the Baylor College of Medicine Sea Urchin Genome Project website (<http://www.hgsc.bcm.tmc.edu/projects/seaurchin/>).

4.2.2 Detection of the putative neuropeptides in sea urchin tissue using mass spectrometry

Like the synthetic echinotocin and NGFFFamide peptides used in Chapter Three, the peptides reported in this chapter were synthesized by the Advanced Biotechnology Centre at Imperial College London for the standardization of mass spectrometric analysis of sea urchin tissue extract. Mass spectrometric parameter settings were determined using the synthetic peptides and the same samples of *S. purpuratus* body wall and visceral tissue that were used for the detection of echinotocin and NGFFFamide were used as analyte for the detection of *in vivo* expression. Furthermore, the same procedures as described in Chapter Two were used to investigate myoactivity of the peptides on sea urchin oesophagus and tube foot tissues using pharmacological *in vitro* bioassays.

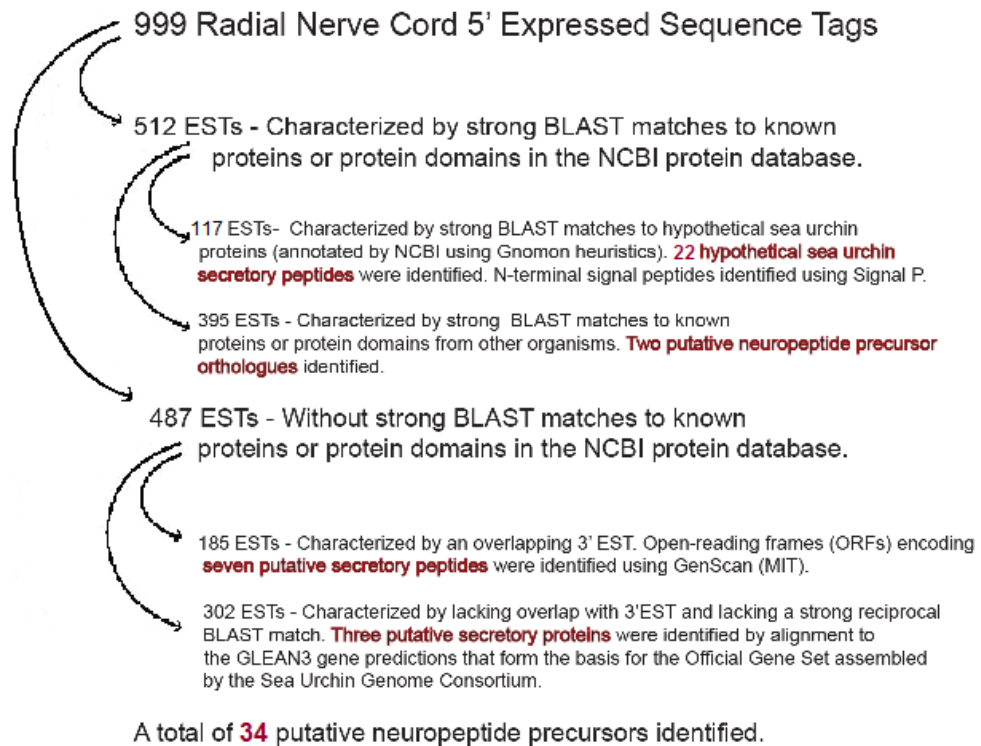


Figure 35. Identification of putative neuropeptide precursors from an adult sea urchin (*S. purpuratus*) radial nerve EST dataset. The search for neuropeptides in the EST library was restricted to 999 five prime reads containing coding regions. Thirty-four short putative genes encoding neurosecretory peptides were identified. The putative precursors were defined by an N-terminal signal peptide, the presence of monobasic and dibasic cleavage sites and a length of less than 500 amino acid residues.

4.3 Results

4.3.1 Results: Identification of putative sea urchin neuropeptide precursors

A set of 34 neural putative peptide-encoding cDNAs containing cleavage sites, an N-terminal signal peptide, and a length of less than 500 residues were identified (Appendix E). Of these, seven showed either homology to known neuropeptide or peptide sequences or to putative peptides encoded by the same precursor gene as repeated variable sequences (Figure 36). The epithet *spp* (*Strongylocentrotus purpuratus* peptide) was used to label each precursor gene. One Gnomon predicted ‘hypothetical protein’ (*spp-1*) (Figures 36 & 37) appears to encode two peptides with similar C-terminal amino acid sequences. Both of these were recognized as SALMFamide family neuropeptides and denominated SpurS8 (NMGSIHSHSGIHF-NH₂) and SpurS9 (MRLHPPGLLF-NH₂), thereby continuing the system of *S. purpuratus* SALMFamide peptide nomenclature begun by Elphick & Thorndyke (2005). Although the putative SpurS8 peptide has 13 residues, an equally possible peptide sequence could be truncated at the N-terminus, incorporating a further four residues including the suggested dibasic cleavage site (GGKKNMIGSIHSHSGIHF-NH₂).

Precursor *spp-2* was identified by a protein BLAST search of the NCBI database as a putative homologue to gonadotropin-releasing hormone (GnRH) (Figures 36 & 38). The precursor’s closest GnRH-like alignment

was with *Oryzias latipes* preprogonadoliberin II (GI:34098706). Similar to other GnRH-like peptides, the putative sea urchin GnRH-like peptide identified has a pyroglutamate residue at the N-terminus. The peptide (pyroGlu-Val-His-His-Arg-Phe-Ser-Gly-Trp-Arg-Pro-Gly-NH₂) was denominated Sp-GnRH. The precursor is also represented by one read from a set of larval ESTs (GI:34745970).

Precursor *spp-3* contains seventeen repeats of a putative thyrotropin-releasing hormone (TRH)-like tetrapeptide (Glu-Tyr-Pro-Gly-NH₂) and was denominated Sp-TRH (Figures 36 & 39). Like vertebrate TRH all peptide repeats are encoded by the third exon of a three exon precursor. The gene is also represented by one read from a set of mesenchyme ESTs. The Glean3 prediction for this gene (Glean3_08352; SPU_008352) has a different exon structure from the Gnomon prediction and one that lacks the coding region for an N-terminal signal peptide. The EST data, however, confirms the structure of the former Gnomon prediction.

Precursor *spp-4* contains a putative peptide sequence displaying homology to calcitonin/CGRP and has been denominated Sp-CGRP (Figures 36 & 40). *Spp-4* is also represented by two reads from a set of larval ESTs.

Some of the sequences lacked homology to known neuropeptides, yet displayed characteristics typical of neuropeptide precursors. Precursor *spp-5*, automatically annotated by NCBI as arginine/serine-rich splicing factor 4, resembles a novel neuropeptide precursor containing thirteen similar repeats of a putative undecapeptide (Ala-Asn-Met-Phe-Arg-Ser-Arg-Leu-

Arg-Gly-Lys-NH₂) and was denominated Sp-GKamide (Figures 36 & 41).

Precursor *spp-6* weakly aligns with molluscan pedal peptide 2

(GI:94434888) (*Aplysia californica*). Although the precursor contains 22

variable repetitions of a non-amidated putative peptide, the gene deviates

from the structural criteria in being 510 residues long. The peptide

consensus sequence (Gly-Phe-Asn-Ser-Gly-Ala-Met-Glu-Pro-Leu-Gly-Ser-

Gly-Phe-Ile) was denominated Sp-PPLN1 (Figures 36 & 42). The gene is

also expressed in ESTs derived from the blastula stage of sea urchin

development, the adult lantern and the primary mesenchyme.

Lastly, precursor *spp-7* displays homology to the pedal peptide-like *spp-6*

gene but without a Gnomon gene prediction corresponding to the EST

(Figure 36). The Glean3_24381 (SPU_024381) was identified by

submitting the EST as a BLAST search query against the Glean3 set of

gene predictions. Interestingly, the exon encoding the signal peptide

sequence and the exon encoding the peptide are located on different

scaffolds. The gene is also represented by one read from a set of larval

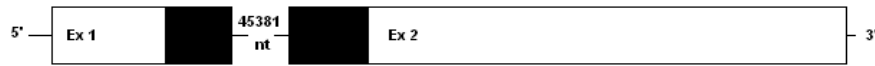
ESTs. The larval EST contains coding sequences from both exons,

confirming that expression of both sequences is combined in a single

mRNA strand.

A) *Spp-1* (GI:115711933; GI:115946155) (encoding SpurS8 and SpurS9).

MQVQOITVFLVACTLSVLVVAYAEQDAETVLLNRLRDLAARAAAGELPDFFADVDDYKRG.G**KK**NMGSIHSHSGI**H**
FGKRRDSESSERARNT**KR**MRLHPPGL**LFGR**APVQKWDQWQAQDTYNPDWELGQFN



B) *Spp-2* (GI:72011734; GI:115975296; SPU_019680) (encoding Sp-GnRH)

MKQIITSLVSI SAALLLFVLI SEYTPRCNGQVHHRFSGWRPGG**KR**RSDDAEVNSNKITIERPQLPICQTTEERQL.
 LEGSDSILGDLRRAANRMLLQLFNLSKTRLN.DLNDATSNEVDERPVYGDYLGTLG



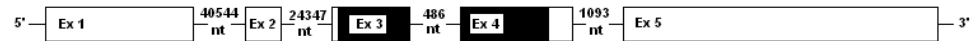
C) *Spp-3* (GI:115657896; SPU_008352) (encoding Sp-TRH)

MWACILGYVTWGGAALPTILGKELVLSNDGP.EIADWVQKEIPLRNQYWGDAEVEEELGMLSPDSE**KRQYP**
GGKRQYPGGKRQYPGGKRQYPGGKRQFPAGKRQFVGGELIPSELRQWPGGKRQWPGGKRQYPGGKRQY
PGGKRQWPVEKRQYPGGKRSEDDQDLLPMEIR**QYPGGKRQWPGGKRQYPGGKRQFPGGKRQFVGG**EALE
 QESNINKRFAPEDDTMDFRRLSQLYDTNDNIVADEGELALEDLDDIMVDRPEFEDPRDLLLGNVDQEDVLALDL
 SALLGDRNPNGW*



D) *Spp-4* (GI:115767208; GI:115945640) (encoding Sp-CGRP)

MKSTVIVTLTICCLLYQTTRAASLTN.RDGLSRQDILDLLQLYEPIRQEGGD**KK**SKGCGSFSGCMQMEVAKNRVA
ALLRNSNAHLFGLNGPGKRRRSVDDLQVNDAAETE



E) *Spp-5* (GI:115933326; GI:115963862; SPU_018666) (encoding Sp-GKamide peptides)

MSRNAYLWAGLLGALCLLITTSIKADGEVTEVDV**KR**ANYFRGR**KPGKR**DEPDAAALVPDDDLSE**KRANMFRS**
RLRGKGRDDPDAAMLPGDWDEE**KRANMFRSRLRGNGKR**DDPDAAMLPGDWDEE**KRANMFRSRLRGKGR**DEPDAA
 EALVPGDWEE**KRANMFRSRLRGKGR**DDPDAAEALVPGDDLSE**KRANMFRSRLRGKGR**DDPDAAEALVPGDDL
 SEE**KRANMFRSRLRFKGR**DDPDAAEALVPGDDLSE**KRANMFRSRLRGKGR**DDPDAAEALVPGDDLSE**KRANM**
FRSRLRGKGRDDPDAAEALVPGDWDEE**KRANMFRSRLRGKGR**DDPDAALVGGDFGDFVDEE**KRANMFRSRLRG**
NGKRDDPDAALVDFEMDEE**KR**ANYFRGR**RPPGKR**DEPDAAALVEDE**KRANFRARQRPKLGK***



F) *Spp-6* (GI:72008820; SPU_03108) (encoding Sp-PPLN1 peptides)

MKFSNGRGAFLVNLIFVLCLVDHMAECRPARKTRDVDE.DLEKEEDSLINALEKVLADEVIDNAENDSDDETG
 ITDRELSLMSMLRDDVSPSRLRGYFGGWRPAYPSELSHVGALEPLATGFLPSRYSG**KKR**FLTGAL**PLSSGF**
IKKGFNTG.AMEPLGSGFIKKGFNSGAMEPLGAGFFKKGFNSGAMEPLGAGFFKKGFNSGAMEPLGAGFFKKGFNSG
AMEPLGAGFFKKGFNSGAMEPLGAGFFKKGFNSGAMEPLGAGFFKKGFNSGAMEPLGAGFFKKGFNSGAMEPLGSGF
IKKGFNSGAMEPLGSGFIKKGFNNGAMEPLGSGFIKKGFNSGAMEPLGSGFIKKGFNSGAMEPLGSGFIKKGFNTG
AMEPLGSGFIKKGFNSGAMEPLGSGFIKKDFNTGAMEPLGSGFIKKGFNSGAMEPLGAGFFKKGFHAGAMEPLSSG
FIDGKRGFYNGAMEPLSAGFHQGRGFHEGEMDKKGFHNGAMEPLKSGFLKD*



g) *Spp-7* (SPU_024381) (encoding Sp-PPLN2 peptides)

MNNYAFLECLACAIQVWTLPIEDK.DGLDIEDQEEAE**KR**FGSMNMEPLVSGFY**KR**FGSGLDSMQSGFY**KR**NFGSG
LNMPEMQSGFYKRNFGGSMPEMQSGFYKRFGGAMEPMSGSGFYKRFSGSLEPMSGSGFYKRNFGSLEPMSGSGFYKRN
FGGANEPMRSGFFKRFSGSLEPMSGSGFYKRNFGSGLDAMQSGFYKRSQEEETD*



Figure 36. Putative neuropeptide precursor sequences (*spp-1* to *spp-7*; A-G). Peptide sequences are shown in blue and cleavage sites in green. Full stops in the amino acid sequence indicate the position of intron splice sites in the coding sequence. Exon and intron structure are shown beneath each precursor sequence. Black infilling denotes a coding region within an exon.

```

1    agaagcgcacagacacacacatagcccccaggaggctgctacttcatTTTTGGTTgaatc
61   aaattcaagagaattatcgagatagggcgggaaagcaaacggccaaacaaaactgtcactt
121  catttataaaaaaacacatactgtgcgcacttggtagtttcaagtaatcacacgaaatcg
181  aacaagatcatacttgaaggcaagctgctctaggatccctcccttccctaaacaagtgctc
241  ttaagactttgaagaaagtagaaggggactcgatttgatTTTTGTataaaaaatcccaatc
301  cteccaaatgcaggtaacaacagataaactgtatTTTTtagtagcatgcacgctgagcgtgtt
      M Q V Q Q I T V F L V A C T L S V L
361  agtagttgcatatgccccagaggacgcagaaacagtacttctcaacagactcagagatat
      V V A Y A Q E D A E T V L L N R L R D I
421  tgccgcccaggcagcgggcgagggtaccggactTTTTcgtgatgtagacgattataa
      A A R A A A A G E L P D F F A D V D D Y K
481  gagagggggaaaaaagaacatgggcagcattcactccactccggcatccacttgggcaa
      R G G K K N M G S I H S H S G I H F G K
541  gaggcgagactctgagagcagtgagcgggcaagggaataccaagagaatgagattacccc
      R R D S E S S E R A R N T K R M R L H P
601  aggcctactTTTTggcaagcgagcccccgcccccaaaaatgggacccaatggcagcagga
      P G L L F G K R A P V Q K W D Q W Q A Q
661  tacatataccccgattgggaacttggacagttcaattaatcatcaatcacataaaaaaat
      D T Y N P D W E L G Q F N
721  aagttagaagatacaggcaacaggtcaacttgagttTTTTgttctgatttattagaaggt
781  ccattgcaatcgtTTTTgtttgtaaaaaatgtctcattctctcttcttcttattttgggtt
841  ttgtaaggacagatatcgatatattggattatattcttctgtagtttcaaaaagggaatgat
901  ggatgttctgattatagattatctcattgcatttcttcttcttcttcttcttcttcttct
961  tctacgaaaaagatcaaatatgatattgatgaattgctactagaatatataatttgaaactg
1021  ggtaattattagaatgacgggcaaggattacattagatgcagaagaatatgaaactgagtg
1081  ataggctcaccggcaagcagcttttatcttgtttcttcttcttcttcttcttcttcttctt
1141  attctgattgtgaaagtctcctacattatatacaatcgtgatctcttcttcttcttcttctt
1201  cgggtacatcgtgtgggttcttcttgaacggaggatctctgctcacttttggcgttctcttag
1261  catcaaaattatcttctgaattatcaatgttgggaatgtgactgccaatcttctcattacc
1321  tctttcaaaaatcagatatagttctataagaagacacgacatgttctctgaatgtgttaagg
1381  aattaagttgcaacgctctgtacacccggtttttctccttcaaaacatatcttctcactctc
1441  caaaacccctttgaaagaatagcccgatgtaagggcagttcgcaatcatctgatgacaaaa
1501  cccgaatttacttttcacccagttcaaaaatccgtgagtgacctgactgtagtggtggacat
1561  ccatacgaactagaagtgattatctgtgctgttaacagtcgccgaaagagcgttccgaag
1621  gcagaaaaaaatgaaagaaaccagttcttctcaattctctgatgtggagatggagtaaacctg
1681  agtttaataaaaaatcgccaccaagcttcatcaagcgatcataaacgctctatcaagcgtg
      tcaaaatcccaattaccctatacagatccattagcaaaaattgatactgagacactttgaggtg
      ttttacctcgatatctcttagtcaattccattaaaaagcccaaggtgtttcaagaagaggggtg
      tcgaaagtacgtatgttttggcctgtattatgcacctctctgtatgatgggttcggttagaag
      gctgtccacccccatgaaaatccacctgctttatagctgtcaatgtaaacactgtccatc
      gctgaaattcttacctgagtg

```

Figure 37. *spp-1*; the putative SpurS8 (NMGSIHSHSGIHF-NH₂) and SpurS9 (MRLHPPGLLF-NH₂) neuropeptide precursor nucleotide and amino acid sequence in full. The signal peptide sequence is shown in blue, the peptide sequence in dark red, and the basic residue cleavage sites in green. An alternative SpurS8 peptide sequence might incorporate a further four residues N-terminal to the predicted sequence (GGKKNMGSIHSHSGIHF-NH₂). Introns were identified by an intron N-terminal GT sequence and C-terminal AG sequence and the intron splice sites are underlined in the exon nucleotide sequence.


```

1      caagcagtcagtccttagtgcatttcgtacaggtggttagaacgtggacagacagagaagggg
61     gaaaacaaaacgaggggaaatcaaacacagggaaaaaacagagaaggcgcctcaaaagtgaaga
121    ggtaccocgaaagtcatcccgctgccggatttcgtcgactttaaggtaaaaactcgtcttatt
181    tgctcgccagaaagaaagagaatcatagatgagcggactaggagaaactaacgcagatcttta
241    cgtgtgggacttagcaagaggacttagcaagagggccggcctggtcctgaagggaactgt
301    gggattatataatgtgggcatgcattctgggatacgtccacatgggggtggagctgcactacc
        M W A C I L G Y V T W G G A A L P
361    aaaccatcctgggaaaaagaactcgtccttagtgagaacgatggggccagagatagctgactg
        T I L G K E L V L S E N D G P E I A D W
421    ggtgcaaggaaaggaatccctttgagggaatcaatattggggcgatgtagcggggaaga
        V Q G K E I P L R N Q Y W G D V A E E E
481    ggaggaagaagaacttggcatgttgcctccctgattcagagaaaacgccagtaccaggtgg
        E E E E L G M L S P D S E K R Q Y P G G
541    taaaagacagtatccaggaggaagcgtcagtatccaggagggaaaacgtcaatcccccg
        K R Q Y P G G K R Q Y P G G K R Q Y P G
601    cggcaagcgtcagtttccggctggcaaaaagcagttcgtcggaggtgaacttataccatc
        G K R Q F P A G K R Q F V G G E L I P S
661    gcctgagctacgtcaatggcccgggtggaagcgtcaatggcccgggtggaagcgtcaatg
        P E L R Q W P G G K R Q W P G G K R W
721    gcccccgtggttaagcgtcaatatacccccgtggaagcgtcaatataccagggtggttaagc
        P G G K R Q Y P G G K R Q Y P G G K R Q
781    atggcccgaagttaagaggcaataccctgggtggcaaacggcggaaagcagatcaagatct
        W P E V K R Q Y P G G K R S E D D Q D L
841    cctgcccatagaaatccgacaataccctggcggtaaacgcccaatggcccgggtggttaacg
        L P M E I R Q Y P G G K R Q W P G G K R
901    acagtatccccggagggaagaggcaataccctggagggaacggcaatctcctgggtggaaa
        Q Y P G G K R Q Y P G G K R Q F P G G K
961    gggcagttcgtcgtggggaagcactcgaacaagaatcgaacatcaacaagcagattcgc
        R Q F V G G E A L E Q E S N I N K R F A
1021   ccctgaagacgacaccatggacttcttttaggctatcccagttatgatgacacgaacgataa
        P E D D T M D F F R L S Q L Y D T N D N
1081   tatagtcgcagatgaaggagaattagctctagaggatctcctgggatgacatcatggttga
        I V A D E G E L A L E D L L D D I M V D
1141   caagaggccagatgttgaagacccaagggaacttgttactagggcaatggtgaccaagaaga
        T R P E F E D P R D L L L G N V D Q E D
1201   tgttcttgccttagatatttcggcctttacttggagatagaacccaacacgggtggta
        V L A L D L S A L L G D R N P N N G W -
1261   a

```

Figure 39. *spp-3*; the putative Sp-TRH (pEYPG-NH₂) neuropeptide precursor nucleotide and amino acid sequence in full. The signal peptide sequence is shown in blue, the peptide sequence in dark red, and the basic residue cleavage sites in green. Introns were identified by an intron N-terminal GT sequence and C-terminal AG sequence and the intron splice sites are underlined in the exon nucleotide sequence.

```

1      aaaaaccgggaaaaagcgaataactcagccctttttgtgaagcgtttgtgaatgagacaacc
61     ggggctgccaaacgaaagagcggggcgaagtatttgggtaagaaacagttattgggaagaa
121    aaacaggataaagaagaaaaacgcaaacgtcattggaatctaatagaaacgacccctccggt
181    gggtaagagcagggtcagttgaaaggaattttgcgagaaagctgtcttgaatctggg
241    gagaacagacgatttctaaataacatccactccccaaaaacagaagcgcctaaagacacag
301    acagccggaaacgaaatccacatcccgaagaggacgaccagaaataatcaaaagacaaagc
361    ttccgcggaaccccaataaaactttataccatcaccatgaaaaagcactgtgattgtaacgctt
                                     M K S T V I V T L
421    acgatatgctgtctttttataccaaaaaactagggcggcgcctcgttaaccaacagagcggg
      T I C C L L Y Q T T R A A S L T N R D G
481    ctctcaagacaagacatcctagatcttctcagctctatgaagaaccaataagacaagag
      L S R Q D I L D L L Q L Y E E P I R Q E
541    ggtggcgacaagagggtcaaaagggttgcgggagttttagtgggtgtatgcaaatgggaagt
      G G D K R S K G C G S F S G C M Q M E V
601    gcaaaagaaacagtggtgctcttttaagaacagtaacgcccatcttttggctaaat
      A K N R V A A L L R N S N A H L F G L N
661    ggaccggcaagagacgacgatcgggtcagcagatcttccccagggtcaacgacgcagagacg
      G P G K R R R S V D D L P Q V N D A E T
721    gagttagattcaacagcgcacatctagagttcaccacacctgtttttgtcctcactcagc
      E
781    cttctctctctctctatttccctgaccacctatgccccccccccatctctctctctctct
841    cctttttttctctctctctctctctctctctctctctctcaaggctctgtcagagatccccaa
901    taggggagtcccccactgtgatcatctcgcacaagcatatctttgtccagctgacaatatcg
961    gactcgggcccctaaccaactcgggtacaaatgatgattgcctctcttaggactcagtcacgga
1021  tagaaaaattaaagagattaatagttttgtgttatcaataatgatggagggtctttagaag
1081  taaaaacctatgaggttaatatatgcatttgcaatgaaatgtatttaaaaacaattttaaa
1141  gttcattttgaactttttgttaacgatttgaagtgagaatcattatgggattcgtttaa
1201  gagcaagacctttgataagtgtaaaaaataatgaaaactaaaaatgggggtaaaactgaaagc
1261  cgctaggatggatatcccaactatgagccaccccttaaaaaacaaaataatcgtttaaatttca
1341  ccgaattcagaccctttttacatttaattgtca

```

Figure 40. *spp-4*; the putative Sp-calcitonin/CGRP-like (SKGCGSFSGCMQMEVAKNRVAALLRNSNAHLFGLNGP-NH₂) neuropeptide precursor nucleotide and amino acid sequence in full. The signal peptide sequence is shown in blue, the peptide sequence in dark red, and the basic residue cleavage sites in green. Introns were identified by an intron N-terminal GT sequence and C-terminal AG sequence and the intron splice sites are underlined in the exon nucleotide sequence.

```

1      ggcagtggaactcgggtgtgcacatcatcatcgatctttcctaccttcgggacccctggacaacgc
61     ggattgaaatcgcgatcctcgccttctctatctctttgtcgtctgcttcgtatttcatac
121    ctggacctgagagactggaatcaaaccttctgaaggaccctcctcgcacagtagctgaagc
181    aagtctgtgacgcacgatagaagtagacaggagctacgatgtcacgaaacgcacatctctat
      M S R N A Y L
241    gggcaggtctgctcctaggagctctatgtctccttatcaccacaaccagtatcaaaagctg
W A G L L L G A L C L L I T T T S I K A
301    acggtgaggttacagaagatgtcgacaaaacgggcaaaactacttccggggagcggggagaa
D G E V T E D V D K R A N Y F R G R G R
361    aaactggaagcgcagacgagcctgacgcagccctggtaaccgatgatgacttgagtgagg
K P G K R D E P D A A L V P D D D L S E
421    acaagagagcgaatattgtttagaagtcggctccgcggtaaaaggaaagcgcagacgcacctg
D K R A N M F R S R L R G K G K R D D P
481    atgcagccatgttaacctgggtgactgggatgagggaaaagagagcgaacatgtttagaagtc
D A A M L P G D W D E E K R A N M F R S
641    ggctccgcggtaacggaaagcgcagacgatcccgatgcagccatgttaacctgggtgactggg
R L R G N G K R D D P D A A M L P G D W
701    atgagggaaaagagagcgaacatgtttagaagtcggctccgcggtaaaaggaaaacgcagacg
D E E K R A N M F R S R L R G K G K R D
761    aaactgacgcagctgaagccctggtaacctgggtgactgggaaagggaaaagagagcgaaca
E P D A A E A L V P G D W E E E K R A N
821    tgtttagaagccggctccgcggtaaaaggaaaagcgcagacgatccctgacgcagccgaagccc
M F R S R L R G K G K R D D P D A A E A
881    tggtaacctgggtgacttaagtgaggaaaagagagcgaacatgtttagaagtcggctcc
L V P G D D L S E E K R A N M F R S R L
941    gcggtaaaaggaaaagcgcagacgatcccgacgcagccgaagccctggtaacctgggtgact
R G K G K R D D P D A A E A L V P G D D
1001  taagtgagggaaaagagagcgaacatgtttagaagccggctccgcggtaaaaggaaaagcgcag
L S E E K R A N M F R S R L R F K G K R
1061  acgatcccgacgcagccgaagccctggtaacctgggtgacttaagtgagggaaaagagag
D D P D A A E A L V P G G D L S E E K R
1021  cgaatattgtttagaagccggctccgcggtaaaaggaaaagcgcagacgatccctgacgcagccg
A N M F R S R L R G K G K R D D P D A A
1081  aagccctggtaacctgggtgacttaagtgagggaaaagagagcgaacatgtttagaagtc
E A L V P G G D L S E E K R A N M F R S
1141  ggctccgcggtaaaaggaaaagcgcagacgatcccgacgcagccgaagccctggtaacctgggtg
R L R G K G K R D D P D A A E A L V P G
1201  actgggatgagggaaaagagagcgaacatgtttagaagtcggctccgcggtaaaaggaaaagc
D W D E E K R A N M F R S R L R G K G K
1261  gagacgatcccgacgcagccctggtaggtgacgatttcgggtgatgaatttcgtggatgaag
R D D P D A A L V G D D F G D E F V D E
1321  aaaaagagagcgaacatgtttagaagtcggctccgcggtaacggaaaagcgcagacgatcccg
E K R A N M F R S R L R G N G K R D D P
1381  atgcagccctggtagatgaattcatggatgagggaaaagagagcgaactacttcagaggac
D A A L V D E F M D E E K R A N Y F R G
1441  ggggaagaagacctggaaaagcagatgagcccgacgcagccctggtagaagatgaaaaaa
R G R R P G K R D E P D A A L V E D E K
1501  gggccaacttcagagccccgcagcccccaattagggaaagtaa
R A N F R A R Q R P K L G K +

```

Figure 41. *spp-5*; the putative Sp-GKamide (consensus sequence ANMFRSRLRGK-NH₂) neuropeptide precursor nucleotide and amino acid sequence in full. The signal peptide sequence is shown in blue, the peptide sequence in dark red, and the basic residue cleavage sites in green. Introns were identified by an intron N-terminal GT sequence and C-terminal AG sequence and the intron splice sites are underlined in the exon nucleotide sequence.

```

1   gatacaatcaatctatctccaaatttgagatacggagcggagatattctcagtatctctca
61  totagaacgaacgtgcaaacgataaactgatacagaattattctcctcogctattctcttacc
121  ctattaccctccogtccaaacogctgctctctggcagagatcccaactcggaaaggtgctct
181  ggtacaaaataactcaaatcaagaagagcaaccaatataatctctcattagttcogctctcc
241  tccatcttaagaattcctctctctctctctctctctctctctctctctctctctctctct
301  gggcaatttaacacogcgttagacacogcgtcogagctctctcggtaaacgggacagggcgc
361  gagctgggtattcctctcagttgtcgaatcgcogcgtatacaaaagtgaagttctctcgg
      M K F S G
421  aaatggcogcagggagcttctctcgtcgtcaaacctgatctctcgtctctctctctctctct
      N G R G A F L V V N L I F V L C L V D H
481  tatggcgggaatgtaggcogcggagaaaagacagagagacogcggatgaggatctcgaaaaaa
      M A E C R P A R K T R D V D E D L E K E
541  agaagacagccttatacaacgctctcggagaaagtctctcagcagatgaagaggtgattgcaaa
      E D S L I N A L E K V L A D E E V I D N
601  cgcagagaaacgattccgacagcaaaacagggataaacagatcgagaaactgtctctctatgct
      A E N D S D D E T G I T D R E L S L M L
661  ttcacatggtggcgtgatgatttagccccagccgctccgagggtaacttggcgggtaaatg
      S M L R D D V S P S R L R G Y F G G K W
721  gagaccagcgtactaccctccagcgtcattcaactggttggtgcttggaaacctcttgctac
      R P A Y Y P S E S L H V G A L E P L A T
781  gggatttttaacaaagtctcactcagccagcaaaaagaggttctctcagcgggtgcttggga
      G F L P S R Y S G Q K K R F L T G A L E
841  acccttagcagtggttctcaagaagaggttcaatcaaggtgcaactggaaacacacttgg
      P L S S G F I K K G F N T G A M E P L G
701  etccagggctcacaagaagggattcaacagtgaggccatggaaaccactggagcgggatt
      S G F I K K G F N S G A M E P L G A G F
761  cttaagaagaggattcaacagtgggccatggaaaccactcggagccggattctcaaga
      F K K G C F N S G A M E P L G A G F F K K
821  gggattcaacagtgggccatggaaaccactcggagcgggattctcaagaagggattcaaa
      G F N S G A M E P L G A G F F K K G F N
881  cagtgggccatggaaaccactggagcgggattctcaagaaggggttcaacagtggggc
      S G A M E P L G A G F F K K G F N S G A
941  catggaaaccactggagcgggattctcaagaagggattcaatagtgggccatggaaac
      M E P L G A G F F K K G F N S G A M E P
1001  acttggagcgggattttcaagaagggattcaacagtgggccatggaaaccactcggagc
      L G A C F F K K G F N S G A M E P L G A
1061  cggattctcaagaaggggttcaacagtggtgcaatggaaaccactcggatcaggggtctat
      C F F K K G F N S G A M E P L G S G F I
1121  caagaaggggttcaacagtggtgcaatggaaaccactcggatcagggattcaatgaagaag
      K K G F N S G A M E P L G S G F I K K G
1181  gttcaacaatgggtgcaatggaaaccactcaggatcaggatttatcaagaaggggttcaaac
      F N N G A M E P L G S G F I K K G F N S
1141  tggtgcaatggaaacctctaggtatcaggattcaatcaagaaggggttcaacagtgggccat
      G A M E P L G S G F I K K G F N S G A M
1201  ggaaccaactcggatcagggattcaatcaagaaggggttcaaccaactggtgcaatggaaac
      E P L G S G F I K K G F N T G A M E P L
1261  aggatcagggattcaatcaagaaggggttcaacagtggtgcaatggaaaccaactcggatcag
      G S G F I K K G F N S G A M E P L G S G
1321  atctcaagaagatttcaaccaactggtgcaatggaaacctcaggatcagggattcaatcaaa
      F I K K D F N T G A M E P L G S G F I K
1381  gaagggcttcaacagtgaggccatggaaaccaactcggagcaggattctcaagaaggggtt
      K G F N S G A M E P L G A G F F K K G C F
1441  ccaacogcgggtgcaatggaaaccaactctctctctctctctctctctctctctctctctct
      H A G A M E P L S S G F I D G K R G F Y
1501  caacogcgtgcaatggagccttcaagtgacggattcaaccaaggaagagaggggttcaacga
      N G A M E P L S A G F H Q G K R G F H E
1561  aggtgcaatggaaacagcaagaaggggttcaataggtgctatggagccgctcaaatc
      G E M D K D K K G F H N G A M E P L K S
1621  aggtttctcaagaagctag
      G F L K D *

```

Figure 42. *spp-6*; the putative Sp-PPLN1 (consensus sequence GFNSGAMEPLGSGFI) neuropeptide precursor nucleotide and amino acid sequence in full. The signal peptide sequence is shown in blue, the peptide sequence in dark red, and the basic residue cleavage sites in green. Introns were identified by an intron N-terminal GT sequence and C-terminal AG sequence and the intron splice sites are underlined in the exon nucleotide sequence.

4.3.2 Results: MS analysis of sea urchin body wall and visceral tissue

Peptide sequences were selected from the identified putative neuropeptide genes including Sp-GnRH, Sp-TRH, as well as peptides representing consensus sequences from Sp-GKamide and Sp-PPLN-1 precursors. All of the peptides were synthesized with an amidated C-terminal with the exception of PPLN-1. In accordance with the structure of human TRH and GnRH, both Sp-GnRH and Sp-TRH were constructed with an N-terminal pyroglutamate residue. Where multiple forms of a peptide were encoded by a precursor gene a representative consensus sequence was chosen. The following sequences, including molecular masses in brackets, were synthesized.

Sp-GnRH (1444.6019)	pEVHHRFSGWRPG-NH ₂
Sp-TRH (444.4675)	pEYPG-NH ₂
GKamide (1334.6151)	ANMFRSRLRGK-NH ₂
PPLN-1 (1483.6653)	GFNSGAMEPLGSGFI

Detection of each of the synthetic peptides using HPLC-MC was standardized, using a 20 µL injection of a 10 µM peptide solution and using several different gradients of organic solvent. The peptides were also subjected to the same extraction protocol that proved effective for detection of the SALMFamide neuropeptides from *Asterias rubens* extracts, and for NGFFFamide from *Strongylocentrotus purpuratus* extracts.

A summary of detection standardization can be found in Table 8.

The synthetic peptide masses identified by mass spectrometry, in protonated form, are compared to the calculated masses in Table 9. As for the detection of NGFFFamide and echinotocin, expert parameters were optimized for each peptide by injecting the mass spectrometer directly with a 1 μM peptide solution, bypassing separation on the HPLC column (Appendix D). Thus, the final method was set up with different voltage settings across an organic gradient corresponding to the point at which the synthetic peptides were found to elute.

Peaks were identified corresponding to Sp-TRH, PPLN-1 (peak average intensity of 2.2×10^5) (Figure 43) and a smaller peak corresponding to Sp-GnRH (at 18.4-19m). The peptide with a mass corresponding to Sp-TRH was detected in the body wall extract at an average intensity of 1.25×10^5 and 1.5×10^5 , but not in the visceral extract (Figure 44).

Table 8: Standardization of synthetic peptide elution points on an acetonitrile (ACN) gradient, including elution points for echinotocin and NGFFFamide.

Peptide	Elution Point (% ACN)			
	0.5%/min	1%/min	1.67%/min	2%/min
GKamide	-	10-40	-	25-35
NGFFFamide	39-46	44-46	-	44-47
Echinotocin	-	32-39	37-39	~42
Sp-TRH	41-46	16-19	13-19	10-25
PPLN1	41.4-45	-	42-44	-
Sp-GnRH	47-60	-	-	-

Table 9: Predicted and observed synthetic peptide masses of Sp-GnRH, Sp-TRH, GKamide, PPLN-1 detected using HPLC-MS. The difference observed between observed and calculated mass gave an error range of 0.082%. Sp-TRH was also identified at high intensity as a sodium-adduct dimer.

Peptide Peptide	Calculated Mass	Observed Mass(es) of Synthetic
Sp-GnRH	1445.6019 (H ⁺)	1445.6 (H ⁺)
Sp-TRH	445.4675 (H ⁺)	446.22 (H ⁺) 913.34 (dimer H ⁺ Na ⁺)
GKamide	1335.6151 (H ⁺)	1334.7 (H ⁺) 667.85 (2H ⁺) 445.57 (3H ⁺)
PPLN-1	1484.6653 (H ⁺)	1483 (H ⁺)

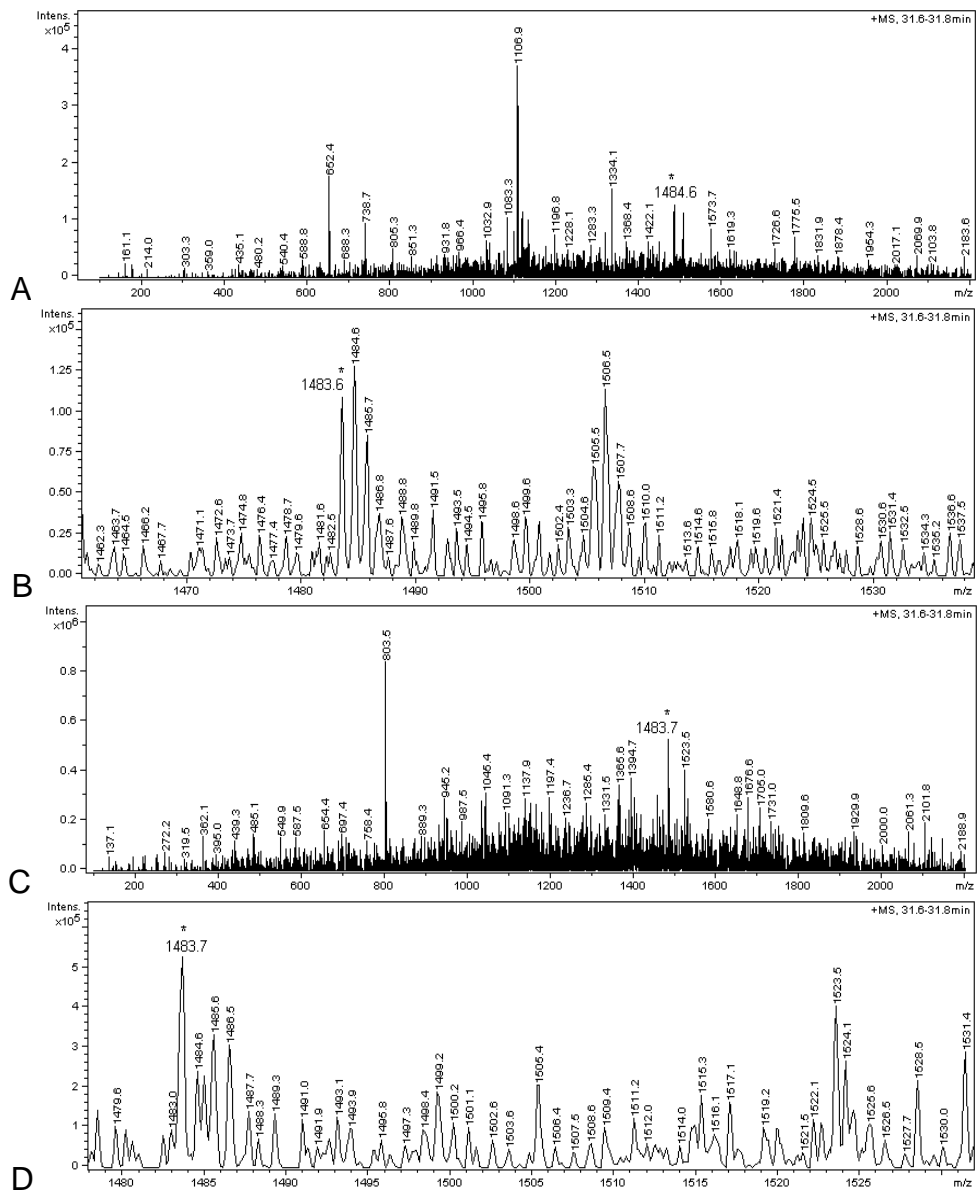


Figure 43. Detection of PPLN-1 consensus sequence peptide using HPLC-MS.

A) Full mass spectrum indicating detection of PPLN-1 consensus sequence synthetic peptide (1484.6) (*).

B) Magnified mass spectrum view of 37A. PPLN-1 consensus sequence synthetic peptide (1483.6) detected at 31.6 to 31.8 minutes (*).

C) Full mass spectrum indicating detection of PPLN-1 peptide (1483.7) (*). Spectrum obtained from sea urchin *S. purpuratus* body wall extract.

D) Magnified mass spectrum view of 37C. Spectrum obtained from sea urchin *S. purpuratus* body wall extract. Putative PPLN-1 peptide (1483.7) detected from 31.6 to 31.8 minutes (*).

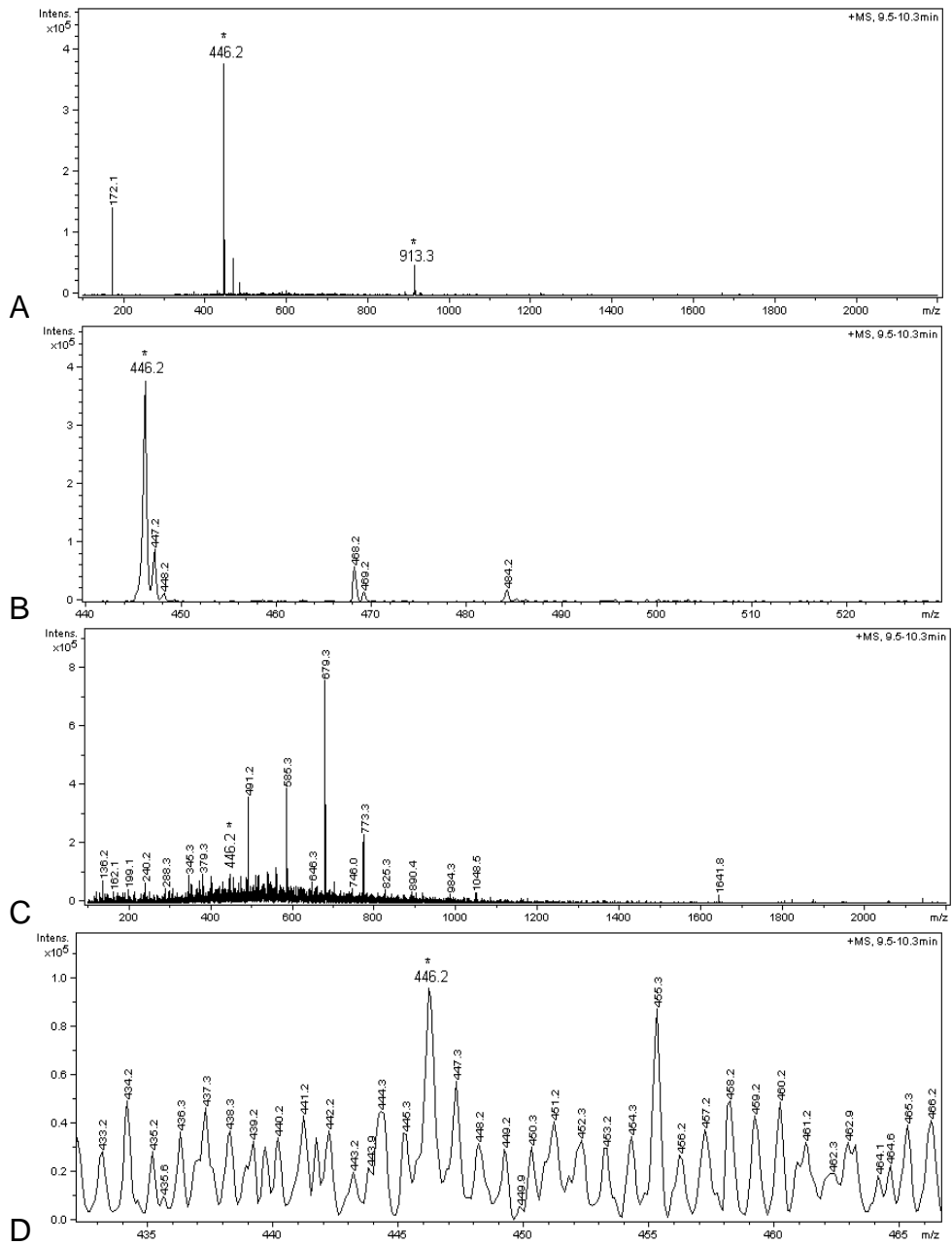


Figure 44. Detection of Sp-TRH using HPLC-MS.

- A) Full mass spectrum indicating detection of Sp-TRH peptide (446.2) detected at 9.5 to 10.3 minutes (*).
- B) Magnified mass spectrum view of 34A. Sp-TRH synthetic peptide (446.2) detected at 9.5 to 10.3 minutes (*).
- C) Full mass spectrum indicating detection of Sp-TRH peptide (446.2) (*). Spectrum obtained from sea urchin *S. purpuratus* body wall extract.
- D) Magnified mass spectrum view of 38C. Spectrum obtained from sea urchin *S. purpuratus* body wall extract. Putative Sp-TRH neuropeptide (446.2) detected from 9.5 to 10.3 minutes (*).

4.3.3 Results: Pharmacology

Sp-GnRH caused contraction of sea urchin oesophagus at a similar efficacy to echinotocin at 3 μ M. The peptide caused contraction at a concentration of 3 μ M with an average force of 0.657 mN ($n = 7$) on an isometric transducer (representative traces are illustrated in Figure 45). However, the synthetic peptide had no observable effect on unstripped sea urchin tube foot preparations ($n = 6$). When tested on starfish muscle preparations the peptide caused contraction of starfish oesophagus and stripped tube foot preparations at concentrations from 100 nM to 3 μ M recorded using an isotonic transducer (Gill, unpublished), but had no observable effect on unstripped preparations (Navaratnam, unpublished).

Sp-TRH caused contraction of sea urchin unstripped tube foot tissue but with lower potency than both echinotocin and NGFFFamide. The peptide caused contractions at concentrations from 3 μ M (0.139 mN, $n = 1$) to 100 μ M (0.493 mN, $n = 1$). The average force of contraction at a concentration of 10 μ M Sp-TRH was 0.637 mN ($n = 11$) (representative traces are illustrated in Figure 46). However, the peptide did not induce contractions of sea urchin oesophagus preparations.

GKamide did not induce contraction of any of the sea urchin preparations tested, including unstripped tube foot (1 μ M, $n = 5$; 10 μ M, $n = 1$), stripped tube foot (3 μ M, $n = 4$), and oesophagus (3 μ M, $n = 3$). PPLN1 also did not induce contraction of the sea urchin preparations tested, including stripped tube foot (1 μ M, $n = 2$), unstripped tube foot (1 μ M, $n = 2$), and oesophagus

(3 μM , $n = 2$). However the peptide caused contraction of starfish apical muscle at 1 μM (0.959 mN, $n = 2$) and 3 μM (2.502 mN, $n = 1$), and unstripped tube foot at 20 μM (0.400 mN, $n = 1$).

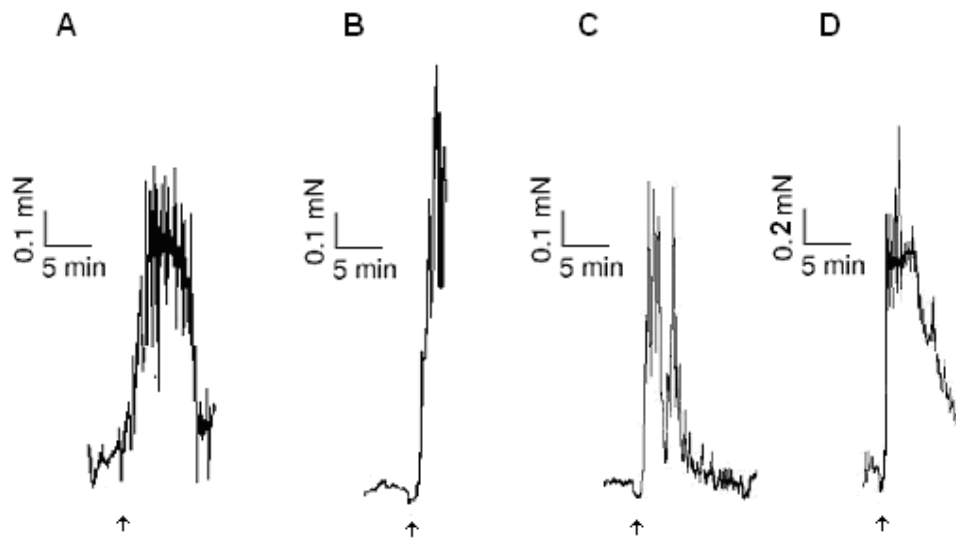


Figure 45. Four representative traces illustrating Sp-GnRH ($3\mu\text{M}$) induced contraction of *E. esculentus* oesophagus preparations. The mean force of contraction was 0.657 mN (s.e.m. = 0.142 , $n = 7$). The individual forces of the contractions illustrated were calculated from displacement on an isometric transducer; 0.626 mN (A), 0.702 mN (B), 0.598 mN (C), and 1.307 mN (D).

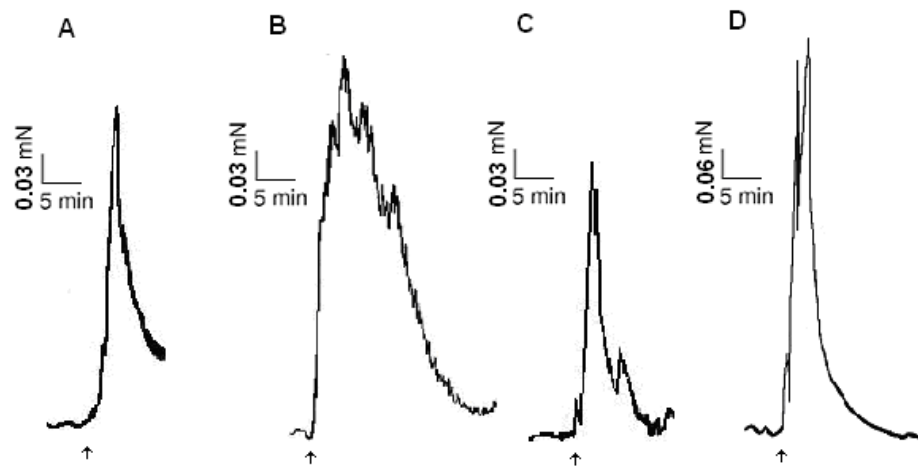


Figure 46. Four representative traces illustrating Sp-TRH ($3 \mu\text{M}$) induced contraction of *E. esculentus* tube foot tissue. The mean force of contraction was 0.637 mN (s.e.m. = 0.171 , $n = 11$). The individual forces of the illustrated contractions were calculated from displacement on an isometric transducer; 0.361 mN (A), 0.452 mN (B), 0.292 mN (C), and 0.959 mN (D).

4.4 Discussion

Ten putative neuropeptide precursor genes have now been identified in the sea urchin genome, including genes encoding homologues of vasotocin and NGIWAYamide, reported by Elphick & Rowe (2009), and the SALMFamide family neuropeptides identified by Elphick & Thorndyke (2005). The novel putative neuropeptides precursors reported in this chapter include homologues of the SALMFamide neuropeptides, gonadotropin-releasing hormone (Sp-GnRH), thyrotropin-releasing hormone (Sp-TRH), calcitonin-gene related peptide (Sp-CGRP), and three novel precursor genes unrelated to known neuropeptides.

4.4.1 SpurS8 & SpurS9: two peptides encoded by a second putative sea urchin SALMFamide neuropeptide precursor

The identification of a second SALMFamide neuropeptide family precursor gene confirms the presence of at least two genes from this neuropeptide family in the sea urchin genome and constitutes the first evidence of SALMFamides being encoded by more than one precursor in an echinoderm.

The two peptides encoded by the gene were recognised by alignment to a consensus sequence suggested by Elphick and Melarange (2001); SxL/FxFamide, where x is a variable residue. However each putative neuropeptide differs from the consensus sequence (Figure 47). If the new peptides are to be considered members of the SALMFamide neuropeptide

family, the consensus sequence may need revising. SpurS8 has an isoleucine residue in the third position from the C-terminus and SpurS9 has a proline residue in place of the serine. Following the sequential nomenclature suggested by Elphick and Thorndyke (2005) for the other sea urchin SALMFamides, the two novel SALMFamide peptides were denoted SpurS8 and SpurS9.

Figure 47 compares the sequences of all currently known SALMFamide family neuropeptides. The first four residues of the SpurS9 nonapeptide resembles the first four residues (Met-Arg-Tyr-His) of a peptide purified from a sea urchin *Echinus esculentus* peptide and exhibiting immunoreactivity to an antiserum raised against starfish SALMFamide 2 (S2) (Elphick and Thorndyke, 2005). Although the sequence is incomplete, the fragment aligns with the N-terminal region of the novel SpurS9 peptide. A single residue substitution, leucine in place of tyrosine, differentiates the two homologues. However, the organisms belong to two separate sea urchin families thought to have separated during the Eocene epoch (30-50 mya). Therefore, some variation in amino acid sequence might be expected (Smith, 1988; Littlewood and Smith, 1995). Although two SALMFamide peptides have been identified in two separate sea cucumber species, in *Holothuria glabberima* (Díaz-Miranda et al., 1995) and *Apostichopus japonicus* (Ohtani et al., 1999), the lack of similarity between the peptides may suggest that homologues may be encoded by separate genes in each species. By comparison with the other echinoderm SALMFamide family members, two structural sub-families have emerged depending on the identity of the third residue from the C-terminus, in the first instance a

hydrophobic residue and in the second instance an aromatic phenylalanine residue. On the basis of this structural division one might predict a second asteroid SALMFamide gene encoding peptides defined by the presence of two C-terminal aromatic residues.

SALMFamide neuropeptides were originally identified by FMRFamide antibody cross-reactivity (Chapter One) (Elphick et al., 1991a).

FMRFamide was originally isolated from the clam *Macrocallista nimbosa* and is partially identical to met-enkephalin opioid peptide (YGGFMRF) (Price and Greenberg, 1977). Subsequently five C-terminal amidated Arg-Phe (RFamide) neuropeptides potentially related to molluscan FMRFamide have been identified in vertebrates. The RFamide peptides appear to have functional homology in the regulation of both protostome and deuterostome feeding behaviour (reviewed in Dockray, 2004). A comprehensive BLAST search of the radial nerve tissue ESTs using the motif RFG_{xx} (where *xx* represents a basic residue cleavage site) failed to identify RFamide peptide C-terminal consensus sequences in the radial nerve tissue EST sequences, suggesting RFamide neuropeptides may have diverged in sequence structure or may have been entirely lost in the divergence of echinoderms from other deuterostomes. Interestingly, the C-terminal seven residues of SpurS6 align with those of the vertebrate RFamide 26RFa (GGSAFVF/GGFSFRF), particularly in the conservation of two C-terminal aromatic residues, and that the terminal four residues of SpurS9 align with those of vertebrate RFamide kisspeptin (GLRF/GLHF), with a basic histidine in place of a basic arginine residue. Notwithstanding the possibility of convergent evolution, the opioids, protostome and

deuterostome RFamides, and echinoderm SALMFamides, may form a super-family of related neuropeptides.

4.4.2 Sea urchin gonadotropin-releasing hormone-like peptide

Sea urchin homologues of gonadotropin-releasing hormone (GnRH) and thyrotropin-releasing hormone (TRH) have been identified. Both vertebrate peptides were amongst the first neuropeptides to be identified in the animal kingdom. Although the peptides share functional homology as vertebrate hypothalamic releasing hormones their secretion and expression appears to be unrelated (Hokfelt et al., 2000).

Similar to other protostome and deuterostome GnRH-like peptides, Sp-GnRH may have a C-terminal Pro-Gly-NH₂ and an N-terminal pyro-glutamate (Rastogi et al., 2002; Tsai, 2006). Figure 48 illustrates a comparison of GnRH-like peptide sequences from a range of bilaterian phyla. All peptide sequences show a high level of homology. Chordate GnRH-family neuropeptides all consist of ten residues, but there is variation in the length of non-vertebrate GnRH-like peptides. Ten GnRH-like peptides have been identified in a chordate deuterostome, the sea squirt *Ciona intestinalis* (reviewed by Tsai, 2006). Only one non-chordate protostome GnRH-like peptide (Oct-GnRH) has been sequenced, from the mollusc *Octopus vulgaris* (Iwakoshi et al., 2002). Like Sp-GnRH, Oct-GnRH is a dodecapeptide with a conserved C-terminal residue sequence of Gly-Trp-Arg/His-Pro-Gly-NH₂.

The human GnRH precursor consists of three coding exons, a pattern that is conserved across vertebrate GnRH-like peptide precursors. An additional gonadotropin-associated peptide (GAP) of 52 residues, and of a currently unknown function, is encoded by all three exons. The echinoderm GnRH-like peptide precursor also consists of three coding exons. Furthermore the precursor encodes a C-terminal peptide of 85 residues, although the peptide has little sequence homology to vertebrate GAP.

Prior to this study the only protostome GnRH-like peptide to be sequenced comes from the octopus *Octopus vulgaris* (Iwakoshi et al., 2002).

Localization studies of oct-GnRH also suggest a reproductive regulatory role in cephalopods (Di Cosmo and Di Cristo, 1998; Di Cristo et al., 2002; Iwakoshi et al., 2002; Iwakoshi-Ukena et al., 2004), bivalves (Pazos and Mathieu, 1999; Gorbman et al., 2003) and gastropods (Goldberg et al., 1993; Young et al., 1999; Zhang et al., 2000; Wayne et al., 2004), where GnRH IR has been observed in the innervation of the vas deferens and muscle of the preputium, or as a neuromodulator (Tsai et al., 2003). In a study by Pazos and Mathieu (1999) on the effects of vertebrate GnRH peptides on bivalve gonad the effective concentration was as low as 10 pM, indicating highly conserved receptor affinity irrespective of peptide sequence divergence. Although protostome GnRH-like peptides have only been identified in molluscs the peptides may be related to insect adipokinetic hormones. Adipokinetic hormones have been shown to bind nematode GnRHR-like receptors and an adipokinetic/GnRH-like peptide has recently been identified in the genome of the nematode *C. elegans* (Lindemans et al., 2009) (Figure 48). The non-amidated putative nematode

adipokinetic/GnRH-like peptide bears structural similarity to arthropod adipokinetic hormone in having nine residues, four of which are identical matches, and a pyroglutamate residue at the N-terminal.

Several studies have hypothesized a cross-phyletic and conserved role for GnRH-like peptides in the direct and indirect regulation of reproduction (Gorbman and Sower, 2003; Tsai, 2006). For example, octopus GnRH-like peptide has been shown to target the optic gland in cephalopods and the vas deferens and preputium muscle in gastropods (Di Cosmo and Di Cristo, 1998; Young et al., 1999). In tunicates, GnRH-like peptide immunoreactivity has been found in the nerve plexus that innervates the gonad and gonoducts (Mackie, 1995).

Molluscs have two GnRH sequences, one distributed in the molluscan CNS, immunoreactive to an antiserum raised against Tunicate-I GnRH, and the other immunoreactive to an antiserum raised against mammalian GnRHs and distributed in the osphradium. The osphradium is one of three major chemosensory organs in the mollusc, the others being rhinophores and tentacles.

In vertebrates the role of GnRH is well established as a regulator of gonadotropin release. In mammals GnRH is involved in stimulating the secretion of LH and FSH from the anterior pituitary. However the peptide is also expressed in the central nervous system, originating from progenitor cells in the vertebrate olfactory placode, thereby also suggesting an ancestral role in chemosensory perception (reviewed by Tsai, 2006). Given

the radical morphological divergence exhibited by echinoderms, the functional roles described by the above studies and the observation of contractile myoactivity on sea urchin oesophagus tissue can only be suggestive until further studies are able to localize GnRH expression in the sea urchin adult and embryo.

4.4.3 Sea urchin thyrotropin-releasing hormone-like peptide

Echinoderm thyrotropin-releasing hormone-like peptide (Sp-TRH) (pEYPG-NH₂) has four residues as opposed to the three residues common to all vertebrate TRH peptides (pEHP-NH₂). It also deviates from a consensus sequence universally observed across chordate TRH neuropeptides in possessing an aromatic residue in the place of a basic histidine residue and an additional C-terminal amidated glycine residue. Furthermore, a peptide corresponding to the predicted structure based on Sp-TRH (pEYPG-NH₂) was detected in *S. purpuratus* body wall extract, suggesting the presence of an N-terminal pyroglutamate residue and C-terminal amidation. Although weakly contractile compared to echinotocin, NGFFFamide, and Sp-GnRH, the peptide's myoactivity when tested on sea urchin tube foot tissues provides an indication of bioactivity.

TRH secreting neurons are found throughout the vertebrate nervous system, being particularly concentrated in the hypothalamus where axons target the median eminence. The neuropeptide targets TRH-specific GPCRs expressed by cells of the anterior pituitary through a phosphatidylinositol second-messenger system (Gershengorn and Thaw, 1985). The anterior

pituitary cells respond by secreting thyroid-stimulating hormone (TSH) and prolactin. The peptide also acts on hypoglossal motor neurons via a G protein-mediated Ca^{2+} -independent second messenger pathway (Bayliss et al., 1994).

Hsieh et al. (1990) have reported TRH immunoreactivity in the shrimp and TRH-like receptors have been identified in the *Drosophila* genome (Hewes and Taghert, 2001), however TRH-like peptides have not yet been identified in protostome or non-chordate species. Therefore, Sp-TRH is the first putative TRH-like peptide to be identified in a non-chordate.

4.4.4 Sea urchin calcitonin/CGRP-like peptide

Like vertebrate calcitonin peptides, the sea urchin calcitonin/CGRP peptide sequence (38 residues) has an amidated C-terminal proline residue and may have an N-terminal disulfide bonded ring structure formed by C⁴ and C¹⁰. Immunoreactivity to calcitonin-like peptides has been shown in both vertebrates and invertebrates (Sasayama et al., 1991). Figure 49 provides a structural comparison of peptide sequences for vertebrate calcitonin, vertebrate CGRP, echinoderm Sp-calcitonin/CGRP, and two arthropod diuretic hormones (Dippu-DH₃₁ and Drome-DH₃₁) with relatively low sequence similarity to CGRP neuropeptides identified in vertebrates (Furuya et al., 2000; Coast et al., 2001; Johnson et al., 2005). Although CGRP-like peptides have been reported in molluscs (Fouchereau-Peron, 1993; Duvail et al., 1997; Lafont et al., 2007), their structure and sequence is not yet known.

Vertebrate calcitonin and calcitonin gene-related peptide are encoded by the same gene. Alternate splicing to produce calcitonin (32 residues) involves the first four exons, whereas the processing of the precursor to produce CGRP (37 residues) involves exons one to three and five and six. The putative echinoderm CGRP precursor (*spp-4*) has five exons, the fourth of which encodes the peptide sequence (Sp-calcitonin/CGRP) (38 amino acids) (Figure 40). There were no other calcitonin-like sequences found in proximity to the gene. The presence of a single calcitonin/CGRP-like peptide sequence in the precursor suggests that calcitonin may have arisen as a later evolutionary development in vertebrates, as suggested by Lafont et al (2007), however, unlike other vertebrate CGRP neuropeptides the sea urchin calcitonin/CGRP-like peptide has an amidated proline residue at the C-terminus (Figure 44).

Calcitonin gene-related peptide (CGRP) has a role as a neuromediator and neurohormone in the nervous systems of invertebrates and chordates, and a role that may have been secondarily-acquired as a regulator of ionic exchange mechanisms in protostomes (Lafont et al., 2007).

Immunoreactivity to calcitonin-like receptors has been demonstrated in the gill and mantle of molluscs, leading Lafont et al. (2007) to hypothesise an ancestral role for the peptide in ionic regulation. Interestingly, activity-modifying proteins, such as RAMPs (receptor activity-modifying proteins) and RCP (receptor coupled protein), are involved in the transport and activity of human CGRP receptors. Protein homologues to RAMP were absent from the *Drosophila* genome (Hewes and Taghert, 2001), although

Johnson et al. (2005) report functional interaction between *Drosophila* RCP (receptor coupling protein) and a calcitonin-like receptor. A putative calcitonin receptor, a member of the secretin family of G-protein coupled receptors, was identified in the *Strongylocentrotus* genome (SPU_018314). However, a BLAST search of the sea urchin genome failed to identify proteins showing homology to human RAMP1 or to *Drosophila* RCP (CG4875). If Sp-CGRP is related to vertebrate CGRP and the *Drosophila* neuropeptide DH₃₁, either the regulatory subunits are unnecessary for physiological activity in the sea urchin or there may be unknown sea urchin receptor-associated regulatory subunits lacking clear homology to those previously described in arthropod and mammalian studies.

4.4.5 Two novel pedal peptide-like peptide precursor genes

Precursor *Spp-6* encodes 22 variable repeats of a non-amidated pentadecapeptide and precursor *spp-7* encodes ten variable repeats of a non-amidated 14 residue peptide (consensus sequence Asn-Phe-Gly-Gly-Ser-Leu-Glu-Pro-Met-Gln-Ser-Gly-Phe-Tyr). The peptides encoded by these last two genes were denominated pedal peptide-like peptide by virtue of weak structural similarity to 'pedal peptide', a neuropeptide unique to molluscs. The peptides share a Ser-Gly-Phe sequence and a non-amidated C-terminus with the three pedal peptide family neuropeptides identified in *Aplysia* (Lloyd and Connolly, 1989; Pearson and Lloyd, 1989; Moroz et al., 2006). A peptide corresponding to the synthetic PPLN1, with a sequence representing a consensus of *spp-6* peptide structures, was detected in *S. purpuratus* body wall tissue extract using mass spectrometry. Furthermore,

although the peptide did not exhibit myoactivity when tested on sea urchin muscle preparations, subsidiary studies have revealed that the peptide appears to be potent in contracting starfish apical muscle at a concentration of 1 μ M (0.959 mN, n = 2).

4.4.6 GKamide: a novel sea urchin peptide

The neuropeptides described in this study were selected on the basis of sequence homology to known neuropeptides and by the presence of homologous peptide repeats. Of the latter, *spp-5* encodes 13 variable repeats of a putative arginine-rich amidated undecapeptide denoted GKamide. Interestingly, although structurally dissimilar, the allatostatin precursor from the American cockroach *Periplaneta americana* encodes 14 allatostatin peptides, a similar number to that encoded by the *S. purpuratus* GKamide precursor. Like PPLN1, the synthesized peptides did not exhibit myoactivity when tested on sea urchin muscle preparations. Although the peptide was not detected using mass spectrometry, it is important to note that GKamide was not identifiable in the standard solutions used for identification of the peptides, suggesting difficulties in retaining the peptide on the column which may have exacerbated over the duration of the study.

Some of the neuropeptides identified in this chapter do provide evidence of the remarkable conservation of neuropeptide sequence structure. However neuropeptides can assume multifunctional roles over evolutionary history, often differentially expressed to fulfil specific functions within an organism. It remains for future research to investigate the neural expression

of these peptides with the hope of elucidating some of their neurobiological, behavioural and physiological functions.

Peptide	Sequence	Species
EscS9 (partial)	Met-Arg-Tyr-His...	<i>E. esculentus</i> (Echinoidea)
SpurS9	Met-Arg-Leu-His-Pro-Gly-Leu-Leu-Phe-NH ₂	<i>S. purpuratus</i> (Echinoidea)
SpurS8	Asn-Met-Gly-Ser-Ile-His-Ser-His-Ser-Gly-Ile-His-Phe-NH ₂	<i>S. purpuratus</i> (Echinoidea)
Kisspeptin (partial)...	Tyr-Tyr-Trp-Tyr-Ser-Phe-Gly-Leu-Arg-Phe-NH ₂	<i>H. sapiens</i> (Mammalia)
S1	Gly-Phe-Asn-Ser-Ala-Leu-Met-Phe-NH ₂	<i>A. rubens</i> (Asteroidea)
S2	Ser-Gly-Pro-Tyr-Ser-Phe-Asn-Ser-Gly-Leu-Thr-Phe-NH ₂	<i>A. rubens</i> (Asteroidea)
GFSKLYFamide	Gly-Phe-Ser-Lys-Leu-Tyr-Phe-NH ₂	<i>H. glaberrima</i> (Holothuroidea)
SGYSVLYFamide	Ser-Gly-Tyr-Ser-Val-Leu-Tyr-Phe-NH ₂	<i>H. glaberrima</i> (Holothuroidea)
GYSPEFamide	Gly-Tyr-Ser-Pro-Phe-Met-Phe-NH ₂	<i>S. japonicus</i> (Holothuroidea)
FKSPFamide	Phe-Lys-Ser-Pro-Phe-Met-Phe-NH ₂	<i>S. japonicus</i> (Holothuroidea)
SpurS1	Pro-Pro-Val-Thr-Thr-Arg-Ser-Lys-Phe-Thr-Phe-NH ₂	<i>S. purpuratus</i> (Echinoidea)
SpurS2	Asp-Ala-Tyr-Ser-Ala-Phe-Ser-Phe-NH ₂	<i>S. purpuratus</i> (Echinoidea)
SpurS3	Gly-Met-Ser-Ala-Phe-Ser-Phe-NH ₂	<i>S. purpuratus</i> (Echinoidea)
SpurS4	Ala-Gln-Pro-Ser-Phe-Ala-Phe-NH ₂	<i>S. purpuratus</i> (Echinoidea)
SpurS5	Gly-Leu-Met-Pro-Ser-Phe-Ala-Phe-NH ₂	<i>S. purpuratus</i> (Echinoidea)
SpurS7	Gly-Asp-Leu-Ala-Phe-Ala-Phe-NH ₂	<i>S. purpuratus</i> (Echinoidea)
SpurS6	Pro-His-Gly-Gly-Ser-Ala-Phe-Val-Phe-NH ₂	<i>S. purpuratus</i> (Echinoidea)
26RFamide (partial)	...Gly-Gly-Phe-Ser-Phe-Arg-Phe-NH ₂	<i>R. esculenta</i> (Amphibia)
[Met ⁵]Enkephalin-Arg ⁶ -Phe ⁷	Tyr-Gly-Gly-Phe-Met-Arg-Phe	<i>H. sapiens</i> (Mammalia)

Figure 47. SALMFamide neuropeptide family sequences assembled from studies of sea urchins (*E. esculentus*, *S. purpuratus*) (this study, Elphick and Thorndyke, 2005), starfish (*A. rubens*) (Elphick et al., 1991a), and holothurians (*H. glaberrima*, *S. japonicus*) (Díaz-Miranda et al., 1995; Iwakoshi et al., 1995). For point of comparison the sequences of human Met-enkephalin-Arg-Phe (GI:49456613), amphibian 26RFamide (GI:50400830), and kisspeptin/metastin (GI:160110095) have been included. Nonpolar residues are labelled in orange, polar residues in green, acidic residues in purple, basic residues in blue, and aromatic residues in red.

Class	Sequence
Mammalia ¹	pyroGlu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH ₂
Chicken-I ²	pyroGlu-His-Trp-Ser-Tyr-Gly-Leu-Gln-Pro-Gly-NH ₂
Chicken-II ³	pyroGlu-His-Trp-Ser-His-Gly-Trp-Tyr-Pro-Gly-NH ₂
Actinopterygii ⁴	pyroGlu-His-Trp-Ser-Tyr-Gly-Trp-Leu-Pro-Gly-NH ₂
Chondrichthyes ⁵	pyroGlu-His-Trp-Ser-His-Asp-Trp-Lys-Pro-Gly-NH ₂
Lamprey-I ⁶	pyroGlu-His-Tyr-Ser-Leu-Glu-Trp-Lys-Pro-Gly-NH ₂
Lamprey-III ⁷	pyroGlu-His-Trp-Ser-His-Asp-Trp-Lys-Pro-Gly-NH ₂
Tunicate-I GnRH ⁸	pyroGlu-His-Trp-Ser-Asp-Tyr-Phe-Lys-Pro-Gly-NH ₂
Tunicate-II GnRH ⁹	pyroGlu-His-Trp-Ser-Leu-Cys-His-Ala-Pro-Gly-NH ₂
Tunicate-III GnRH ⁹	pyroGlu-His-Trp-Ser-Tyr-Glu-Phe-Met-Pro-Gly-NH ₂
Tunicate-IV GnRH ⁹	pyroGlu-His-Trp-Ser-Asn-Gln-Leu-Thr-Pro-Gly-NH ₂
Tunicate-V GnRH ⁹	pyroGlu-His-Trp-Ser-Tyr-Glu-Tyr-Met-Pro-Gly-NH ₂
Tunicate-VI GnRH ⁹	pyroGlu-His-Trp-Ser-Lys-Gly-Tyr-Ser-Pro-Gly-NH ₂
Tunicate-VII GnRH ⁹	pyroGlu-His-Trp-Ser-Tyr-Ala-Leu-Ser-Pro-Gly-NH ₂
Tunicate-VIII GnRH ⁹	pyroGlu-His-Trp-Ser-Leu-Ala-Leu-Ser-Pro-Gly-NH ₂
Tunicate-IX GnRH ⁹	pyroGlu-His-Trp-Ser-Asn-Lys-Leu-Ala-Pro-Gly-NH ₂
Tunicate-X GnRH ⁹	pyroGlu-His-Phe-Ser-Asn-Trp-Trp-Ile-Pro-Gly-Ala-Pro-Gly-Tyr-Asn-Gly-NH ₂
Echinoidea Sp-GnRH ¹⁰	pyroGlu-Val-His-His-Arg-Phe-Ser-Gly-Trp-Arg-Pro-Gly-NH ₂
Cephalopoda Oct-GnRH ¹¹	pyroGlu-Asn-Tyr-His-Phe-Ser-Asn-Gly-Trp-His-Pro-Gly-NH ₂
Arthropoda BomAKH	pyroGlu-Leu-Thr-Phe-Thr-Ser-Ser-Trp-Gly-NH ₂
Nematoda AKH-GnRH	pyroGlu-Met-Thr-Phe-Thr-Asp-Gln-Trp-Thr

Figure 48. A comparison of GnRH-like neuropeptide sequences from four chordate classes, as well as one molluscan GnRH-like peptide, and the putative echinoderm neuropeptide Sp-GnRH. For point of comparison with adipokinetic hormones, *Bombyx* adipokinetic hormone and an AKH-GnRH-like peptide from *C. elegans* have been included. Hydrophobic residues are shown in orange, hydrophilic residues are shown in green, aromatic residues are shown in red, acidic residues are shown in purple, and basic residues are shown in blue. GnRH-like peptides sequences were obtained from ¹*Homo sapiens* (GI:31956, Hayflick et al., 1989), ²*Gallus gallus* (GI:124249294, King and Millar, 1982), ³*Gallus gallus* (GI:89274005), ⁴*Oncorhynchus nerka* (GI:1030714, Ashihara et al., 1995), ⁵*Squalus acanthias* (GI:121520, Lovejoy et al., 1992), ⁶*Petromyzon marinus* (GI:8698621, Suzuki et al., 2000), ⁷*Petromyzon marinus* (GI:37542417, Silver et al., 2004), ⁸*Chelyosoma productum* (Powell et al., 1996), ⁹*Ciona intestinalis* (Tello et al., 2005; review by Tsai, 2006), ¹⁰*Strongylocentrotus purpuratus*, ¹¹*Octopus vulgaris* (GI:19697930, Iwakoshi et al., 2002), ¹²*Bombyx mori* (GI:193248566, Roller et al., 2008), and ¹³*Caenorhabditis elegans* (GI:3329621, Lindemans et al., 2009).

Gene	Sequence
<i>H. sapiens</i> CT ¹	<u>CG</u> NLSTCMLGTYTQDFNKFH ^T FPQTAIGVGP-NH ₂
Sp-CT/CGRP ²	SKGCGSFSGCMQMEVAKNRVAALLRNSNAHLFGLNGP-NH ₂
<i>Rana</i> CGRP ³	ACNTATCVTHRLADFLSRSGGMAKNNFVPTNVGSKAF-NH ₂
<i>S. scrofa</i> CGRP ⁴	<u>SC</u> NTATCVTHRLAGLLSRSGGMVKSNFVPTDVGSEAF-NH ₂
Dippu-DH31 ⁵	GLDLGLSRGFSGSQAAKHLMGLAAANYAGGP-NH ₂
Drome-DH31 ⁶	TVDFGLARGYSGTQEAKHRMGLAAANFAGGP-NH ₂

Figure 49. Calcitonin and CGRP-like peptide sequences from ¹human (GI:179820), ²sea urchin (*S. purpuratus*), ³anuran (*R. ridibunda*) (GI:399232) (Conlon et al., 1993), ⁴pig (*Sus scrofa*) (GI:156447412) (Kimura et al., 1987), ⁵cockroach (Dippu-DH31) (*Diploptera punctata*) (GI:8928102) (Furuya et al., 2000), and ⁶fruit fly (Drome-DH31) (*D. melanogaster*) (GI:17647327). Nonpolar residues are labelled in orange, polar residues in green, acidic residues in purple, basic residues in blue, and aromatic residues in red. The underlined conserved cysteine residues may join to form a disulfide bond.

Chapter 5

DISCUSSION

Nine putative neuropeptide precursors were identified in the genome of the Californian sea urchin *Strongylocentrotus purpuratus* using bioinformatic techniques and several of the putative peptides identified were detected using mass spectrometry. Six of the putative peptides were synthesized and demonstrated myoactivity on *in vitro* muscle preparations from sea urchin (*Echinus esculentus*) and starfish (*Asterias rubens*). The identified genes include the precursors to the peptides echinotocin, NGFFFamide, Sp-TRH, Sp-GnRH, SALMFamides S8 and S9, Sp-calcitonin/CGRP and PPLN1, PPLN2 peptides, and GKamide.

Before the current post-genomic era, immunological methods were used to detect the presence of peptides immunoreactive to antibodies raised against known neuropeptides as a prologue to the identification and sequencing of novel peptides (Robb et al., 1989; Díaz-Miranda et al., 1991; Keller, 1992; Ghyoot et al., 1994). Bioinformatic '*in silico*' approaches to neuropeptide identification have found 22 putative neuropeptide-encoding genes in the *Drosophila* genome (Hewes and Taghert, 2001; Vanden Broeck, 2001b) and up to 35 putative neuropeptide-encoding genes in the *Anopheles gambiae* genome (Riehle et al., 2002). MALDI-TOF MS has been used to identify and sequence neuropeptides from invertebrate neurosecretory tissues, including the atrial gland of the sea slug *Aplysia californica* (Li et al., 1999b) and the crustacean pericardial organs of the crustacean *Cancer borealis* (Li et al., 2003). Subsequently a total of 119 putative arthropod neuropeptides have been identified in the *Drosophila* genome (Liu et al., 2006).

Nematode neuropeptides have also been identified using a combination of genomic and biochemical techniques (Husson et al., 2005; Husson et al., 2007). An 'in silico' approach, including the use of preproprotein characteristics as identifying criteria, was used by Nathoo et al (2001) to identify 32 putative neuropeptide-like genes in the *Caenorhabditis elegans* genome. The neuropeptides identified in the roundworm genome show considerable diversification and expansion and have been divided into several novel gene families, including two large families of FMRF-related neuropeptides and insulin-like neuropeptides.

A second putative SALMFamide neuropeptide precursor

In this study, a second putative SALMFamide neuropeptide precursor gene was identified in the *Strongylocentrotus* genome, encoding two novel SALMFamide peptides, S8 and S9. The two putative neuropeptides supplement the seven encoded by the other SALMFamide neuropeptide identified in the sea urchin genome (Elphick and Thorndyke, 2005). The family may be unique to the echinoderm phylum. However the new sequences identified modify the consensus SALMFamide neuropeptide sequence previously suggested by Elphick & Melarange (2001), SxL/FxFamide, to one set of SALMFamide neuropeptides defined by a hydrophobic residue in the third position from the C-terminus (S8 & S9) and a second set defined by an aromatic residue in the third position from the C-terminus (S1-S7). The presence of similar peptide sequences identified in both holothurians and asteroid species supports the hypothesis of two separate SALMFamide neuropeptide precursors.

The SALMFamide neuropeptide family may be evolutionarily linked to the FMRFamide-related peptides and -RFamides, which have been identified in both vertebrates and invertebrates and have been shown to have multiple roles in behaviour and physiology (Dockray, 2004). There were no putative -RFamide neuropeptides identified in the sea urchin genome, although a number of receptors homologous to vertebrate -RFamide receptors were identified. Although the physiological roles of the SALMFamide neuropeptides are not yet known in the sea urchin, the family members have been shown to be myoactive in other echinoderms, notably inducing relaxation of starfish stomach tissue (Elphick and Melarange, 2001). However, the sea urchin genome encodes nine putative SALMFamide peptides, some of which may not be myoactive and all of which may vary in physiological role. The roles of the SALMFamides in sea urchins and the variation in peptide structure and function of the SALMFamides across the five echinoderm classes remains to be determined.

Neurophysin-containing neuropeptide precursors: echinotocin and NGFFFamide

One NGIWYamide-like peptide, NGFFFamide, has been identified. Both of the peptides may represent another family of neuropeptides unique to echinoderms. However it is unclear if NGFFFamide is a true functional orthologue of holothurian NGIWYamide without further pharmacological characterization of the neuropeptide in both sea urchins and starfish. Most interestingly NGFFFamide may be related to the vasopressin/oxytocin

family of neuropeptides because the NGFFFamide precursor gene also encodes a C-terminal neurophysin domain. Furthermore, the precursor to a vasotocin-like peptide was identified and denominated echinotocin. Consistent with all known vasotocin-like peptide precursors a neurophysin domain was also identified C-terminal to the peptide encoding region. The echinotocin neurophysin molecule is sequentially similar to the oxytocin neurophysin molecule however the NGFFFamide neurophysin molecule lacks important residues in its peptide binding region, which may reflect modifications essential for binding NGFFFamide. Phylogenetic analysis of neurophysin sequence structure and the absence of other neurophysin-associated peptides in both echinoderms and other protostomes suggest that the neurophysin-associated peptide may have originated during echinoderm evolution. It remains to be determined whether other neurophysin domains not associated with vasopressin/oxytocin-like peptides are to be found in the genomes of species from other echinoderm classes or even within the genomes of other invertebrate species.

NGFFFamide was detected in both the neural EST library and in *S. purpuratus* body wall extract using mass spectrometry. Due to the small size of the radial nerve EST library and the low sensitivity of the mass spectrometric analysis, the peptide may be expressed at high quantities throughout the sea urchin nervous system. Interestingly echinotocin was not identified in the EST library, the transcriptome database, or the body wall or visceral extracts analysed using mass spectrometry, suggesting that the peptide may not be expressed, or may only be expressed at low levels, and not necessarily as a neurohormone or neurotransmitter. Pharmacological

investigation demonstrates that both NGFFFamide and echinotocin induce contraction of sea urchin tube foot and oesophagus tissues but NGFFFamide was the more potent of the two peptides. The observations of myoactivity are similar to the responses induced by the holothurian NGIWIamide on both holothurian, echinoid and asteroid tissue (Inoue et al., 1999; Saha et al., 2006). The effects of both NGFFFamide and echinotocin on tissues from other echinoderm classes await thorough investigation.

Putative novel neuropeptides homologous to vertebrate neuropeptides

GnRH-like peptides are expressed throughout the animal kingdom and one GnRH-like peptide encoding precursor was identified in the sea urchin genome. In vertebrates the peptide indirectly regulates reproduction via the release of glycoprotein hormones from the pituitary. The peptide is also associated with reproduction in molluscs. Sp-GnRH is the first GnRH-like neuropeptide identified in echinoderms. The echinoderm peptide has myoactive properties, inducing contraction of sea urchin oesophagus tissue. It remains to be determined whether the peptide has similar myoactive properties in other echinoderm classes.

Homologues to TRH, another tropic hormone released by the pituitary, heretofore have not been identified in non-chordates. The TRH-like peptide identified in this study bears resemblance to vertebrate TRH in the relatively high copy number encoded by the precursor, precursor structure and residue similarity, including C-terminal amidation and an N-terminal

pyroglutamate residue. Furthermore, the peptide was identified in a pyroglutamated and amidated form by mass spectrometric analysis of body wall tissue extract. The peptide induced contraction of sea urchin tube feet, but at a lower potency than both NGFFFamide and echinotocin.

A gene encoding a putative neuropeptide homologous to calcitonin/calcitonin-gene related peptide was identified. In vertebrates both calcitonin and CGRP are encoded by the same gene however a CGRP-encoding gene is thought to have been the common ancestor of the peptides before modification through intragenic duplication (Lafont et al., 2007). Thus, the identification of a single calcitonin/CGRP-like peptide in a deuterostome species provides additional evidence that the duplication event producing calcitonin and CGRP may have occurred in the vertebrate lineage. Immunohistochemical and pharmacological studies will be needed to address the role of Sp-calcitonin/CGRP in the sea urchin and whether the peptide has a role as a neuromediator and/or regulator of calcium metabolism.

Putative novel sea urchin neuropeptide families

Three more putative neuropeptide-encoding precursors were identified by virtue of their structural characteristics. The three putative precursors may encode neuropeptides representing novel echinoderm neuropeptide families. However, two precursors encode peptides showing structural similarity to molluscan pedal peptide. One of the pedal peptide-like peptides was detected in *S. purpuratus* tissue extract analysed using mass

spectrometry. Two of the peptides were synthesized for pharmacological investigation, but neither demonstrated myoactivity when tested on sea urchin muscle preparations.

Conclusions

Investigation of the expression and function of the novel neuropeptides will benefit future echinoderm neurobiological, physiological and behavioural research. Echinoderms are a global economic resource. Sea urchin roe and sea cucumbers are both important in luxury regional cuisines of Japan, South East Asia, China, France, and Peru with approximately 80% of sea urchin consumption occurring in Japan (Yokota, 2002). However, as active predators of marine bivalves, introduced echinoderms transported by ballast water, including the starfish *Asterias amurensis* and *Acanthaster planci*, have damaged the shellfish industry and the endemic fauna of coral reefs in the South Pacific.

Neuropeptides and their receptors are likely to play a role in the regulation of echinoderm reproductive behaviour. Studies suggest that both SALMFamide and NGIWYamide-like peptides may have a role in the regulation of reproduction (Maruyama, 1985; Mita et al., 2004; Mita et al., 2009). Both NGIWYamide and the dimeric gonad-stimulating substance (GSS) stimulate follicle cells surrounding oocytes and the GSS pathway has been well defined. Oocyte meiosis is triggered by the release of maturation-inducing substance (MIS), α 1-methyl adenine, by the follicle cells which then binds to the oocyte membrane initiating a pathway involving a

maturation-promoting factor, conserved across eukaryotes (Kanatani, 1983; Shirai, 1986; Yoshikuni et al., 1988; Tadenuma et al., 1992; Okumura et al., 2002; Hiraoka et al., 2004). A full understanding of the neurohormonal regulation of echinoderm reproduction may be of benefit both to the control of echinoderms as introduced ecological pests and as organisms for aquaculture.

Furthermore, echinoderms have been vital model organisms for research into the early stages of animal development (Briggs and Wessel, 2006; Burke et al., 2006) and several of the putative neuropeptides identified in this study are also expressed by the developing echinoderm larva.

The role of neuropeptides in understanding neural signalling in both vertebrates and invertebrates cannot be underestimated. Neuropeptides are important to understanding the processes of neuromodulation (Brinton, 1990), the biochemical mechanisms involved in intercellular signalling and GPCR activation, immunity, stress (Herman et al., 2003), development (Briggs and Wessel, 2006; Howard-Ashby et al., 2006) and social behaviour. Both oxytocin and vasopressin have been shown to have roles in the regulation of social behaviour in vertebrates, including vasopressin in facilitating aggression and in the maintenance of monogamous pair bonding (Winslow et al., 1993; Caldwell et al., 2008) and oxytocin in maternal-infant bonding and the regulation of social interaction (Donaldson and Young, 2008). Variation in vasopressin and oxytocin receptor structure may be responsible for pathological social behavioural traits in humans (Israel et al., 2008).

Understanding the biochemical activity of neuropeptides and their receptors will be integral to understanding the processes of neural integration and signal processing. Alongside other signalling molecules, neuropeptides may serve as molecular synchronization molecules in neural networks generating rhythmic output, including central pattern generators, in circadian rhythm synchronization, and in learning and memory. For example, neuropeptides released by the pericardial organs may be involved in regulation of the crustacean stomatogastric ganglion (Li et al., 2003) and pigment-dispersing factor is an integral component to the regulation of *Drosophila* circadian rhythms (Taghert and Shafer, 2006). Circadian abnormalities have a disruptive and detrimental effect on longevity, development, and reproductive success. The identification of rhythmic centres regulating periodicity in reproduction and motor action in the dispersed echinoderm nervous system will assist in unravelling the mysteries of echinoderm neural integration.

There is an argument that neuropeptides are amongst the most ancestral neurotransmitters in animals, which is apparent by the presence of a host of cnidarian neuropeptides with neurotransmitter capabilities (Anctil and Grimmelikhuijzen, 1989; McFarlane et al., 1991; Schmutzler et al., 1992). Comparative studies of neuropeptide function and expression will continue to inform studies of molecular and morphological evolution. Future comparative studies of neuropeptide structure and function will also benefit from novel and expanded techniques, including reverse pharmacology, for

the identification of endogenous ligands to orphan GPCRs, and MALDI-TOF, for the identification of novel bioactive neuropeptides.

In summary, the thesis reports the identification of two novel neuropeptide families, homologous peptides to three vertebrate neuropeptides, and one novel precursor encoding two members of the echinoderm SALMFamide neuropeptide family. Since the divergence of echinoderms from other deuterostomes in the late Precambrian supereon, neuropeptides and their receptors in echinoderms are likely to have expanded and acquired novel structures and diverse functions necessary for a highly divergent and unusual morphology and physiology. It may be that, notwithstanding the comparative advantages of a relatively close ancestral relationship to vertebrates, the value of echinoderm neuroendocrine research lies in understanding how radically different echinoderms are from vertebrates in anatomy, physiology, and biochemistry, and how related signalling molecules have acquired novel structures and functions. This project introduces a number of candidate molecules for the regulation of echinoderm physiology and behaviour which will be of high importance to future students and researchers investigating the complexities of echinoderm biology.

References

- Aikins, M. J., D. A. Schooley, K. Begum, M. Detheux, R. W. Beeman and Y. Park (2008). "Vasopressin-like peptide and its receptor function in an indirect diuretic signaling pathway in the red flour beetle." *Insect Biochem Mol Biol* **38**(7): 740-748.
- Allard, M., S. Labrousche, A. Nosjean and R. Laguzzi (1995). "Mechanisms underlying the cardiovascular responses to peripheral administration of NPF in the rat." *J. Pharmacol. Exp. Ther.* **274**: 577-583.
- Anctil, M. and C. J. Grimmelikhuijzen (1989). "Excitatory action of the native neuropeptide antho-RFamide on muscle in the pennatulid *Renilla kollikera*." *Gen. Pharmacol.* **20**(3): 381-384.
- Ashihara, M., M. Suzuki, K. Kubokawa, Y. Yoshiura, M. Kobayashi, A. Urano and K. Aida (1995). "Two differing precursor genes for the salmon-type gonadotropin-releasing hormone exist in salmonids." *J Mol Endocrinol* **15**(1): 1-9.
- Bargmann, C. I. (1998). "Neurobiology of the *Caenorhabditis elegans* genome." *Science* **282**(5396): 2028-2033.
- Bateman, A., E. Birney, R. Durbin, S. R. Eddy, K. L. Howe and E. L. L. Sonnhammer (2000). "The Pfam protein families database." *Nucleic Acids Res.* **28**: 263-266.
- Bayliss, D. A., F. Viana and A. J. Berger (1994). "Effects of thyrotropin-releasing hormone on rat motoneurons are mediated by G proteins." *Brain Res* **668**(1-2): 220-229.
- Belles, X., L. A. Graham, W. G. Bendena, Q. I. Ding, J. P. Edwards, R. J. Weaver and S. S. Tobe (1999). "The molecular evolution of the allatostatin precursor in cockroaches." *Peptides* **20**(1): 11-22.
- Bendena, W. G., B. C. Donly and S. S. Tobe (1999). "Allatostatins: a growing family of neuropeptides with structural and functional diversity." *Ann N Y Acad Sci* **897**: 311-329.
- Bielsky, I. F., S. B. Hu, K. L. Szegda, H. Westphal and L. J. Young (2004). "Profound impairment in social recognition and reduction in anxiety-like behavior in vasopressin V1a receptor knockout mice." *Neuropsychopharmacology* **29**(3): 483-493.
- Birenheide, R., M. Tamori, T. Motokawa, M. Ohtani, E. Iwakoshi, Y. Muneoka, T. Fujita, H. Minakata and K. Nomoto (1998). "Peptides controlling stiffness of connective tissue in sea cucumbers." *Biol Bull* **194**(3): 253-259.
- Blanchard, R. J., G. Griebel, C. Farrokhi, C. Markham, M. Yang and D. C. Blanchard (2005). "AVP V1b selective antagonist SSR149415 blocks

- aggressive behaviors in hamsters." Pharmacol Biochem Behav **80**(1): 189-194.
- Blevins, E. and S. Johnsen (2004). "Spatial vision in the echinoid genus *Echinometra*." J. Exp. Biol. **207**: 4249-4253.
- Boeckxstaens, G. E. and P. A. Pelckmans (1997). "Nitric oxide and the non-adrenergic non-cholinergic neurotransmission." Comp Biochem Physiol A Physiol **118**(4): 925-937.
- Boler, J., F. Enzmann, K. Folkers, C. Y. Bowers and A. V. Schally (1969). "The identity of chemical and hormonal properties of the thyrotropin releasing hormone and pyroglutamyl-histidyl-proline amide." Biochem. Biophys. Res. Commun. **37**: 705-710.
- Boltz, R. E. and D. W. Ewer (1963). "Studies on the myoneural physiology of echinodermata- V. The lantern retractor muscle of *Parechinus*: responses to drugs." J. Exp. Biol. **40**: 727-733.
- Bono, M. d. and C. I. Bargmann (1998). "Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*." Cell **94**: 679-689.
- Bosch, T. C. and T. Fujisawa (2001). "Polyps, peptides and patterning." Bioessays **23**(5): 420-427.
- Bourlat, S. J., T. Juliusdottir, et al. (2006). "Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida." Nature **444**: 85-88.
- Bradbury, A. F., M. D. Finnie and D. G. Smyth (1982). "Mechanism of C-terminal amide formation by pituitary enzymes." Nature **298**(5875): 686-688.
- Bremaeker, N. D., F. Baguet, M. C. Thorndyke and J. Mallefet (1999). "Modulatory effects of some amino acids and neuropeptides on luminescence in the brittlestar *Amphipholis squamata*." J. Exp. Biol. **202**(13): 1785-1791.
- Briggs, E. and G. M. Wessel (2006). "In the beginning...animal fertilization and sea urchin development." Dev Biol **300**(1): 15-26.
- Brinton, R. E. (1990). "Neuromodulation: associative and nonlinear adaptation." Brain Res Bull **24**(5): 651-658.
- Brinton, R. E. and B. S. McEwen (1989). "Vasopressin neuromodulation in the hippocampus." J Neurosci **9**(3): 752-759.
- Brody, T. and A. Cravchik (2000). "Drosophila melanogaster G protein-coupled receptors." J Cell Biol **150**(2): F83-88.

- Bromham, L. D. and B. M. Degnan (1999). "Hemichordates and deuterostome evolution: robust molecular phylogenetic support for a hemichordate + echinoderm clade." Evol. Dev. **1**(3): 166-171.
- Brusca, R. C. and G. J. Brusca (2003). Invertebrates. Sunderland, Massachusetts, Sinauer Associates, Inc.
- Bullock, T. H. and G. A. Horridge (1965). Structure and function in the nervous systems of invertebrates. San Francisco, Freeman.
- Burge, C. and S. Karlin (1997). "Prediction of complete gene structures in human genomic DNA." J. Mol. Biol. **268**: 78-94.
- Burke, R. D., L. M. Angerer, et al. (2006). "A genomic view of the sea urchin nervous system." Dev Biol **300**(1): 434-460.
- Byrne, M. (2001). "The morphology of autotomy structures in the sea cucumber *Eupentacta quinquesemita* before and during evisceration." J. Exp. Biol. **204**: 849-863.
- Byrne, M. and P. Cisternas (2002). "Development and distribution of the peptidergic system in larval and adult *Patiriella*: comparison of sea star bilateral and radial nervous systems." J. Comp. Neurol. **451**(2): 101-114.
- Cai, W. W., R. Chen, R. A. Gibbs and A. Bradley (2001). "A clone-array pooled shotgun strategy for sequencing large genomes." Genome Res. **11**(10): 1619-1623.
- Caldwell, H. K., H. J. Lee, A. H. Macbeth and W. S. Young, 3rd (2008). "Vasopressin: behavioral roles of an "original" neuropeptide." Prog Neurobiol **84**(1): 1-24.
- Cameron, R. A., G. Mahairas, et al. (2000). "A sea urchin genome project: sequence scan, virtual map, and additional resources." Proc Natl Acad Sci U S A **97**(17): 9514-9518.
- Cameron, R. A., M. Samanta, A. Yuan, D. He and E. Davidson (2009). "SpBase: the sea urchin genome database and web site." Nucleic Acids Res **37**(Database issue): D750-754.
- Campbell, A. C. (1983). "Form and Function of Pedicellariae." Echinoderm Studies **1**: 139-167.
- Cazzamali, G. and C. J. Gimmelikhuijzen (2002). "Molecular cloning and functional expression fo the first insect FMRamide receptor." Proc. Natl. Acad. Sci. USA **99**: 12073-12078.
- Chaet, A. B. and R. A. McConnaughy (1959). "Physiological activity of nerve extracts." Biol. Bull. **117**: 407.
- Chartrel, N., F. Bruzzone, C. Dujardin, J. Leprince, H. Tollemer, Y. Anouar, M. Vallarino, J. Costentin and H. Vaudry (2005). "Identification

of 26RFa from Frog Brain: A Novel Hypothalamic Neuropeptide with Orexigenic Activity in Mammals." Ann. N.Y. Acad. Sci. **1040**: 80-83.

Chini, B., B. Mouillac, et al. (1995). "Molecular basis for agonist selectivity in the vasopressin/oxytocin receptor family." Adv Exp Med Biol **395**: 321-328.

Coast, G. M., S. G. Webster, K. M. Schegg, S. S. Tobe and D. A. Schooley (2001). "The *Drosophila melanogaster* homologue of an insect calcitonin-like diuretic peptide stimulates V-ATPase activity in fruit fly Malpighian tubules." J. Exp. Biol. **204**: 1795-1804.

Cobb, J. L. (1978). "An ultrastructural study of the dermal papulae of the starfish, *Asterias rubens*, with special reference to innervation of the muscles." Cell Tissue Res **187**(3): 515-523.

Cobb, J. L. S. and M. S. Laverack (1967). "Neuromuscular systems in echinoderms." Symp. Zool. Soc. Lond. **20**: 25-51.

Cobb, J. L. S. and A. Moore (1989). "Studies on the integration of sensory information by the nervous system of the brittlestar *Ophiura ophiura*." Mar. Behav. Physiol. **14**: 211-222.

Cobb, J. L. S. and V. W. Pentreath (1977). "Anatomical studies of simple invertebrate synapses using stage rotation electron microscopy and densitometry." Tissue Cell **9**: 125-135.

Coleman, R. (1969). "Ultrastructure of the tube foot wall of a regular echinoid, *Diadema antillarum* Philippi." Z. Zellforsch. mikrosk. Anat. **96**: 162-172.

Conlon, J. M., M. C. Tonon and H. Vaudry (1993). "Isolation and structural characterization of calcitonin gene-related peptide from the brain and intestine of the frog, *Rana ridibunda*." Peptides **14**(3): 581-586.

Coppard, S. E. and A. C. Campbell (2005). "The lunar periodicity of diadematid echinoids breeding in Fiji." Coral Reefs **24**: 324-332.

Dale, J. (1999). "Coordination of chemosensory orientation in the starfish *Asterias forbesi*." Mar. Fresh. Behav. Physiol. **32**: 57-71.

Darmer, D., C. Schmutzler, D. Diekhoff and C. J. Grimmlikhuijzen (1991). "Primary structure of the precursor for the sea anemone neuropeptide Antho-RFamide (less than Glu-Gly-Arg-Phe-NH₂)." Proc. Natl. Acad. Sci. USA **88**: 2555-2559.

De Bree, F. M. (2000). "Trafficking of the vasopressin and oxytocin prohormone through the regulated secretory pathway." J Neuroendocrinol **12**(6): 589-594.

de Bree, F. M. and J. P. Burbach (1998). "Structure-function relationships of the vasopressin prohormone domains." Cell Mol Neurobiol. **18**(2): 173-191.

De Lange, R. P. J., F. A. Van Golen and J. V. Van Minnen (1997). "Diversity in cell specific co-expression of four neuropeptide genes involved in control of male copulation behaviour in *Lymnaea stagnalis*." Neurosci. **78**(1): 289-299.

De Lecea, L., T. S. Kilduff, et al. (1998). "The hypocretins: hypothalamus-specific peptides with neuroexcitatory activity." Proc Natl Acad Sci U S A **95**(1): 322-327.

De Loof, A., G. Baggerman, M. Breuer, I. Claeys, A. Cerstiaens, E. Clynen, T. Janssen, L. Schoofs and J. Vanden Broeck (2001). "Gonadotropins in insects: an overview." Arch Insect Biochem Physiol **47**(3): 129-138.

Delsuc, F., H. Brinkmann, D. Chourrout and H. Philippe (2006). "Tunicates and not cephalochordates are the closest living relatives of vertebrates." Nature **439**(7079): 965-968.

Devlin, C. L. (2001). "The pharmacology of gamma-aminobutyric acid and acetylcholine receptors at the echinoderm neuromuscular junction." J Exp Biol **204**(Pt 5): 887-896.

Di Cosmo, A. and C. Di Cristo (1998). "Neuropeptidergic control of the optic gland of *Octopus vulgaris*. FMRF-amide and GnRH immunoreactivity." J. Comp. Neurol. **398**: 1-12.

Di Cristo, C., M. Paolucci, J. Iglesias, J. Sanchez and A. Di Cosmo (2002). "Presence of two neuropeptides in the fusiform ganglion and reproductive ducts of *Octopus vulgaris*: FMRFamide and gonadotropin-releasing hormone (GnRH)." J Exp Zool **292**(3): 267-276.

Díaz-Miranda, L., R. E. Blanco and J. E. García-Arrarás (1995). "Localization of the heptapeptide GFSKLYFamide in the sea cucumber *Holothuria glaberrima* (Echinodermata): A light and electron microscopic study." J. Comp. Neurol. **352**: 626-640.

Díaz-Miranda, L., G. Escalona de Motta and J. E. García-Arrarás (1991). "Localization of neuropeptides in the nervous system of the marine annelid *Sabellastarte magnifica*." Cell Tissue Res. **266**: 209-217.

Díaz-Miranda, L. and J. E. García-Arrarás (1995). "Pharmacological action of the heptapeptide GFSKLYFamide in the muscle of the sea cucumber *Holothuria glaberrima* (Echinodermata)." Comp. Biochem. Physiol. **110C**: 171-176.

Díaz-Miranda, L., D. A. Price, M. J. Greenberg, T. D. Lee, K. E. Doble and J. E. García-Arrarás (1992). "Characterization of two novel

neuropeptides from the sea cucumber *Holothuria glaberrima*." Biol. Bull. **182**: 241-247.

Dibas, A. I., A. J. Mia and T. Yorio (1998). "Aquaporins (water channels): role in vasopressin-activated water transport." Proc Soc Exp Biol Med **219**(3): 183-199.

Disturnal, J. E., W. L. Veale and Q. J. Pittman (1987). "Modulation by arginine vasopressin of glutamate excitation in the ventral septal area of the rat brain." Can J Physiol Pharmacol **65**(1): 30-35.

Dockray, G. J. (1980). "Cholecystokinins in rat cerebral cortex: identification, purification and characterization by immunochemical methods." Brain Res **188**(1): 155-165.

Dockray, G. J. (2004). "The expanding family of -RFamide peptides and their effects on feeding behaviour." Exp. Physiol. **89**(3): 229-235.

Donaldson, Z. R. and L. J. Young (2008). "Oxytocin, vasopressin, and the neurogenetics of sociality." Science **322**(5903): 900-904.

Dreifuss, J. J., M. Dubois-Dauphin, H. Widmer and M. Raggenbass (1992). "Electrophysiology of oxytocin actions on central neurons." Ann N Y Acad Sci **652**: 46-57.

Dufresne, M., C. Seva and D. Fourmy (2006). "Cholecystokinin and gastrin receptors." Physiol Rev **86**(3): 805-847.

Dunn, C. W., A. Hejnol, et al. (2008). "Broad phylogenomic sampling improves resolution of the animal tree of life." Nature **452**: 745-749.

Duvail, L., E. Lopez and M. Fouchereau-Peron (1997). "Characterization of a calcitonin gene-related peptide in the abalone gill." Peptides **20**: 361-366.

East, P., K. Tregenza, H. Duve and A. Thorpe (1996). "Identification of the dipteran Leu-callatostatin peptide family: characterisation of the prohormone gene from *Calliphora vomitoria* and *Lucilia cuprina*." Regul Pept **67**(1): 1-9.

Edmondson, C. H. (1935). "Autotomy and regeneration in Hawaiian starfishes." Occ Pap Bernice P Bishop Mus **11**: 1-29.

Elekonich, M. M. and F. M. Horodyski (2003). "Insect allatotropins belong to a family of structurally-related myoactive peptides present in several invertebrate phyla." Peptides **24**(10): 1623-1632.

Elphick, M. R., I. C. Green and M. O'Shea (1993). "Nitric oxide synthesis and action in an invertebrate brain." Brain Res. **619**: 344-346.

Elphick, M. R. and R. Melarange (2001). "Neural Control of Muscle Relaxation in Echinoderms." J. Exp. Biol. **204**: 875-885.

Elphick, M. R., S. J. Newman and M. C. Thorndyke (1995). "Distribution and action of SALMFamide neuropeptides in the starfish *Asterias rubens*." J. Exp. Biol. **198**: 2519-2525.

Elphick, M. R., D. A. Price, T. D. Lee and M. C. Thorndyke (1991a). "The SALMFamides: a new family of neuropeptides isolated from an echinoderm." Proc. R. Soc. Lond. B **243**: 121-127.

Elphick, M. R., J. R. Reeve, R. D. Burke and M. C. Thorndyke (1991b). "Isolation of the neuropeptide SALMFamide-1 from starfish using a new antiserum." Peptides **12**: 455-459.

Elphick, M. R. and M. L. Rowe (2009). "NGFFFamide and echinotocin: structurally unrelated myoactive neuropeptides derived from neurophysin-containing precursors in sea urchins." J Exp Biol **212**(Pt 8): 1067-1077.

Elphick, M. R. and M. C. Thorndyke (2005). "Molecular characterisation of SALMFamide neuropeptides in sea urchins." J. Exp. Biol. **208**(22): 4273-4282.

Elsik, C. G., A. J. Mackey, J. T. Reese, N. V. Milshina, D. S. Roos and G. M. Weinstock (2007). "Creating a honey bee consensus gene set." Genome Biol. **8**(1): R13.

Emanuelsson, O., S. Brunak, G. von Heijne and H. Nielsen (2007). "Locating proteins in the cell using TargetP, SignalP and related tools." Nat Protoc **2**(4): 953-971.

Epelbaum, J., P. Dournaud, M. Fodor and C. Viollet (1994). "The neurobiology of somatostatin." Crit Rev Neurobiol **8**(1-2): 25-44.

Florey, E. and M. A. Cahill (1977). "Ultrastructure of sea urchin tube feet." Cell. Tiss. Res. **177**: 195-214.

Florey, E. and M. A. Cahill (1980). "Cholinergic motor control of sea urchin tube feet: evidence for chemical transmission without synapses." J Exp Biol **88**: 281-292.

Florey, E., M. A. Cahill and M. Rathmayer (1975). "Excitatory actions of GABA and of acetyl-choline in sea urchin tube feet." Comp Biochem Physiol C **51**(1): 5-12.

Flower, D. R. (1999). "Modelling G-protein-coupled receptors for drug design." Biochim Biophys Acta **1422**(3): 207-234.

Fouchereau-Peron, M. (1993). "Characterization of a molecule related to calcitonin gene related peptide (CGRP) in the scallop, *Pecten maximus*." Comp. Biochem. Physiol. **105B**: 707-711.

Fredriksson, R., M. C. Lagerstrom, L. Lundin and H. B. Schioth (2003). "The G-Protein-Coupled Receptors in the Human Genome Form Five

Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints." Molecular Pharmacology **63**: 1256-1272.

Fujino, Y., T. Nagahama, et al. (1999). "Possible functions of oxytocin/vasopressin-superfamily peptides in annelids with special reference to reproduction and osmoregulation." J Exp Zool **284**(4): 401-406.

Furuya, K., R. J. Milchak, K. M. Schegg, J. Zhang, S. S. Tobe, G. M. Coast and D. A. Schooley (2000). "Cockroach diuretic hormones: characterization of a calcitonin-like peptide in insects." Proc Natl Acad Sci U S A **97**(12): 6469-6474.

García-Arrarás, J. E., M. Rojas-Soto, L. B. Jimenez and L. Diaz-Miranda (2001). "The enteric nervous system of echinoderms: unexpected complexity revealed by neurochemical analysis." J Exp Biol **204**(Pt 5): 865-873.

Gershengorn, M. C. and C. Thaw (1985). "Thyrotropin-releasing hormone (TRH) stimulates biphasic elevation of cytoplasmic free calcium in GH3 cells. Further evidence that TRH mobilizes cellular and extracellular Ca²⁺." Endocrinology **116**(2): 591-596.

Ghyoot, M., J. L. S. Cobb and M. C. Thorndyke (1994). "Localisation of neuropeptides in the nervous system of the brittle star *Ophiura ophiura*." Phil. Trans. R Soc. London **346**: 434-444.

Gimpl, G. and F. Fahrenholz (2001). "The oxytocin receptor system: structure, function, and regulation." Physiol Rev **81**(2): 629-683.

Goldberg, J. I., R. Garofalo, C. J. Price and J. P. Chang (1993). "Presence and biological activity of a GnRH-like factor in the nervous system of *Helisoma trevolvis*." J. Comp. Neurol. **336**: 571-582.

Goodson, J. L. and A. H. Bass (2000). "Forebrain peptides modulate sexually polymorphic vocal circuitry." Nature **403**: 769-772.

Gorbman, A. and S. A. Sower (2003). "Evolution of the role of GnRH in animals (Metazoan) biology." Gen. Comp. Endocrinol. **134**: 207-213.

Gorbman, A., A. Whiteley and S. Kavanaugh (2003). "Pheromonal stimulation of spawning release of gametes by gonadotropin releasing hormone in the chiton, *Mopalia* sp." Gen Comp Endocrinol **131**(1): 62-65.

Gozdowska, M., A. Kleszczynska, E. Sokolowska and E. Kulczykowska (2006). "Arginine vasotocin (AVT) and isotocin (IT) in fish brain: diurnal and seasonal variations." Comp Biochem Physiol B Biochem Mol Biol **143**(3): 330-334.

Green, E. D. (2001). "Strategies for systematic sequencing of complex genomes." Nat. Rev. Genet. **2**: 573-583.

Greenberg, M. J., S. D. Painter, K. E. Doble, G. T. Naggle, D. A. Price and H. K. Lehman (1983). "The molluscan neurosecretory peptide FMRFamide: comparative pharmacology and relationship to the enkephalins." Fed. Proc. **42**: 82-86.

Greenberg, M. J. and D. A. Price (1983). "Invertebrate neuropeptides: native and naturalized." Annu. Rev. Physiol **45**: 271-288.

Grimmelikhuijzen, C. J., F. Hauser, K. K. Eriksen and M. Williamson (1999). "Invertebrate neurohormones and their receptors." Results Probl Cell Differ **26**: 339-362.

Gudermann, T., T. Schoneberg and G. Schultz (1997). "Functional and structural complexity of signal transduction via G-protein-coupled receptors." Annu Rev Neurosci **20**: 399-427.

Gutman, G. A., K. G. Chandy, et al. (2005). "International Union of Pharmacology. LIII. Nomenclature and Molecular Relationships of Voltage-Gated Potassium Channels." Pharmac. Rev. **57**: 473-508.

Haken, H. (1983). Synergetics: An Introduction. Berlin, Springer-Verlag.

Hartenstein, V. (2006). "The neuroendocrine system of invertebrates: a developmental and evolutionary perspective." J Endocrinol **190**(3): 555-570.

Hayflick, J. S., J. P. Adelman and P. H. Seeburg (1989). "The complete nucleotide sequence of the human gonadotropin-releasing hormone gene." Nucleic Acids Res **17**(15): 6403-6404.

Herman, J. P., H. Figueiredo, N. K. Mueller, Y. Ulrich-Lai, M. M. Ostrander, D. C. Choi and W. E. Cullinan (2003). "Central mechanisms of stress integration: hierarchical circuitry controlling hypothalamo-pituitary-adrenocortical responsiveness." Front Neuroendocrinol **24**(3): 151-180.

Hewes, R. S. and P. H. Taghert (2001). "Neuropeptides and neuropeptide receptors in the *Drosophila melanogaster* genome." Genome Res. **11**(6): 1126-1142.

Higgins, D. G. and P. M. Sharp (1988). "CLUSTAL: a package for performing multiple sequence alignment on a microcomputer." Gene **73**: 237-244.

Hille, B. (2001). Ion Channels of Excitable Membranes, Sinauer.

Hiraoka, D., S. Hori-Oshima, T. Fukuhara, K. Tachibana, E. Okumura and T. Kishimoto (2004). "PDK1 is required for the hormonal signaling pathway leading to meiotic resumption in starfish oocytes." Dev Biol **276**(2): 330-336.

Hokfelt, T., U. Arvidsson, S. Cullheim, D. Millhorn, A. P. Nicholas, V. Pieribone, K. Serogy and B. Ulfhake (2000). "Multiple messengers in

descending serotonin neurons: localization and functional implications." J Chem Neuroanat **18**(1-2): 75-86.

Holland, P. W., J. Garcia-Fernandez, N. A. Williams and A. Sidow (1994). "Gene duplications and the origins of vertebrate development." Dev Suppl: 125-133.

Howard-Ashby, M., S. C. Materna, C. T. Brown, Q. Tu, P. Oliveri, R. A. Cameron and E. H. Davidson (2006). "High regulatory gene use in sea urchin embryogenesis: Implications for bilaterian development and evolution." Dev Biol **300**(1): 27-34.

Hoyle, C. H. (1999). "Neuropeptide families and their receptors: evolutionary perspectives." Brain Res **848**(1-2): 1-25.

Hsieh, Y. S., M. C. Fann, W. C. Wan and P. S. Wang (1990). "The presence of a thyrotropin-releasing hormone-like factor in the grass shrimp (black tiger prawn, *Penaeus monodon*)." Chin. J. Physiol. **33**(2): 179-190.

Hsu, S. Y., K. Nakabayashi and A. Bhalla (2002). "Evolution of glycoprotein hormone subunit genes in bilateral metazoa: identification of two novel human glycoprotein hormone subunit family genes, GPA2 and GPB5." Mol. Endocrinol. **16**(7): 1538-1551.

Hummon, A. B., T. A. Richmond, et al. (2006). "From the Genome to the Proteome: Uncovering Peptides in the *Apis* Brain." Science **314**: 647-649.

Husson, S. J., E. Clynen, G. Baggerman, A. De Loof and L. Schoofs (2005). "Discovering neuropeptides in *Caenorhabditis elegans* by two dimensional liquid chromatography and mass spectrometry." Biochem. Biophys. Res. Commun. **335**: 76-86.

Husson, S. J., I. Mertens, T. Janssen, M. Lindemans and L. Schoofs (2007). "Neuropeptidergic signaling in the nematode *Caenorhabditis elegans*." Progress in Neurobiology **82**: 33-55.

Hyman, L. (1955). Echinodermata. New York and London, McGraw-Hill.

Inoue, M., R. Birenheide, O. Koizumi, Y. Kobayakawa, Y. Muneoka and T. Motokawa (1999). "Localization of the neuropeptide NGIYWamide in the holothurian nervous system and its effect on muscular contraction." Proc. R. Soc. Lond. B **266**: 993-1000.

Israel, S., E. Lerer, et al. (2008). "Molecular genetic studies of the arginine vasopressin 1a receptor (AVPR1a) and the oxytocin receptor (OXTR) in human behaviour: from autism to altruism with some notes in between." Prog Brain Res **170**: 435-449.

Iwakoshi-Ukena, E., K. Ukena, K. Takuwa-Kuroda, A. Kanda, K. Tsutsui and H. Minakata (2004). "Expression and distribution of octopus gonadotropin-releasing hormone in the central nervous system and peripheral organs of the octopus (*Octopus vulgaris*) by in situ

hybridization and immunohistochemistry." J. Comp. Neurol. **477**: 310-323.

Iwakoshi, E., M. Ohtani, T. Takahashi, Y. Moneoka, T. Ikeda, T. Fujita, H. Minakata and K. Nomoto (1995). Comparative Aspects of Structure and Action of Bioactive Peptides Isolated from the Seacucumber *Stichopus japonicus*. Peptide Chemistry 1994. M. Ohno. Osaka, Protein Research Foundation: 261-264.

Iwakoshi, E., K. Takuwa-Kuroda, T. Fujisawa, M. Hisada, K. Ukena, K. Tsutsui and H. Minakata (2002). "Isolation and characterization of a GnRH-like peptide from *Octopus vulgaris*." Biochem. Biophys. Res. Commun. **291**: 1187-1193.

Johnson, E. C., O. T. Shafer, J. P. Trigg, J. Park, J. A. Schooley and P. H. Taghert (2005). "A novel diuretic hormone receptor in *Drosophila*: evidence for conservation of CGRP signaling." J. Exp. Biol. **208**: 1239-1246.

Jojart, I., F. Joo, L. Siklos and F. A. Laszlo (1984). "Immunoelectronhistochemical evidence for innervation of brain microvessels by vasopressin-immunoreactive neurons in the rat." Neurosci Lett **51**(2): 259-264.

Kamesh, N., G. K. Aradhyam and N. Manoj (2008). "The repertoire of G protein-coupled receptors in the sea squirt *Ciona intestinalis*." BMC Evol Biol **8**: 129.

Kanatani, H. (1979). Hormones in echinoderms. Hormones and Evolution. E. J. W. Barrington. London, Academic Press: 273-207.

Kanatani, H. (1983). "Nature and action of the mediators inducing maturation of the starfish oocyte." Ciba Found Symp **98**: 159-170.

Kanatani, H., S. Ikegami, H. Shirai, H. Oide and S. Tamura (1971). "Purification of gonad-stimulating substance obtained from radial nerves of the starfish *Asterias amurensis*." Development, Growth and Differentiation **13**(3): 151-164.

Kato, S., S. Tsurumaru, M. Taga, T. Yamane, Y. Shibata, K. Ohno, A. Fujiwara, K. Yamano and M. Yoshikuni (2009). "Neuronal peptides induce oocyte maturation and gamete spawning of sea cucumber, *Apostichopus japonicus*." Dev Biol **326**(1): 169-176.

Kawada, T., T. Sekiguchi, Y. Itoh, M. Ogasawara and H. Satake (2008). "Characterization of a novel vasopressin/oxytocin superfamily peptide and its receptor from an ascidian, *Ciona intestinalis*." Peptides **29**: 1672-1678.

Kawaguti, S. (1964). "Electron microscopic structures of the podial wall of an echinoid with special references to the nerve plexus and the muscle." Biol. J. Okyama Univ. **10**: 1-12.

- Keller, R. (1992). "Crustacean neuropeptides: Structures, functions and comparative aspects." Experientia **48**: 439-448.
- Kelso, J. A. S. (1995). Dynamic patterns : the self-organization of brain and behavior. Cambridge, Mass ; London, MIT Press.
- Kerkut, G. A. (1954). "The mechanisms of coordination of the starfish tube feet." Behaviour **6**: 206-232.
- Kimura, S., Y. Sugita, I. Kanazawa, A. Saito and K. Goto (1987). "Isolation and amino acid sequence of calcitonin gene related peptide from porcine spinal cord." Neuropeptides **9**(1): 75-82.
- King, J. A. and R. P. Millar (1982). "Structure of chicken hypothalamic luteinizing hormone-releasing hormone. II. Isolation and characterization." J Biol Chem **257**(18): 10729-10732.
- Kolakowski, L. F. (1994). "GCRDb: a G-protein-coupled receptor database." Recept. Channels **2**: 1-7.
- Koob, T. J., M. M. Koob-Emunds and J. A. Trotter (1999). "Cell-derived stiffening and plasticizing factors in sea cucumber (*Cucumaria frondosa*) dermis." J Exp Biol **202 (Pt 17)**: 2291-2301.
- Kosfeld, M., M. Heinrichs, P. J. Zak, U. Fischbacher and E. Fehr (2005). "Oxytocin increases trust in humans." Nature **435**(7042): 673-676.
- Kuczer, M., G. Rosinski and D. Konopinska (2007). "Insect gonadotropic peptide hormones: some recent developments." J Pept Sci **13**(1): 16-26.
- Lafont, A.-G., S. Dufour and M. Fouchereau-Peron (2007). "Evolution of the CT/CGRP family: Comparative study with new data from models of teleosts, the eel, and cephalopod molluscs, the cuttlefish and the nautilus." Gen. Comp. Endocrinol. **153**: 155-169.
- Lapraz, F., E. Rottinger, et al. (2006). "RTK and TGF-beta signaling pathways genes in the sea urchin genome." Dev Biol **300**(1): 132-152.
- Legros, J. J. and V. Geenen (1996). "Neurophysins in central diabetes insipidus." Horm Res **45**(3-5): 182-186.
- Li, C., K. Kim and L. S. Nelson (1999a). "FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*." Brain Res. **848**: 26-35.
- Li, L., R. W. Garden, E. V. Romanova and J. V. Sweedler (1999b). "In situ sequencing of peptides from biological tissues and single cells using MALDI-PSD/CID analysis." Anal Chem **71**(24): 5451-5458.
- Li, L., W. P. Kelley, C. P. Billimoria, A. E. Christie, S. R. Pulver, J. V. Sweedler and E. Marder (2003). "Mass spectrometric investigation of the

neuropeptide complement and release in the pericardial organs of the crab, *Cancer borealis*." J Neurochem **87**(3): 642-656.

Lichtarge, O., H. R. Bourne and F. E. Cohen (1996). "An evolutionary trace method defines binding surfaces common to protein families." J Mol Biol **257**(2): 342-358.

Light, A. and V. Du Vigneaud (1958). "On the nature of oxytocin and vasopressin from human pituitary." Proc Soc Exp Biol Med **98**(4): 692-696.

Lindemans, M., F. Liu, T. Janssen, S. J. Husson, I. Mertens, G. Gade and L. Schoofs (2009). "Adipokinetic hormone signaling through the gonadotropin-releasing hormone receptor modulates egg-laying in *Caenorhabditis elegans*." Proc Natl Acad Sci U S A **106**(5): 1642-1647.

Littlewood, D. T. J. and A. B. Smith (1995). "A combined morphological and molecular phylogeny for sea urchins (Echinoidea: Echinodermata)." Phil. Trans. R. Soc. Lond. B **347**: 213-234.

Liu, F., G. Baggerman, E. D'Hertog, P. Verleyen, L. Schoofs and G. Wets (2006). "In silico identification of new secretory peptide genes in *Drosophila melanogaster*." Mol. Cell. Proteomics **5**(3): 510-522.

Lloyd, P. E. and C. M. Connolly (1989). "Sequence of Pedal Peptide: A Novel Neuropeptide from the Central Nervous System of *Aplysia*." Journal of Neuroscience **9**(1): 312-317.

Lovejoy, D. A., W. H. Fischer, S. Ngamvongchon, A. G. Craig, C. S. Nahorniak, R. E. Peter, J. E. Rivier and N. M. Sherwood (1992). "Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain provides insight into GnRH evolution." Proc Natl Acad Sci U S A **89**(14): 6373-6377.

Lundin, L. G. (1993). "Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse." Genomics **16**(1): 1-19.

Luo, C. W., E. M. Dewey, S. Sudo, J. Ewer, S. Y. Hsu, H. W. Honegger and A. J. Hsueh (2005). "Bursicon, the insect cuticle-hardening hormone, is a heterodimeric cystine knot protein that activates G protein-coupled receptor LGR2." Proc Natl Acad Sci U S A **102**(8): 2820-2825.

Mackie, G. O. (1995). "On the 'visceral nervous system' of *Ciona*." J. Mar. Biol. Assoc. UK **75**: 141-151.

Mackie, G. O., R. M. Marx and R. W. Meech (2003). "Central circuitry in the jellyfish *Aglantha digitale* IV. Pathways coordinating feeding behaviour." J Exp Biol **206**(Pt 14): 2487-2505.

- Martin, V. J. (1992). "Characterization of a RFamide-positive subset of ganglionic cells in the hydrozoan planular nerve net." Cell Tissue Res **269**(3): 431-438.
- Maruyama, K. Y. (1985). "Holothurian oocyte maturation induced by radial nerve." Biol. Bull. **168**: 249-262.
- McFarlane, I. D., P. A. Anderson and C. J. Grimmelikhuijzen (1991). "Effects of three anthozoan neuropeptides, Antho-RWamide I, Antho-RWamide II and Antho-RFamide, on slow muscles from sea anemones." J. Exp. Biol. **156**: 419-431.
- Melarange, R. and M. R. Elphick (2003). "Comparative analysis of nitric oxide and SALMFamide neuropeptides as general muscle relaxants in starfish." J Exp Biol **206**(Pt 5): 893-899.
- Melarange, R., D. J. Potton, M. C. Thorndyke and M. R. Elphick (1999). "SALMFamide neuropeptides cause relaxation and eversion of the cardiac stomach in starfish." Proc. R. Soc. Lond. B **266**: 1785-1789.
- Mendes, E. G., L. Abbud and A. A. Ancona-Lopez (1970). "Pharmacological studies on the invertebrate non-striated muscles- I. The response to drugs." Comp. gen. Pharmac. **1**: 11-22.
- Migita, M., E. Mizukami and Y. P. Gunji (2005). "Flexibility in starfish behavior by multi-layered mechanism of self-organization." Biosystems **82**(2): 107-115.
- Mita, M., H. Oka, M. C. Thorndyke, Y. Shibata, M. Yoshikuni and Y. Nagahama (2004). "Inhibitory effect of a SALMFamide neuropeptide on secretion of gonad-stimulating substance from radial nerves in the starfish *Asterina pectinifera*." Zoolog Sci **21**(3): 299-303.
- Mita, M., M. Yoshikuni, K. Ohno, Y. Shibata, B. Paul-Prasanth, S. Pitchayawasin, M. Isobe and Y. Nagahama (2009). "A relaxin-like peptide purified from radial nerves induces oocyte maturation and ovulation in the starfish, *Asterina pectinifera*." Proc Natl Acad Sci U S A **106**(23): 9507-9512.
- Mladenov, P. V., S. Igdoura, S. Asotra and R. D. Burke (1989). "Purification and partial characterization of an autotomy-promoting factor from the seastar *Pycnopodia helianthoides*." Biol. Bull. **176**: 169-175.
- Mohr, E., M. Hillers, R. Ivell, I. D. Haulica and D. Richter (1985). "Expression of the vasopressin and oxytocin genes in human hypothalami." FEBS Lett. **193**(1): 12-16.
- Moroz, L. L., J. R. Edwards, et al. (2006). "Neuronal Transcriptome of *Aplysia*: Neuronal Compartments and Circuitry." Cell **127**: 1453-1467.
- Morris, B. J. and H. M. Johnston (1995). "A role for hippocampal opioids in long-term functional plasticity." Trends Neurosci **18**(8): 350-355.

Moss, C., R. D. Burke and M. C. Thorndyke (1994). "Immunolocalization of the neuropeptide S1 and serotonin in the larvae of the starfish *Pisaster ochraceus* and *Asterias rubens*." 1994 J. Mar Biol. Assoc. UK(74): 61-71.

Mousley, A., N. J. Marks, D. W. Halton, T. G. Geary, D. P. Thompson and A. G. Maule (2004). "Arthropod FMRamide-related peptides modulate muscle activity in helminths." Int J Parasitol **34**(6): 755-768.

Nachman, R. J., G. M. Coast, S. E. Tichy, D. H. Russell, J. A. Miller and R. Predel (2002). "Occurrence of insect kinins in the flesh fly, stable fly and horn fly: mass spectrometric identification from single nerves and diuretic activity." Peptides **23**: 1885-1894.

Nassel, D. R. (2002). "Neuropeptides in the nervous system of *Drosophila* and other insects: multiple roles as neuromodulators and neurohormones." Prog. Neurobiol. **68**(1): 1-84.

Nathoo, A. N., R. A. Moeller, B. A. Westlund and A. C. Hart (2001). "Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species." Proc. Natl. Acad. Sci. USA **98**: 14000-14005.

Newman, S. J., M. R. Elphick and M. C. Thorndyke (1995). "Tissue distribution of the SALMamide neuropeptides S1 and S2 in the starfish *Asterias rubens* using novel monoclonal and polyclonal antibodies. I. Nervous and locomotory systems." Proc Biol Sci **261**(1360): 139-145.

Nichols, R., W. G. Bendena and S. S. Tobe (2002). "Myotropic peptides in *Drosophila melanogaster* and the genes that encode them." J Neurogenet **16**(1): 1-28.

Nicolas, P., M. Camier, M. Lauber, M. J. Masse, J. Mohring and P. Cohen (1980). "Immunological identification of high molecular weight forms common to bovine neurophysin and vasopressin." Proc Natl Acad Sci U.S.A. **77**(5): 2587-2591.

Nicolas, P., J. Wolff, M. Camier, C. Di Bello and P. Cohen (1978). "Importance of neurophysin dimer and of tyrosine-49 in the binding of neurohypophyseal peptides." J. Biol. Chem. **253**: 2633-2639.

Nordström, K. J., R. Fredriksson and H. B. Schioth (2008). "The amphioxus (*Branchiostoma floridae*) genome contains a highly diversified set of G protein-coupled receptors." BMC Evol Biol **8**: 9.

O'Donnell, M. E., V. Duong, J. Suvatne, S. Foroutan and D. M. Johnson (2005). "Arginine vasopressin stimulation of cerebral microvascular endothelial cell Na-K-Cl cotransporter activity is V1 receptor and [Ca] dependent." Am J Physiol Cell Physiol **289**(2): C283-292.

O'Shea, M. and M. Schaffer (1985). "Neuropeptide function: the invertebrate contribution." Annu. Rev. Neurosci. **8**: 171-198.

- Ohtani, M., E. Iwakoshi, Y. Muneoka, H. Minakata and K. Nomoto (1999). Isolation and characterization of bioactive peptides from the sea cucumber, *Stichopus japonicus*. Peptide Science - Present and Future. Y. Shimonishi: 419-420.
- Okada, Y., K. S. Kiyotugu and M. Yanagihara (1984). "Synchronized rhythmic contractions among five gonadal lobes in the shedding sea urchins: coordinative function of the aboral nerve ring." Biol. Bull. **166**: 228-236.
- Okumura, E., T. Fukuhara, H. Yoshida, S. Hanada Si, R. Kozutsumi, M. Mori, K. Tachibana and T. Kishimoto (2002). "Akt inhibits Myt1 in the signalling pathway that leads to meiotic G2/M-phase transition." Nat Cell Biol **4**(2): 111-116.
- Olson, G. A., R. D. Olson and A. J. Kastin (1995). "Endogenous opiates: 1994." Peptides **16**(8): 1517-1555.
- Ota, Y., H. Ando, H. Ueda and A. Urano (1999). "Differences in seasonal expression of neurohypophysial hormone genes in ordinary and precocious male masu salmon." Gen Comp Endocrinol **116**(1): 40-48.
- Oumi, T., K. Ukena, O. Matsushima, T. Ikeda, T. Fujita, H. Minakata and K. Nomoto (1994). "Annetocin: an oxytocin-related peptide isolated from the earthworm, *Eisenia foetida*." Biochem. Biophys. Res. Commun. **198**: 393-399.
- Palmer, R. M., A. G. Ferrige and S. Moncada (1987). "Nitric oxide release accounts for the biological activity of endothelium-derived relaxing factor." Nature **327**(6122): 524-526.
- Panopoulou, G., S. Hennig, D. Groth, A. Krause, A. J. Poustka, R. Herwig, M. Vingron and H. Lehrach (2003). "New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes." Genome Res **13**(6A): 1056-1066.
- Panula, P., E. Kalso, M. Nieminen, V. K. Kontinen, A. Brandt and A. Pertovaara (1999). "Neuropeptide FF and modulation of pain." Brain Res. **848**: 191-196.
- Patel, Y. C. (1999). "Somatostatin and its receptor family." Front Neuroendocrinol **20**(3): 157-198.
- Pazos, A. J. and M. Mathieu (1999). "Effects of five natural gonadotropin-releasing hormones on cell suspensions of marine bivalve gonad: stimulation of gonial DNA synthesis." Gen Comp Endocrinol **113**(1): 112-120.
- Pearson, W. L. and P. E. Lloyd (1989). "Immunocytochemical localization of pedal peptide in the central nervous system and periphery of *Aplysia*." J Neurosci **9**(1): 318-325.

- Pentreath, V. W. and G. A. Cottrell (1968). "Acetylcholine and cholinesterase in the radial nerve of *Asterias rubens*." Comp Biochem Physiol **27**(3): 775-785.
- Pesole, G., S. Liuni, G. Grillo and C. Saccone (1997). "Structural and compositional features of untranslated regions of eukaryotic mRNAs." Gene **205**(1-2): 95-102.
- Peters, B. H. (1981). The neurobiology of spines and pedicellariae from *Echinus esculentus* L. University of London.
- Pinelli, C., B. D'Aniello, P. Sordino, D. L. Meyer, M. Fiorentino and R. K. Rastogi (2000). "Comparative immunocytochemical study of FMRamide neuronal system in the brain of *Danio rerio* and *Acipenser ruthenus* during development." Brain Res Dev Brain Res **119**(2): 195-208.
- Poels, J., R. T. Birse, R. J. Nachman, J. Fichna, A. Janecka, J. Vanden Broeck and D. R. Nassel (2009). "Characterization and distribution of NKD, a receptor for *Drosophila* tachykinin-related peptide 6." Peptides **30**(3): 545-556.
- Powell, J. F., S. M. Reska-Skinner, M. O. Prakash, W. H. Fischer, M. Park, J. E. Rivier, A. G. Craig, G. O. Mackie and N. M. Sherwood (1996). "Two new forms of gonadotropin-releasing hormone in a protochordate and the evolutionary implications." Proc Natl Acad Sci U S A **93**(19): 10461-10464.
- Pradayrol, L., J. A. Chayvialle, M. Carlquist and V. Mutt (1978). "Isolation of a porcine intestinal peptide with C-terminal somatostatin." Biochem Biophys Res Commun **85**(2): 701-708.
- Price, D. A. and M. J. Greenberg (1977). "Structure of a molluscan cardioexcitatory neuropeptide." Science **197**(4304): 670-671.
- Proux, J. P., C. A. Miller, J. P. Li, R. L. Carney, A. Girardie, M. Delaage and D. A. Schooley (1987). "Identification of an arginine vasopressin-like diuretic hormone from *Locusta migratoria*." Biochem. Biophys. Res. Commun. **149**(1): 180-186.
- Rastogi, R. K., M. M. D. Fiore, A. D'Aniello, L. Iela and M. Fiorentino (2002). "GRH in the invertebrates: an overview." Prog. Brain Res. **141**: 19-29.
- Reese, E. S. (1966). The complex behavior of echinoderms. Physiology of Echinodermata. New York, Interscience Publishers: 157-218.
- Reich, G. (1992). "A new peptide of the oxytocin/vasopressin family isolated from nerves of the cephalopod *Octopus vulgaris*." Neurosci. Lett. **134**(2): 191-194.

- Riehle, M. A., S. F. Garczynski, J. W. Crim, C. A. Hill and M. R. Brown (2002). "Neuropeptides and Peptide Hormones in *Anopheles gambiae*." Science **298**: 172-175.
- Robb, S., L. C. Packman and P. D. Evans (1989). "Isolation, primary structure and bioactivity of schistoFLRF-amide, a FMRF-amide-like neuropeptide from the locust, *Schistocerca gregaria*." Biochem. Biophys. Res. Commun. **160**: 850-856.
- Rogers, C., V. Reale, K. Kim, H. Chatwin, C. Li, P. Evans and M. de Bono (2003). "Inhibition of *Caenorhabditis elegans* social feeding by FMRFamide-related peptide activation of NPR-1." Nat Neurosci **6**(11): 1178-1185.
- Roller, L., N. Yamanaka, K. Watanabe, I. Daubnerova, D. Zitnan, H. Kataoka and Y. Tanaka (2008). "The unique evolution of neuropeptide genes in the silkworm *Bombyx mori*." Insect Biochem Mol Biol **38**(12): 1147-1157.
- Saha, A. K., M. Tamori, M. Inoue, Y. Nakajima and T. Motokawa (2006). "NGIWyamide-induced Contraction of Tube Feet and Distribution of NGIWyamide-like Immunoreactivity in Nerves of the Starfish *Asterina pectinifera*." Zool. Sci. **23**: 627-632.
- Sakarya, O., K. A. Armstrong, M. Adamska, M. Adamski, I. F. Wang, B. Tidor, B. M. Degnan, T. H. Oakley and K. S. Kosik (2007). "A post-synaptic scaffold at the origin of the animal kingdom." PLoS One **2**(6): e506.
- Sakurai, T., A. Amemiya, et al. (1998). "Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior." Cell **92**(4): 573-585.
- Samson, W. K., M. M. Taylor and J. R. Baker (2003). "Prolactin-releasing peptides." Regul. Peptides **114**: 1-5.
- Santos, R., D. Haesaerts, M. Jangoux and P. Flammang (2005). "The tube feet of sea urchins and sea stars contain functionally different mutable collagenous tissues." J. Exp. Biol. **208**(12): 2277-2288.
- Sasayama, Y., A. Katoh, C. Oguro, A. Kambegawa and H. Yoshizawa (1991). "Cells showing immunoreactivity for calcitonin or calcitonin gene-related peptide (CGRP) in the central nervous system of some invertebrates." Gen. Comp. Endocrinol. **83**: 565-585.
- Schmutzler, C., D. Darmer, D. Diekhoff and C. J. Grimmlikhuijzen (1992). "Identification of a novel type of processing sites in the precursor for the sea anemone neuropeptide Antho-RFamide (<Glu-Gly-Arg-Phe-NH₂) from *Anthopleura elegantissima*." J. Biol. Chem. **267**(31): 22534-22541.

- Shirai, H. (1986). Gonad-stimulating and maturation inducing substance. Methods in Cell Biology. K. Wilson, Academic Press Inc. **99**: 73-88.
- Shirai, H., P. Bulet, N. Kondo, M. Isobe, K. Imai, T. Goto and I. Kubota (1987). "Gonad-stimulating substance of starfish." Gen. Comp. Endocrinol. **66**: 50.
- Shirai, H., H. Kanatani and S. Taguchi (1972). "1-Methyladenine Biosynthesis in Starfish Ovary: Action of Gonad-Stimulating Hormone in Methylation." Science **175**(4028): 1366-1368.
- Shul'gina, G. I. (2006). "[Behavioral training of starfish *Asterias rubens*.]" Zh Evol Biokhim Fiziol **42**(2): 130-133.
- Silver, M. R., H. Kawauchi, M. Nozaki and S. A. Sower (2004). "Cloning and analysis of the lamprey GnRH-III cDNA from eight species of lamprey representing the three families of Petromyzoniformes." Gen Comp Endocrinol **139**(1): 85-94.
- Singh, G. and A. P. Davenport (2006). "Neuropeptide B and W: neurotransmitters in an emerging G-protein-coupled receptor system." Br J Pharmacol **148**(8): 1033-1041.
- Siviter, R. J., G. M. Coast, A. M. Winther, R. J. Nachman, C. A. Taylor, A. D. Shirras, D. Coates, R. E. Isaac and D. R. Nassel (2000). "Expression and functional characterization of a *Drosophila* neuropeptide precursor with homology to mammalian preprotachykinin A." J Biol Chem **275**(30): 23273-23280.
- Smith, A. B. (1988). "Phylogenetic Relationship, Divergence Times, and Rates of Molecular Evolution for Camarodont Sea Urchins." Mol. Biol. Evol. **5**(4): 345-365.
- Smith, G. N. and M. J. Greenberg (1973). "Chemical control of the evisceration process in *Thyone briareus*." Biol. Bull. **144**: 421-436.
- Smith, J. E. (1945). "The role of the nervous system in some activities of starfishes." Biol. Rev. **20**: 29-43.
- Smith, J. E. (1965). Echinodermata. Structure and Function in the Nervous Systems of Invertebrates. T. H. Bullock and G. A. Horridge. London, W. H. Freeman & Co: 1519-1558.
- Sodergren, E., G. M. Weinstock, et al. (2006). "The genome of the sea urchin *Strongylocentrotus purpuratus*." Science **314**(5801): 941-952.
- Southward, E. C. and A. C. Campbell (2006). Echinoderms : keys and notes for the identification of British species. Shrewsbury, FSC Publications.
- Stafflinger, E., K. K. Hansen, F. Hauser, M. Schneider, G. Cazzamali, M. Williamson and C. J. Grimmelikhuijzen (2008). "Cloning and

identification of an oxytocin/vasopressin-like receptor and its ligand from insects." Proc. Natl. Acad. Sci. **105**(9): 3262-3267.

Starratt, A. N. and B. E. Brown (1975). "Structure of the pentapeptide proctolin, a proposed neurotransmitter in insects." Life Sci **17**(8): 1253-1256.

Stay, B. and S. S. Tobe (2007). "The role of allatostatins in juvenile hormone synthesis in insects and crustaceans." Annu Rev Entomol **52**: 277-299.

Strand, F. L. (1999). Neuropeptides : regulators of physiological processes. Cambridge, Mass. ; London, MIT.

Suzuki, K., R. L. Gamble and S. A. Sower (2000). "Multiple transcripts encoding lamprey gonadotropin-releasing hormone-I precursors." J Mol Endocrinol **24**(3): 365-376.

Suzuki, M., K. Kubokawa, H. Nagasawa and A. Urano (1995). "Sequence analysis of vasotocin cDNAs of the lamprey, *Lampetra japonica*, and the hagfish, *Eptatretus burgeri*: evolution of cyclostome vasotocin precursors." J. Mol. Endocrinol. **14**(1): 67-77.

Syed, N., C. A. Martens and W. H. Hsu (2007). "Arginine vasopressin increases glutamate release and intracellular Ca²⁺ concentration in hippocampal and cortical astrocytes through two distinct receptors." J Neurochem **103**(1): 229-237.

Tadenuma, H., K. Takahashi, K. Chiba, M. Hoshi and T. Katada (1992). "Properties of 1 methyladenine receptors in starfish oocyte membranes: involvement of pertussis toxin sensitive GTP binding protein in the receptor mediated signal transduction." **186**: 114-121.

Taghert, P. H. and O. T. Shafer (2006). "Mechanisms of clock output in the *Drosophila* circadian pacemaker system." J Biol Rhythms **21**(6): 445-457.

Taghert, P. H. and J. A. Veenstra (2003). "Drosophila neuropeptide signaling." Adv Genet **49**: 1-65.

Takayasu, S., T. Sakurai, et al. (2006). "A neuropeptide ligand of the G protein-coupled receptor GPR103 regulates feeding, behavioural arousal, and blood pressure in mice." Proc. Natl. Acad. Sci. USA **103**(19): 7438-7443.

Tamori, M., A. K. Saha, A. Matsuno, S. C. Noskor, O. Koizumi, Y. Kobayakawa, Y. Nakajima and T. Motokawa (2007). "Stichopin-containing nerves and secretory cells specific to connective tissue of the sea cucumber." Proc Biol Sci **274**: 2279-2285.

Taylor, M. H. (1986). "Environmental and endocrine influences on reproduction of *Fundulus heteroclitus*." Am. Zool. **26**: 159-171.

- Tello, J. A., J. E. Rivier and N. M. Sherwood (2005). "Tunicate gonadotropin-releasing hormone (GnRH) peptides selectively activate *Ciona intestinalis* GnRH receptors and the green monkey type II GnRH receptor." Endocrinology **146**(9): 4061-4073.
- Tessmar-Raible, K., F. Raible, F. Christodoulou, K. Guy, M. Rembold, H. Hausen and D. Arendt (2007). "Conserved sensory-neurosecretory cell types in annelid and fish forebrain: insights into hypothalamus evolution." Cell **129**(7): 1389-1400.
- Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins (1997). "The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools." Nucleic Acids Res. **25**: 4876-4882.
- Thompson, K. S. and J. P. Bacon (1991). "The vasopressin-like immunoreactive (VPLI) neurons of the locust, *Locust migratoria*. II. Physiology." J. Comp. Physiol. **168**: 619-630.
- Thorndyke, M. C., W. C. Chen, P. W. Beesley and M. Patruno (2001). "Molecular approach to echinoderm regeneration." Microsc Res Tech **55**(6): 474-485.
- Thorndyke, M. C. and G. J. Goldsworthy (1988). Neurohormones in invertebrates. Cambridge, Cambridge University Press.
- Tipper, J. P., G. Lyons-Levy, M. A. L. Atkinson and J. A. Trotter (2003). "Purification, characterization and cloning of tensilin, the collagen-fibril binding and tissue-stiffening factor from *Cucumaria frondosa* dermis." Matrix Biol. **21**: 625-635.
- Troemel, E. R., J. H. Chou, N. D. Dwyer, H. A. Colbert and C. I. Bargmann (1995). "Divergent seven transmembrane receptors are candidate chemosensory receptors in *C. elegans*." Cell **83**(2): 207-218.
- Tsai, P.-S. (2006). "Gonadotropin-releasing hormone in invertebrates: Structure, function and evolution." Gen. Comp. Endocrinol. **148**: 48-53.
- Tsai, P. S., T. A. Maldonado and J. B. Lunden (2003). "Localization of gonadotropin-releasing hormone in the central nervous system and a peripheral chemosensory organ of *Aplysia californica*." Gen Comp Endocrinol **130**(1): 20-28.
- Twarog, B. W. (1954). "Responses of a molluscan smooth muscle to acetylcholine and 5-hydroxytryptamine." J. Cell Physiol. **44**: 141-163.
- Ukena, K., E. Iwakoshi-Ukena and A. Hikosaka (2008). "Unique Form and Osmoregulatory Function of a Neurohypophysial Hormone in a Urochordate." Endocrinology **149**(10): 5254-5261.

- Urano, A., S. Hyodo and M. Suzuki (1992). "Molecular evolution of neurohypophysial hormone precursors." Prog. Brain Res. **92**: 39-46.
- Valentinčič, T. (1983). Innate and learned responses to external stimuli in asteroids. Echinodem Studies. M. Jangoux and J. M. Lawrence. **1**: 111-137.
- van den Hooff, P., M. A. Seger and J. P. Burbach (1990). "The carboxyterminal glycopeptide of propressophysin potentiates excitatory transmission in the rat lateral septum." Neuroscience **37**(3): 647-653.
- van den Hooff, P., I. J. Urban and D. de Wied (1989). "Vasopressin maintains long-term potentiation in rat lateral septum slices." Brain Res **505**(2): 181-186.
- Van Kesteren, R. E., A. B. Smit, R. P. De Lange, K. S. Kits, F. A. Van Golen, R. C. Van Der Schors, N. D. De With, J. F. Burke and W. P. Geraerts (1995). "Structural and functional evolution of the vasopressin/oxytocin superfamily: vasopressin-related conopressin is the only member present in *Lymnaea*, and is involved in the control of sexual behavior." J Neurosci **15**(9): 5989-5998.
- Van Kesteren, R. E., A. B. Smit, R. W. Dirks, N. D. De With, W. P. Geraerts and J. Joosse (1992). "Evolution of the vasopressin/oxytocin superfamily: characterization of a cDNA encoding a vasopressin-related precursor, preproconopressin, from the mollusc *Lymnaea stagnalis*." Proc Natl Acad Sci U.S.A. **89**(10): 4593-4597.
- Vanden Broeck, J. (2001a). "Insect G protein-coupled receptors and signal transduction." Arch Insect Biochem Physiol **48**(1): 1-12.
- Vanden Broeck, J. (2001b). "Neuropeptides and their precursors in the fruitfly, *Drosophila melanogaster*." Peptides **22**: 241-254.
- Vanden Broeck, J., D. Velaert, W. G. Bendena, S. S. Tobe and A. De Loof (1996). "Molecular cloning of the precursor cDNA for schistostatins, locust allatostatin-like peptides with myoinhibiting properties." Mol Cell Endocrinol **122**(2): 191-198.
- Veenstra, J. A. (2009). "Allatostatin C and its paralog allatostatin double C: the arthropod somatostatins." Insect Biochem Mol Biol **39**(3): 161-170.
- Voronezhskaya, E. E., S. A. Tyurin and L. P. Nezlin (2002). "Neuronal development in larval chiton *Ischnochiton hakodadensis* (Mollusca: Polyplacophora)." J Comp Neurol **444**(1): 25-38.
- Walker, R. J. (1986). Transmitters and modulators. The Mollusca, Vol. 9, Part 2. A. O. D. Willows. New York, Academic Press.
- Walsh, J. H. and G. J. Dockray (1994). Gut peptides: Biochemistry and Physiology. New York, Raven Press.

- Wayne, N. L., W. Lee, S. Michel and S. B. De Quintana (2004). "Post-afterdischarge depolarization does not stimulate prolonged neurohormone secretion but is required for activity-dependent stimulation of neurohormone biosynthesis from peptidergic neurons." Endocrinology **145**(4): 1678-1684.
- Weber, J. L. and E. W. Myers (1997). "Human whole-genome shotgun sequencing." Genome Res. **7**: 401-409.
- Welsh, J. H. (1966). Neurohumors and neurosecretion. Physiology of Echinodermata. R. A. Boolootian. New York, Interscience.
- Wilkie, I. C. (2001). "Autotomy as a prelude to regeneration in echinoderms." Microsc Res Tech **55**(6): 369-396.
- Wilkie, I. C. (2002). "Is muscle involved in the mechanical adaptability of echinoderm mutable collagenous tissue?" J. Exp. Biol. **205**(Pt 2): 159-165.
- Wilkie, I. C. (2005). "Mutable collagenous tissue: overview and biotechnological perspective." Prog. Mol. Subcell. Biol. **39**: 221-250.
- Wilson, S. and D. Bergsma (2000). "Orphan G-protein coupled receptors: novel drug targets for the pharmaceutical industry." Drug Des Discov **17**(2): 105-114.
- Winslow, J. T., N. Hastings, C. S. Carter, C. R. Harbaugh and T. R. Insel (1993). "A role for central vasopressin in pair bonding in monogamous prairie voles." Nature **365**(6446): 545-548.
- Yokota, Y. (2002). Fishery and consumption of the sea urchin in Japan. The Sea Urchin: From Basic Biology to Aquaculture. Y. Yokota, V. Matranga and Z. Smolenicka, CRC Press.
- Yoshida, M. (1966). Photosensitivity. Physiology of Echinoderms. R. A. Boolootian. New York, Interscience: 435-464.
- Yoshikuni, M., K. Ishikawa, M. Isobe, T. Goto and Y. Nagahama (1988). "Characterization of 1-methyladenine binding in starfish oocyte cortices." Proc Natl Acad Sci U S A **85**(6): 1874-1877.
- Young, K. G., J. P. Chang and J. I. Goldberg (1999). "Gonadotropin-releasing hormone neuronal system of the freshwater snails *Helisoma trivolvis* and *Lymnaea stagnalis*: possible involvement in reproduction." J. Comp. Neurol. **404**: 427-437.
- Zhang, L., N. L. Wayne, N. M. Sherwood, H. R. Postigo and P. S. Tsai (2000). "Biological and immunological characterization of multiple GnRH in an opisthobranch mollusk, *Aplysia californica*." Gen Comp Endocrinol **118**(1): 77-89.

Appendix A: Publications

Burke RD, Angerer LM, Elphick MR, Humphrey GW, Yaguchi S, Kiyama T, Liang S, Mu X, Agca C, Klein WH, Brandhorst BP, **Rowe M**, Wilson K, Churcher AM, Taylor JS, Chen N, Murray G, Wang D, Mellott D, Olinski R, Hallböök F, Thorndyke MC. (2006). A genomic view of the sea urchin nervous system. *Dev Biol* 300, 434-60.

Elphick, M. R. and **Rowe, M. L.** (2009). NGFFFamide and echinotocin: structurally unrelated myoactive neuropeptides derived from neurophysin-containing precursors in sea urchins. *J Exp Biol* 212, 1067-77.

Sodergren E, Weinstock GM, Davidson EH, Cameron RA, Gibbs RA, Angerer RC, Angerer LM, Arnone MI, Burgess DR, Burke RD, Coffman JA, Dean M, Elphick MR, Etensohn CA, Foltz KR, Hamdoun A, Hynes RO, Klein WH, Marzluff W, McClay DR, Morris RL, Mushegian A, Rast JP, Smith LC, Thorndyke MC, Vacquier VD, Wessel GM, Wray G, Zhang L, Elsik CG, Ermolaeva O, Hlavina W, Hofmann G, Kitts P, Landrum MJ, Mackey AJ, Maglott D, Panopoulou G, Poustka AJ, Pruitt K, Sapojnikov V, Song X, Souvorov A, Solovyev V, Wei Z, Whittaker CA, Worley K, Durbin KJ, Shen Y, Fedrigo O, Garfield D, Haygood R, Primus A, Satija R, Severson T, Gonzalez-Garay ML, Jackson AR, Milosavljevic A, Tong M, Killian CE, Livingston BT, Wilt FH, Adams N, Bellé R, Carbonneau S, Cheung R, Cormier P, Cosson B, Croce J, Fernandez-Guerra A, Genevière AM, Goel M, Kelkar H, Morales J, Mulner-Lorillon O, Robertson AJ, Goldstone JV, Cole B, Epel D, Gold B, Hahn ME, Howard-Ashby M, Scally M, Stegeman JJ, Allgood EL, Cool J, Judkins KM, McCafferty SS, Musante AM, Obar RA, Rawson AP, Rossetti BJ, Gibbons IR, Hoffman MP, Leone A, Istrail S, Materna SC, Samanta MP, Stolc V, Tongprasit W, Tu Q, Bergeron KF, Brandhorst BP, Whittle J, Berney K, Bottjer DJ, Calestani C, Peterson K, Chow E, Yuan QA, Elhaik E, Graur D, Reese JT, Bosdet I, Heesun S, Marra MA, Schein J, Anderson MK, Brockton V, Buckley KM, Cohen AH, Fugmann SD, Hibino T, Loza-Coll M, Majeske AJ, Messier C, Nair SV, Pancer Z, Terwilliger DP, Agca C, Arboleda E, Chen N, Churcher AM, Hallböök F, Humphrey GW, Idris MM, Kiyama T, Liang S, Mellott D, Mu X, Murray G, Olinski RP, Raible F, **Rowe M**, Taylor JS, Tessmar-Raible K, Wang D, Wilson KH, Yaguchi S, Gaasterland T, Galindo BE, Gunaratne HJ, Juliano C, Kinukawa M, Moy GW, Neill AT, Nomura M, Raisch M, Reade A, Roux MM, Song JL, Su YH, Townley IK, Voronina E, Wong JL, Amore G, Branno M, Brown ER, Cavalieri V, Duboc V, Duloquin L, Flytzanis C, Gache C, Lapraz F, Lepage T, Locascio A, Martinez P, Matassi G, Matranga V, Range R, Rizzo F, Röttinger E, Beane W, Bradham C, Byrum C, Glenn T, Hussain S, Manning G, Miranda E, Thomason R, Walton K, Wikramanayake A, Wu SY, Xu R, Brown CT, Chen L, Gray RF, Lee PY, Nam J, Oliveri P, Smith J, Muzny D, Bell S, Chacko J, Cree A, Curry S, Davis C, Dinh H, Dugan-Rocha S, Fowler J, Gill R, Hamilton C, Hernandez J, Hines S, Hume J, Jackson L, Jolivet A, Kovar C, Lee S, Lewis L, Miner G, Morgan M, Nazareth LV, Okwuonu G, Parker D, Pu LL, Thorn R, Wright R. (2006). The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science* 314, 941-52.

Appendix B: Putative sea urchin voltage-gated ion channel genes.

The VGIC gene predictions are arranged by ion selectivity and type (Gutman et al., 2005).

Putative sea urchin K⁺ voltage-gated ion channels

Kv1	SPU_026003, SPU_006470
Kv3	SPU_019341, SPU_006470, SPU_009204
Kv4	SPU_012871
Kv7	SPU_004907
Kv10	SPU_000092
Kv11	SPU_016505
Kv12	SPU_010067

Putative sea urchin Ca²⁺ voltage-gated ion channels

L-type	SPU_007770
N/P/Q/R-type	SPU_011692
T-type	SPU_017773, SPU_025833

Putative sea urchin Na⁺ voltage-gated ion channels

Na ⁺ VGIC	SPU_025997
----------------------	------------

Appendix C: Human neuropeptides organized by receptor family

Neuropeptides binding secretin-family GPCRs

Calcitonin (CT)

CGNLSTCMLGTYTQDFNKFHTFPQTAIGVVGAP-NH₂

Corticotropin-releasing hormone (CRH)

SQEPPISLDLTFHLLREVLEMTKADQLAQQAHSNRKLLDIA-NH₂

Glucagon

HSQGTFTSDYSKYLDSRRAQDFVQWLMNT

Gastric inhibitory polypeptide (GIP)

YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGGKNDWKHNITQ

Glucagon-like peptide-1

HDEFERHAEGTFTSDVSSYLEGQAAKEFI

Glucagon-like peptide-2

HADGSFSDEMNTILDNLAARDFINWLIQTKIT

Growth-hormone releasing hormone (GHRH)

YADAIFYNSYRKVLGQLSARKLLQDIMSRQQGESNQERGARAR
L-NH₂

Pituitary adenylyl cyclase-activating protein (PACAP)

PACAP 27 HSDGIFTDSYSRYRKQMAVKKYLA AVL

PACAP 38 HSDGIFTDSYSRYRKQMAVKKYLA AVLG

KRYKQRVKNK-NH₂

Parathyroid hormone

SVSEIQLMHNLGKHLNSMERVEWLRKKLQDVHNFVALGAPLAP
RDAGSQRPRKKEDNVLVESHEKSLGEADKADVNVLTAKASQ

Secretin

HSDGTFTSELSRLREGARLQRLQLGLV-NH₂

Vasoactive intestinal peptide (VIP)

HSDAVFTDNYTRLRQMAVKKYLYLNSILN-NH₂

Neuropeptides binding rhodopsin-family GPCRs (α)

Melanocortins;

Adrenocorticotropin (ACTH)

SYSMEHFRWGKPVGKKRRPVKVYPNGAEDESAEAFPLEF

Alpha-MSH

SYSMEHFRWGKPV-NH₂

Beta-MSH

DEGPYKMEHFRWGSPPKD

Neuropeptides binding rhodopsin-family GPCRs (β)

Orexin/hypocretin

QLLPDCCRQKTCSCRLYELLHGAGNHAAGILTTL-NH₂

RSGPPGLQGRLQRLQLQASGNHAAGILTM-NH₂

Neuropeptide FF

SQAFLFQPQRF-NH₂

Tachykinins;

Substance P (SP)

RPKPQQFFGLM-NH₂

Neurokinin A (NKA)

HKTDSFVGLM-NH₂

Neurokinin B (NKB)

DMHDFVGLM-NH₂

Cholecystokinin
KAPSGRMSIVKNLQNLDPNDRDY(-SO₃)-MGWMDF-NH₂

Neuropeptide Y
YPSKPDNPGEDAPAEDMARYYSALRHYINLITRQRY-NH₂

Pancreatic polypeptide (PP)
APLEPVYPGDNATPEQMAQYAADLRRYINMLTRPRY-NH₂

Peptide YY (PYY)
YPIKPEAPGEDASPEELNRYASLRHYLNLVTRQRY-NH₂

Gastrin-releasing peptide
VPLPAGGGTVLTKMYPRGNHWAVGHLM-NH₂

Neuromedin C
GNHWAVGHLM-NH₂

Neuromedin B
GNLWATGHFM-NH₂

Bombesin
pEQRLGNQWAVGHLM-NH₂

Neurotensin
pELYENKPRRPYIL

Neuromedins;
Thyrotropin-releasing hormone (TRH)
pEHP-NH₂

Ghrelin
GS-S(n-octanoyl)-FLSPEHQRVQQRKESKKPPAKLQPR

Gonadotropin-releasing hormone
pEHWSYGLRPG-NH₂

Arginine vasopressin
CYFQNCPRG-NH₂

Oxytocin
CYIQNCPLG-NH₂

Neuropeptides binding rhodopsin-family GPCRs (γ)

Galanin
GWTLNSAGYLLGPHAVGNHRFSKDKNGLTS

Kisspeptin/metastin
-APQGAVLVQREKDLPNYYWYSFGLRF-NH₂

Somatostatin-28
SANSNPAMAPRERKAGCKNFFWKTFTSC

Endogenous opioids;
Met-enkephalin
YGGFM

Leu-enkephalin
YGGFL

Dynorphin A (1-17)
YGGFLRRIRPKLKWNG

Neuropeptide W
WYKHAVSPRYHTVGRAAGLLMGLRRSPYLW

Melanin-concentrating hormone
DFDMLRCMLGRVYRPCWQV

Angiotensin 1
DRVYIHPFHL

Bradykinin
RPPGFSPFR

Neuropeptides binding delta rhodopsin GPCRs

Chorionic gonadotropin alpha (CGA)

APDVQDCPECTLQENPFFSQPGAPILQCMGCCFSRAYPTPLRS
KKTMLVQKNVTSESTCCVAKSYNRVTVMGGFKVENHTACHCS
TCYYHKS

Follicle-stimulating hormone beta (FSH)

NSCELTNITIAIEKEEERFCISINTTWCAGYCYTRDLVYKDPARPK
IQKTCTFKELVYETVRVPGCAHHADSLYTPVATQCHCGKCDS
DSTDCTVRGLGPSYCSFGEMKE

Thyroid-stimulating hormone beta (TSH)

FCIPTEYTMHIERRECAAYCLTINTTICAGYCMTRDINGKLFLPKYA
LSQDVCTYRDFIYRTVEIPGCPLHVAPYFSPVALSCKCGKCNT
DYSDCIHEAIKTNCTKPQKSYLVGFSV

Luteinizing hormone-releasing hormone (LHRH)

pEHWSYGLRPG-NH₂

Appendix D: Mass Spectrometry Expert Parameter Settings

MS Expert Parameter Settings	Peptides					
	NGFFFamide	Echinotocin	Sp-GnRH	Sp-TRH	GKamide	PPLN1
Capillary (V)	-4500	-4500	-4500	-4500	-4500	-4500
Skimmer (V)	49.84	38.69	100	27.54	52.62	37.3
Capillary Exit (V)	205.74	300	300	103.28	189.34	300
Octopole 1 DC	16.86	14.3	10.05	20	17.15	9.25
Octopole 2 DC	0	0	3.97	3.02	0	4.67
Trap Drive	57.99	64.14	124.17	103.41	62.5	128.41
Oct RF	267.21	300	300	300	226.23	226.23
Lens1	-6.64	-6.89	-4.18	-6.89	-5.9	-5.56
Lens 2	-52.79	-69.02	-66.07	-91.15	-58.69	-79.34

Appendix E: 34 Putative Sea Urchin Neuropeptide Precursors

Spp-1 GI:115711933 gi:115946155

MQVQQITVFLVACTLSVLVVAYAQEDAETVLLNRLRDI AARAAAAGELPDFFADVDDY**KRG**
GKKNMGSIHSHSGIHFGKRRDSESSERARNT**KRMRLHPGLLF**GKRAPVQKWDQWQAQDTY
NPDWELGQF

Spp-2 GI:72011734 gi:115975296 Glean3_19680

MKQIITSLVSI SAALLLVLI SEYTPRCNGQVHHRFSGWRPG**GKKR**SDAAEVNSNKITIE
RPQLPICQTTEERQLLEGDS DILGDLRRAAN**RMRL**LQLFNLSKTRLNLDNDATSNEVDER
PVYGDYLGTL

Spp-3 GI:115657896 Glean3_08352

MWACILGYVTWGGAALPTILGKELVLSNDGPEIADWV**QGKE**IPLRNQYWGDAVEEEEE
ELGMLSPDSE**KRQYPGGKRQYPGGKRQYPGGKRQYPGGKRQFPAGKRQFVGGELIP**SP
ELRQWP**GKRQWP**GG**KRQWP**GG**KRQYP**GG**KRQYP**GG**KRQWP**EV**KRQYP**GG**KR**SEDDQDLLPM
EIRQYP**GKRQWP**GG**KRQYP**GG**KRQYP**GG**KRQFP**GG**KRQFV**GG**EAL**EQESNIN**KR**FAPED
DTMDFRLSQLYDTNDNIVADEGELALEDLLDDIMVDTRPEFEDPRDLLLGNVDQEDVLA
LDLSALLGDRNPNG

Spp-4 GI:115767208 gi:115945640

MKSTVIVTLTICCLLYQTTRAASLTNRDGLSRQDILDLLQLYE EPIRQEGG**KR**SKGCGS
FSGCMQMEVAKNRVAALLRNSNAHLFGLNGP**GKRRR**SVDDLQVNDAAETE

Spp-5 GI:115933326 gi:115963862 Glean3_18666

MSRNAYLWAGLLL GALCLLITTTSIKADGEVTEVD**KR**ANYFRGR**GRK**P**GKR**DEPDAAALV
PDDDLSE**KRANM**FRSRLRGK**GKR**DDPDAAALPGDWDEE**KRANM**FRSRLRG**KR**DDPDA
AMLPGDWDEE**KRANM**FRSRLRGK**GKR**DEPDAAEALVPGDWEEE**KRANM**FRSRLRG**KR**D
DPDAAEALVPGDDLSEE**KRANM**FRSRLRGK**GKR**DDPDAAEALVPGDDLSEE**KRANM**FRSR
LRGK**GKR**DDPDAAEALVPGDDLSEE**KRANM**FRSRLRGK**GKR**DDPDAAEALVPGDDLSEE**K**
RANMFRSRLRGK**GKR**DDPDAAEALVPGDWDEE**KRANM**FRSRLRGK**GKR**DDPDAAALVGDDF
GDEFVDEE**KRANM**FRSRLRG**KR**DDPDAAALVDFEFMDEE**KR**ANYFRGRGR**R**P**GKR**DEPDA
ALVEDE**KRAN**FRARQR**P**KL**GK**

Spp-6 GI:72008820 Glean3_03108

MKFSNGRGAFLVVNLI FVLCLVDHMAECR**PARK**TRDVDEDELEKEEDSLINALEKVLADE
EVIDNAENDSDDETGITDRELSLMLSMLRDDVSPSRLRGYF**GK**WRPAYYPSESLHVGAL
EPLATGFLPSRYSG**QK**RFLTGALEPLSSGF**I****KK**GFNTGAMEPLGSGF**I****KK**GFNSGAMEP
LGAGFF**KK**GFNSGAMEPLGAGFF**KK**GFNSGAMEPLGAGFF**KK**GFNSGAMEPLGAGFF**KK**G
FNSGAMEPLGAGFF**KK**GFNSGAMEPLGAGFF**KK**GFNSGAMEPLGAGFF**KK**GFNSGAMEPL
GSGF**I****KK**GFNSGAMEPLGSGF**I****KK**GFNNGAMEPLGSGF**I****KK**GFNSGAMEPLGSGF**I****KK**GF
NSGAMEPLGSGF**I****KK**GFNTGAMEPLGSGF**I****KK**GFNSGAMEPLGSGF**I****KK**DFNTGAMEPLG
SGF**I****KK**GFNSGAMEPLGAGFF**KK**GFHAGAMEPLSSGF**ID****GKR**GFYNGAMEPLSAGFH**QK**
RGFHEGEMDK**KK**GFHNGAMEPLKSGFLKD

Spp-7 Glean3_24381

MNNYAFLFCLACAIGQVWTLPIEDKDGLDIEDQEEAE**KR**FGSMNMEPLVSGFY**KR**FGSGL
DSMQSGFY**KK**NFGSGLNMEPMQSGFY**KK**NFGGSMEPMQSGFY**KR**FGGAMEPMSGFGY**KR**F
GSGSLEPMSGFGY**KK**NFGGSLEPMQSGFY**KR**FGGANEPMRSGFF**KR**FGSGLPMSGFGY
KKNFGGSLDAMQSGFY**KR**SQEETD

Spp-8 GI:115764725 GI:115935324

MANRQLLALAFIVSLALAVVEARNFHAAMGGPRPWQAGMKQOSALPKGTNPFLKRLKQI
VFQPDGFYDPGMDHFAFGEAFNADE

Spp-9

MRSSLAVLLLAACLAATISRESPVQAVPRIRPAILQHGMPPF**GKR**GYSGNNARDCFH**R**ALND
DKNSEELVNLIEAWYRMKVEDGLSCMNGLSAFDEAAA

Spp-10 GI:115647054 GI:115936357

MKSVMYQVVLAFVAVLVCVAWTCQAYGLDQDEY**RR**GAAENALDEQEIYEIIESLEHAMS**KR**
GSVKHLGLANVDNWRMMKNVNR**LR**NLLNS**GKR**SDQQQLDSQ

Spp-11 GI:115666438 GI:115939808

MNSLILVVMGLLLLTAELIPAAPAPYFDEDAMDLMDFVFNFKDDSAV**KR**S**P**MLQKSCIYT
CLACSKNTQMTMPECIYGCQS**AGR**DPSQARAYNACHKYLH**SGR**

Spp-12 GI:115620334 GI:115939916

MDSNMTVRSVLVLSVLLAVVSCHAHNTFSFK**GR**SRYPF**GKR**AITDGS**AV**DTASQRFESI
NLDDF**QK**PESQTLREMLTEL**RGY**CDFLLKLLDGVRPDL**PQQRK**

Spp-13 GI:115660734 GI:115930865

MTTSTTMELRFLVLLVFCALATSLPANLARE**RR**TTNPVLRDK**GR**ESMKT**KQF**RIGYRYG
RAWQPPTTLDNVDNYGADNYDNEAF**QFR**NLPLEKLI**AQ**LEKADENGGY

Spp-14 GI:115958765

MEPHQLTLTVFILSLSVLMVSTGAFPQEV**RGDR**TGHMIDGFSNDIDLLPLQ**E**TALIRL
LSNLQSSSEYASGEDETYPMVAS**KR**SR**SGR**KLRF**CM**DVIR**NT**WR**LCR**NT**RSN**

Spp-15 GI:115920974 GI:115978337

MNTLSQYLLLICSLLVFIQSYALP**T**YDKQNVDELQGDNDID**E**QQ**LE**MWDAMQGGDND**D**VF
SRL**TR**GG**EAF**SRD**RRR**V**CV**SDCS**FC**HSFF**P**TYKLG**NC**FG**CRK**GFHD**L**G**CKQF**RY

Spp-16 GI:115898497

MNLTTCYLAIALAAILAVAAGRTLDLGLPVMELQ**EED**FPQ**ME**Q**Q**N**ME**HQ**S**MR**DM**V**S**AR**L**WS
II**Q**RL**K**MD**Q**AVDL**K**DELDTLD**Q**GA**E**K**M**LS**ED**FN**KR**GR**RR**P**A**R**K**IC**I**ND**I**W**K**GR**GG**GL**RCN**

Spp-17 GI:115751547 GI:115977303

MNSTISTLLSLAALLII**AVQ**M**SSAL**SIT**E**GP**Q**GS**A**W**A**LED**N**EE**P**V**D**Y**R**SV**L**K**M**K**Y**E**I**L
L**K**L**M**ND**L**C**D**E**L**DM**C**PP**S**Q**V**PAR**Q**AP**V**V**RR**GD**NN**Q**ERR**GG**A**HL**F**W**R**T**G**V**L**N**K**S**P**IM**K**A**A**N

Spp-18 GI:115839524 GI:115976432

MQPNSIISVAVVMTLL**AT**L**FT**Q**AVCS**LQ**F**ETTQ**DR**V**PA**K**R**L**F**W**V**D**KK**D**H**P**V**D**T**D**F**F**T**VR**A**N
D**A**E**E**V**L**D**C**F**V**E**V**C**I**A**D**F**V**N**C**A**KK**C**L**F**Y**EN**G**NT**C**L**P**T**C**R**H**T**RS**I**C**S**V**Q**C**F**KR**Y**D**V**D**V**S**D**S**V
H

Spp-19 GI:115722995 GI:115946386]

MRCYTWVFTVSVFLTSA**V**L**A**I**A**SPRW**P**GGNS**Q**QR**P**W**E**L**G**D**A**D**F**SS**P**IT**D**T**S**F**V****KR**LL**G**R
I**H**E**D**L**R**Q**K**SN**Q**A**A**D**L**R**D**A**T**S**R**G**F**E**T**V**D**L**K**Q**L**S**D**NG**A**GL**Q**V**H**G**V**R**Q**T**RG**K**CM**GR**F**GP**Y**ML**N**
C**K**R**S**G**P**T**T**I

Spp-20 GI:115899431 GI:115955088

MQSSLVVISAVFAAVLLGLATATDVLP**Q**QR**ERR**APS**DE**ME**G**F**E**PG**T**W**D**DA**E**E**V**E**D**F**R**T**N**E
F**A**K**R**I**A**D**N**D**F**A**A**M**R**H**Q**E**R**S**N**S**M****RR**T**R**LL**Q**A**M**N**E**L**A**K**A**G**K**R**S**S**T**D**KR**S**L**A**Q**N**D**Y**M**M**V**R**Q**D
L**A**NG**R**L**Y**R**S**L**M**D**R**M**L**S**E**A**G**K**R**

Spp-21 GI:115916164 GI:115944736 Glean3_03170

MYVNALTFPLAVAI FLVGLA I L P G M L A E D G M E L T H D E Q P L N A M E I R S P Q D E Q I D L R Y L L Q
NFLDNRDTRSWSPLRNSKQCIGDKCRYAWKRGFETPASSRINSRGAPAWATSQFSNGGCS
GSSCLTGWKRGFRVLPQLNDDDN

Spp-22 GI:115611897 GI:115611899 GI:115615405 GI:115615407

MLSYGQVGFVAVSMLFLLQIVTGLNIHRLTQSQDPVQSAAEQEIAHLLQRYPEQQVLDRN
KRFPKFLTCEITDDDCARLDRPACCAKSVMGPFVWCKALGELNAACNSYSMPHMWGGQRNG
HFCPCQHLACVGHRRGGCDLISDK

Spp-23 GI:72015992 GI:115940672 Glean3_14142 Glean3_03878

MTSQLVTLVLAVFVCSAAVVYSQSPSSPPSASPPTVLATEPITTPRPAVATTPPPVDNGT
PAPSANGTDAPTPVVTDA PMTSAKDGGDDGMKGDGDGQKGHDDDEEGGGLRRGDIALAI
LATILVVAVICTFIGLCYWKYKGN SYVTVTADTTYRQ

Spp-24 GI:115640656 Glean3_14640

MAVSTPKWLLVGGFLVAALAISSRGAWAEGIRAADAMDAINS GKQQFLQGLRGALAAVA
AQKDNLVEELLADVEDEEENEVEKIEEELGVLDPNVKTVEIRLPNAEIAESGTTLC
TSTSLTEALQDIVGFSPLINQNAVTHMALYGC PDQAASTQDYW

Spp-25 GI:115674678 GI:115709525

MKLTICFIISVILFTP SLGIFEFWHDIRDHGGHSPVYHLSIDL LYLECNVPPAYRGLKF
KKGKRDV RVGKRVLNHSMVYFRGKTYEFGANSLVLHEGGVGR TNPTSCPTV RKT TAGRSL
CTEQATELAVRYHQDHNFNLF SKNCHMFAEWLINRLLNNKCTI

Spp-26 GI:115749015

MGYERRILR TLLSILIVLASFVTVYGERDSNFMQHKLFRNIVPSPLIQKWRENRMGPAEE
KTSNEQWRDELLSNLRNVLRKHNASPSRSIDRS DITGYGLQEPMQQLPADVTADQLFIL
EGAVNSPTENYEEETPIDEDKRNNGFFF GKRNNGFFF GKRS DSDASSTKMDDRLPKYKS
SGSFDKCRPCGPRQGTMRNGX

Spp-27 GI:115627958 GI:115940296 Glean3_20714

MNKKCIVTATHILLFLQALSSINALPILPEEPEHDVVHDYHYETS FVTPYEEIDELAHLK
WSVDYTTSLRLRYMLSVPLFDDVTADDAGSRSENFPALSRAGREWLAVGMSPYGASTDADM
VTFIRGSNGEVTILDYTDENGVLTVDDKQDYKLWNAWIRQPLGEDISDDEDDHR
VKRKELVIEYGRQIDTCDDKDYLVVRFISI

Spp-28 GI:115747988 GI:115918610 Glean3_02352

MGMLEWKLPLLLLLLVHVAQQA PVDPKPAEEDRDQNEGEGIAGEEILNIAEGDGIADLGL
EYEGYLREVIQVLEADPEMKSHMDEM TDNDLLVKT SIGNAPISKNEMEALITNHVKEYVL
TIKYIMITILQGNRLIDIFGKLVYLLTGKFGKQLNNVGS HIRGQDELK KKEIQRLRII
ARKAMEQQALGKAKPADQKFLDNLIGHLD PSSMDAFSPDDITKLIQKVDNY

Spp-29 GI:72007133 GI:115946242 Glean3_15798

MWYVILSMLLLGALASSEYSGMSLRDRIALRNL MNSNYDFPLAPSQRDLENTATNSKV
VIPPLTFIAGGAGEGVQHLGAEGDIPNRQNP IPEVSSQYDNPPNCP PGGELTGKVKVYS
KVHRHQVCACEKGWEAGERRDCSSAPKCLPNMPND AEFVNQYQLSERSLKNRQLFTQTR
NKYSVGDKRS DHMAKSPVYKRSINSYLP GDMVRHVSKKSAYPIRTYVNPYTYNQPHLKS
VVAKKAPVYSGAKPIM

Spp-30 GI:115650747 Glean3_21555

MRVLVALALCLCFIAPSPVLSFTMP EEFVENK MADVGE EGTGQNNINSIAKSLIREVFG
AAEEREMEAENEAEDEAELSLSKRTTGSTRPQREIRARAQYAAARRPPVTTRSKFTFGKRS
SPTPVISRPLAEQLLEELQRNAEMSDDWRESDKLALLNDAALYDSLVD SHQVQK DAYSAF

SFG**KR**GMSAFSFG**KRA**QPSFAFG**KR**GLMPSFAFG**KR**PHGGSFAVFG**R**RDWAPREQDFANA
AEESGPY**KR**GDLAFAFG**K**REDQ

Spp-31 GI:115762091 GI:115953199

MGYLVPLHVLVTLVLSVSGTVSTNQNQVETVFIQDRVPSLLCMPKNSSDS**AGK**VASHYDVEN
EPYGSMDCTFTMNAKQGQRINIRFDYFDVARSYDS**REGRCR**SEGDSLISYESGNIFLIK
SNILTPEVRLCGGT**GK**FPDDYQSVSNVTVRFKTFDEYSPDNGIKFTYTAFLQPQERTPE
CFMCNDGTMCIDPLLLTCNGMPNCNDDSDDEEPSNCKASTNIFEKGIETLGVPMMAAIGAAV
GLLLLCLIVLCIVCCCC**KR**NNPSDQKAQYRAPPTPRPGSYPSHNSNYSSYSQASNGN
TNGTAG**GR**PPYPNFPFPGYPPFHNGGYSVGYNYENGRMEFPQKV

Spp-32 Glean3_14973

MSMYVFLIFLTFNRPSACTIIHVREGEAAKLNFSPCNITKTTLQHSQAPFYSSVNP
SNPLPPRYELANHTKSDDDTTCFILLIIDPILRNDVGTIILLVYNHSQLLEYPKIDLRV
AYPP**GK**ASCESSNHVDGAWIKLQCTAPLGNLAGQIVCYQNGMRLPPLGTPVEN**GRR**LSQ
VILARIADHSV**RCC**SSTLDQTKTLQQCRDCGWDPIONIALKIDITDLSTTVSPTQHSKSTQ
NFTLEGETPSSPPHMITNMATNKDVWSNFKRPVYIALSFLISSYVIYSITLLLKQVKTRC
LKRSLIPQ**R**DPVTLQCQILKQNSGFSSCCNDNFNPLVFDRIHPHLL**RSK**LHYQGIRGDSL
WTMSLLANRLQAVVNGACSAWMPVTSVGPQGSVIGQLLFLLYINDIKLKSICR

Spp-33 GI:115741907 GI:115741905 Glean3_08863

MEIMGLSIYIALCSVLLFLGAPRAEGCSCEGAHPQQQFC**R**ASYVIGARVL**KREY**IPKNP
RGLGFSTGLLGSDHAAIEPLGERTRPYPGTGLHRLPDE**R**LPSEQFGDGGRQLQDLEPLEP
LPYPDGLGPPILPEAADPFQNDGLFAPIPQSDGLQYPQNDGLEPFPQAGLQPLQOPEG
LQPLPFEGGPEPLGQPEGLLPLPHPEGLQPLPEEPHTHEALPPLPLPEDMKPLEDTITTL
TTPPPDLVMTDESHTMTLPPQEKPSNETVLEPTIGAEPIAEEERIIDPESMEAWNYK
IIYTLKLSHIYKTGMGQEGWKQGDVQVDSSAASSL**GK**TDLKT**GK**NYLIAGSSANQVVL
CDWVQEHKLTTRQKQGLKTIY**GK**SCDVCQIHSCTSSRH**CKK**PKENRATTCLYDASQVHI
GDAYSPLYDCIASHSRCTMRHGKCKWQKNNDYHKCVKHQLP

Spp-34 GI:115966309 Glean3_26982

MCCGLGALACCFGTAACSCCCSACPCKNSTSTRIVYGLFLLVAVVSAVFLIPQVDQAL
SNSPLLCKDVAVVQGLIPSEVsCERLAGYRSVYRVVSGVAAFFFLSLIMINVKSSKDP
SPIQNGFWFFKFLVMCGLCVAAFFIPNGSFENVFMYFGMVGAFAFI IQLVLLVDFAHSW
NESWV**GR**MEETEHEKGWYCALMSSTVVMYLIALTGFILFFIFYIGT**GK**ECSLHKFFISFNL
VLCVMSVISILPKVQEAMP**RS**GLLQSAVISMYTMYLTWSAMSNNPDDTCNPSITTIQ
IGPSGNNTNVHNQDVGSAENWASFAIWLICLIYACIRTASTNNV**GK**LTGSEDNLQYGTNE
KTLGSTNSSGGDSKPADGDAEKWGQEVYDNEEDTVSYSTFFHIMLMALAAFYMMMLTS
WFQAGANFDSLAAANSAMWVKISSSWCVVALYVWTLVAPIILSEREFS