

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

**Molecular biological studies on neuropeptides of the
adipokinetic hormone/red pigment-concentrating
hormone family and the neuroparsin
family in the arthropod sister groups of insects and
crustaceans**

Lance Anders

Dissertation presented for the degree of

MASTER OF SCIENCE

in the Department of Zoology

Faculty of Science

University of Cape Town

August 2012

Supervised by Prof. Dr. Gerd Gäde, Dr. Heather G. Marco
and Dr. Rob Ingle

Plagiarism declaration:

I, Lance Anders, know the meaning of plagiarism and declare that all of the work in this dissertation, save for that which is properly acknowledged is my own.

August 2012

Acknowledgements:

I would like to thank my supervisor, Prof. Gerd Gäde for giving me the opportunity to investigate the fascinating world of insect and crustacean neuropeptides, as well as critically reading my thesis manuscripts.

Dr. Heather G. Marco for developing my critical thinking and scientific technique to overcome many of the challenges encountered and without whom I would most certainly not have made it thus far with so much.

Dr. Rob Ingle, thank you for critically reading my thesis manuscript, developing my scientific method and for welcoming me into his lab.

Many thanks to all my friends especially my lab colleagues, Delroy Guzah, Anastasia Diener, Emang Molojwane and Michael Wolf for their support both academic and non-academic.

Funding during my research was obtained from the National Research Foundation (NRF), KW Johnston Bequest, UCT Equity Scholarship and the Masters Research Scholarship.

I would like to especially thank two of the most important people in life, my mother and father, because without their support and encouragement, my dream would not become my reality.

Lastly but most importantly of all, I would like to thank Vanessa Lakay for whom I ultimately did it all for. Words just cannot express my gratitude for the amount of love and support you have given me, so in simple words, thank you for everything you've done for me.

Abstract:

This study describes the identification of three novel precursor transcripts which includes the adipokinetic hormone (AKH) / red pigment-concentrating hormone (RPCH) and the neuroparsin (NP) from the South African spiny lobster, *Jasus lalandii*, the RPCH and the RPCH receptor (RPCHR) from the water flea *D. pulex* and the NP from the southern green stinkbug, *Nezara viridula*. The study also investigates the localisation and expression profiles of the AKH/RPCH and NP transcripts within crustaceans and insects. In *J. lalandii* the RPCH cDNA consists of a 297 bp open-reading frame (ORF) that encodes a preprohormone of 98 amino acid residues and an NP cDNA consisting of 312 bp ORF that encodes a preprohormone of 103 amino acid residues. In *N. viridula* NP cDNA consists of 321 bp ORF that encodes a preprohormone of 106 amino acid residues. Localisation studies involving the PCR amplification of RPCH and NP transcripts within various *J. lalandii* tissues and the localisation of NP within various *N. viridula* tissues were investigated. The results show that within *J. lalandii*, the RPCH transcript is limited to the X-organ and the suboesophageal ganglion (SOG). Whereas in both *J. lalandii* and *N. viridula*, the NP transcript was detected not only in neural tissues like the brain and SOG but also in muscle tissue (specifically in the heart muscle of *J. lalandii*) and reproductive organs (depending on the sex). In addition, a preliminary study involving the use of RNA probes in *in situ* hybridisation experiments was used to determine where the *Daphnia pulex* RPCH and RPCH receptor (RPCHR) mRNA transcripts localise. The results of the preliminary study suggest that the *D. pulex* RPCH is likely synthesised in or near the ovaries. The results of the preliminary study involving *D. pulex* RPCHR show that the probe is located in tissue which appears to be the thoracopods.

Distribution of the RPCH in *J. lalandii* is exclusively expressed in brain and suboesophageal ganglion tissues whereas preliminary localisation studies in *D. pulex* show that the RPCH associates with the ovaries. The *D. pulex* RPCHR associated with the thoracopods suggests that expression of RPCH is linked to movement involving reproductive behaviour. The distribution of NP in *J. lalandii* and *N. viridula* suggests that the expression profile of this gene is conserved between advanced crustaceans like *J. lalandii* and the stinkbug and locust families.

Table of Contents:

Plagiarism declaration:	i
Acknowledgements:	ii
Abstract:	iii
List of figures:	viii
List of tables:	xi
List of abbreviations:	xii
1. Introduction	1
1.1 Neurosecretion	2
1.1.1 Insect and crustacean neuroendocrine complexes	2
1.1.2 Neuropeptides.....	3
1.2 Animals used in the study.....	3
1.3 Adipokinetic hormone / Red pigment-concentrating hormone family	5
1.3.1 The crustacean RPCH	5
1.3.2 The insect AKH.....	7
1.3.3 Locust physiology and AKH	8
1.4 Mobilisation of energy stores during insect flight	9
1.5 Introduction to the parsin neuropeptide family	14
1.5.1 Insulin-related parsin.....	14
1.5.2 Ovary maturing parsin.....	15
1.5.3 Pacifastin.....	15
1.5.4 Neuroparsin	16
1.5.5 Neuroparsin homologous genes in <i>L. migratoria</i> and <i>S. gregaria</i>	18
1.5.6 Neuroparsin-like peptides in other organisms	18
1.6 Aims of the study.....	22

2.	Materials and Methods.....	24
2.1	Animals	24
2.1.1	<i>Nezara viridula</i> (Southern green stink bug).....	24
2.1.2	<i>Jasus lalandii</i> (South African west coast spiny lobster)	25
2.1.3	<i>Daphnia</i> spp. (waterflea)	25
2.2	Dissections.....	26
2.2.1	Dissection of <i>Jasus lalandii</i>	26
2.2.2	Dissection of <i>Nezara viridula</i>	26
2.2.3	Dissection of <i>Daphnia pulex</i>	26
2.3	DNA techniques	28
2.3.1	Genomic DNA extraction method.....	28
2.3.2	Splinkerette PCR system	29
2.4	Polymerase chain reaction	31
2.4.1	Splinkerette PCR	31
2.4.2	28s ribosomal DNA amplification	31
2.4.3	Agarose gel electrophoresis.....	31
2.4.4	Ligations	32
2.4.5	Transformation.....	32
2.4.6	Plasmid DNA extraction	33
2.4.7	Bioinformatics.....	33
2.5	RNA Techniques	34
2.5.1	RNA extraction.....	34
2.5.2	Total RNA quality and quantity determination.....	34
2.5.3	DNase treatment of total RNA.....	34
2.5.4	First strand cDNA synthesis	35
2.5.5	3' Rapid Amplification of cDNA Ends PCR system	35
2.5.6	5' Rapid Amplification of cDNA Ends PCR system	39

2.6	Whole mount <i>in situ</i> hybridization for the detection of mRNA transcripts	44
2.6.1	<i>Nezara viridula</i> and <i>Daphnia pulex</i> tissues	44
2.6.2	MEMFA fixation	44
2.6.3	Prehybridisation treatment.....	44
2.6.4	Hybridisation procedure.....	45
2.6.5	Washing and detection	46
3.	Results.....	47
3.1	PCR-based cloning and sequence analysis of <i>Daphnia pulex</i> (German and South African ecotypes).....	47
3.1.1	Actin	47
3.1.2	Verification of the red pigment-concentrating hormone	49
3.1.3	Verification of the red pigment-concentrating hormone receptor	52
3.1.4	Localisation of RPCH and RPCHR in <i>Daphnia pulex</i> tissues	60
3.2	PCR-based cloning and sequence analysis of <i>Nezara viridula</i>	62
3.2.1	Splinkerette PCR	62
3.2.2	Amplification of NP in 3' RACE and 5' RACE PCR.....	62
3.2.3	Actin	65
3.2.4	Localisation of NP in <i>Nezara viridula</i> tissues.....	66
3.3	PCR-based cloning and sequence analysis of <i>Jasus lalandii</i>	71
3.3.1	Actin	71
3.3.2	Crustacean hyperglycaemic hormone II	72
3.3.3	Red-pigment concentrating hormone	73
3.3.4	Neuroparsin	74
3.3.5	Localisation of NP and RPCH in <i>Jasus lalandii</i>	77

4.	Discussion	79
4.1	AKH/ RPCH homology.....	80
4.2	AKH/ RPCH phylogeny and localisation	83
4.3	NP homology.....	87
4.4	NP phylogeny and localisation.....	88
4.5	RPCHR homology and localisation	94
5.	Future perspectives.....	98
6.	References	99

University of Cape Town

List of figures:

Figure 1: Comparison between the primary structure of Panbo-RPCH and Dappu-RPCH.	6
Figure 2: Diagram illustrating the position of elements that comprise the open reading frame (5' to 3') for the Aedae-AKH-I preprohormone.....	8
Figure 3: A general overview of AKH controlled mobilisation of trehalose in the fatbody of insects.....	11
Figure 4: A general overview of AKH controlled mobilisation of proline in the fatbody of the beetle, <i>P. sinuata</i>	12
Figure 5: An overview of AKH controlled mobilisation of lipids (DAG) in the fatbody of locusts during flight when carbohydrate energy stores are exhausted.	13
Figure 6: An alignment of <i>L. migratoria</i> and <i>Schistocerca gregaria</i> neuroparsin prohormone.	16
Figure 7: An alignment showing <i>L. migratoria</i> / <i>S. gregaria</i> NP precursor 1 (NPP) vs. <i>A. aegypti</i> / <i>A. gambiae</i> "ovary ecdysteroidogenic hormone (OEH).	19
Figure 8: Diagram illustrating certain body parts of <i>D. pulex</i> (South African ecotype)...	27
Figure 9: A modified diagram taken from Devon et al. (1995) of the (A) vectorette and (B) splinkerette system	30
Figure 10: Diagram illustrating the 3' RACE PCR system.....	36
Figure 11: Diagram illustrating 5' RACE PCR.....	40
Figure 12: Represents the highly conserved partial actin cDNA and deduced amino acid sequence of <i>D. pulex</i> (both the German and South African ecotype).....	48
Figure 13: An alignment of the <i>D. pulex</i> preprohormone RPCH sequences	50
Figure 14: The RPCH preprohormone cDNA and deduced amino acid sequence of <i>D. pulex</i> from A) German and B) South African ecotypes	51
Figure 15: An alignment of the different amplicons in the elucidation of the Dappu- RPCHR gene	53
Figure 16: Represents the consensus nucleotide sequence for the Dappu-RPCHR and its deduced peptide sequence from the <i>D. pulex</i> (German ecotype).....	58
Figure 17: (A) Depicts an agarose electrophoresis gel showing the actin and the partial (~750 bp) amplification of the Dappu-RPCHR gene and (B) represents the Dappu-	

RPCHR cDNA and its deduced peptide sequence from <i>D. pulex</i> (South African ecotype).....	59
Figure 18: An alignment of the complete preprohormone RPCHR peptide sequence amplified from the German ecotype (GER) <i>D. pulex</i> and partial sequence from the South African ecotype.....	59
Figure 19: Images from the <i>in situ</i> hybridisation experiments of control (A), Dappu-RPCH (B) and Dappu-RPCHR (C).....	61
Figure 20: shows an agarose electrophoresis gel with first partial amplification of the NP gene in 3' RACE.....	64
Figure 21: Nucleotide sequence of the Hemipteran insect <i>N. viridula</i> NP prohormone cDNA and its deduced amino acid sequence.....	64
Figure 22: Represents the ~900bp actin sequence with the primers Nvactin f and Nvactin r.....	65
Figure 23: Represents the DNA contamination actin PCRs for various <i>N. viridula</i> (A) mature female, (B) mature male, (C) immature female and (D) immature male tissues.....	68
Figure 24: Represents the expression profile for the NP gene in various mature female <i>N. viridula</i> tissues.....	69
Figure 25: Represents the expression profile for the NP gene in various mature male <i>N. viridula</i> tissues (A) and the Actin control (B).....	69
Figure 26: Represents the expression profile for the NP gene in various immature female <i>N. viridula</i> tissues.....	70
Figure 27: Represents the expression profile for the NP gene in various immature male <i>N. viridula</i> tissues.....	70
Figure 28: shows the amplification of the <i>J. lalandii</i> cHH-II gene from x-organ tissue.....	71
Figure 29: Represents the cDNA and its deduced amino acid sequence of the decapod crustacean <i>J. lalandii</i> partial actin gene.....	72
Figure 30: Represents the cDNA and its deduced amino acid sequence of the crustacean <i>J. lalandii</i> RPCH preprohormone.....	74
Figure 31: shows a RACE PCR reaction whereby the <i>J. lalandii</i> NP was amplified.....	75
Figure 32: Nucleotide sequence of the decapod crustacean <i>J. lalandii</i> NP prohormone cDNA and its deduced amino acid sequence.....	76

Figure 33: Represents the DNA contamination actin PCRs for various <i>J. lalandii</i> tissues.....	77
Figure 34: The expression profile of the RPCH preprohormone and NP prohormone for <i>J. lalandii</i> in various tissues.....	78
Figure 35: Alignment showing the RPCH preprohormone amino acid sequence from <i>D. pulex</i> against the crustaceans RPCH preprohormone (A) and the insects AKH preprohormone (B)	82
Figure 36: Phylogenetic analysis of the AKH/RPCH preprohormone gene family in insect and crustaceans. The phylogenetic unrooted tree was constructed using the minimum evolution method.....	86
Figure 37: Alignment showing the <i>D. pulex</i> NP prohormone amino acid sequence against the crustaceans NP prohormone and the insects NP prohormone.....	89
Figure 38 Phylogenetic analysis of the NP gene family in insect and crustaceans.....	91
Figure 39: Alignment showing the <i>D. pulex</i> RPCHR amino acid sequence against insect AKH receptor sequences.....	97

University of Cape Town

List of tables:

Table 1: Oligonucleotides used in splinkerette construction and the splinkerette PCR system to amplify the Nezvi-AKH gene.....	30
Table 2: Primers used in general and 3'/5' RACE PCR.....	37
Table 3: Primers to amplify target genes in general PCR.....	43
Table 4: Expression of NP in various <i>N. viridula</i> and <i>J. lalandii</i> tissues	92

University of Cape Town

List of abbreviations:

°C	degrees Celsius
μ	micro
AA	amino acid(s)
AKH	adipokinetic hormone
AKHR	AKH receptor
Amp	ampicillin
AAP	abridged anchor primer
AUAP	abridged universal amplification primer
bp	base pair(s)
CC	corpora cardiaca
cDNA	complementary DNA
cfu	colony forming unit(s)
cHH-II	crustacean hyperglycaemic hormone-II
cm	centimetre(s)
DEPC	diethylpyrocarbonate
dH ₂ O	distilled water
DNA	deoxyribonucleic acid
dNTP	deoxynucleosidetriphosphate
DTT	dithiothreitol
EDTA	ethylene diamine tetraacetic acid
g	gram(s)
gsfp	gene specific forward primer
gsrp	gene specific reverse primer
h	hour(s)
hybsol	hybridisation solution
IRP	insulin-related peptide
IPTG	isopropyl-β-D-thiogalactopyranoside
ILGFBS	insulin-like growth factor binding proteins
JH	juvenile hormone
kb	kilobase(s)
l	litre(s)

LB	Luria Bertani broth
M	molar
μl	microliter(s)
min	minute(s)
ml	millilitre(s)
MOPS	4-morpholinepropanesulfonic acid
mRNA	messenger RNA
MW	molecular weight
NP	neuroparsin
NPLP	neuroparsin-like peptide
NPP	neuroparsin-related peptide
OEH	ovary ecdysteroidogenic hormone
OMP	ovary maturing parsin
PCR	polymerase chain reaction
PTTH	prothoracicotropic hormone
RACE	rapid amplification of cDNA Ends
RNA	ribonucleic acid
RPCH	red pigment-concentrating hormone
RPCHR	RPCH receptor
RT-PCR	reverse transcriptase PCR
S	second(s)
U	unit(s)
UCT	University of Cape Town
UI3R	used in 3' RACE
UI5R	used in 5' RACE
v/v	volume per volume
w/v	weight per volume
X-Gal	5-bromo-4-chloro-2-indolyl-β-D-galactoside

1. Introduction

The phylum Arthropoda contains 4 distinct subphyla, the Chelicerata (e.g. spiders and scorpions), Myriapoda (e.g. centipedes and millipedes), Crustacea (e.g. crabs and shrimps) and Hexapoda (e.g. insects and springtails). The most well-known of these arthropods are probably the crustaceans and insects, because of their high economic value. The decapod crustaceans such as shrimp, crayfish, prawns, etc. all form a lucrative commercial industry because their meat is a sought-after delicacy world-wide. Crustaceans are also of ecological importance because they dominate marine and fresh water habitats and are part of complex food chains. Insects are also very important economically: they are known for many beneficial aspects like pollination, but also for more disastrous aspects like crop damage (e.g. locusts) and as vectors for spreading diseases (e.g. mosquitoes). The adaptability of insects and crustaceans are very interesting, together they dominate land and sea and the sheer variety is amazing. Understanding what makes these two groups similar and/ or different could help in the management of pests and disease transmission or for economical purposes, involving the health and growth of aquaculture in the seafood industry. One strategy to determine the relationship between insects and crustaceans is to compare the peptides which control bodily functions. Such regulatory peptides are synthesised in modified neurons and are thus named neuropeptides. Analysing these neuropeptides in terms of their structure and function, will better serve to answer questions about the relatedness between crustaceans and insects, and allow us to better identify evolutionary aspects between the two which can be exploited. Comparing crustacean and insect neuropeptides will allow for the potential design of safer species-specific molecules for the intention of insect pest management. By using the knowledge gained there is the possibility that non-target beneficial insects and crustaceans will not be harmed or affected to the same degree [reviews by (Gäde & Marco, 2006; Gäde, 1997)].

1.1 Neurosecretion

1.1.1 Insect and crustacean neuroendocrine complexes

In insects the neuropeptide-synthesising neural complexes are located in the cerebral ganglia (brain) and more specifically in the pars intercerebralis, the median and lateral parts of the protocerebrum and the retrocerebral corpora cardiaca. Other neuronal tissues of interest that synthesise neuropeptides are the thoracic ganglia and the abdominal ganglia. Neuropeptide-producing cells are typically arranged symmetrically, and their axons end in neurohaemal complexes where the neuropeptides are stored and also released (Gäde & Marco, 2006).

Similar to insects, physiological processes in crustaceans are also under the control of neurohormones. In earlier years, studies on the discovery and identification of crustacean neuroendocrine tissues involved techniques such as tissue ablation, tissue implantation or injection of crude extracts from specific tissues. The changes observed due to these experimental approaches at the macro- and microscopic level, allowed for the discovery of certain neural tissues. One of these tissues, discovered in the eyestalks of the decapod crustacean, is important in the control of glucose metabolism, moulting, reproduction and epithelial pigmentation (Carlisle & Knowles, 1959). These physiological processes were shown to be controlled by a neuroendocrine complex in the eyestalk consisting of the X-organ and sinus gland. In this neuroendocrine complex, the X-organ is the site where neuropeptides are synthesised. The axonal endings from the X-organ project into the sinus gland; this tissue is used as a storage site for these neuropeptides and their release into the circulation. Other neuroendocrine tissues in Crustaceans that are of interest are: the pericardial organs, post-commissural organs and the central nervous system (brain, sub-oesophageal ganglion and the thoracic and abdominal ganglia) (Cooke & Sullivan, 1982; Gäde & Marco, 2006).

1.1.2 Neuropeptides

Neuropeptides are considered one of the most diverse category of extracellular messengers within neural tissues and are secreted into the general circulation of arthropods, called the haemolymph (Gäde & Marco, 2006; Kodrík, 2008). These neuropeptides control and regulate many physiological processes and are used to signal change between cells, tissues and organs from external and internal stimuli. Signalling in insects and crustaceans are controlled by the endocrine and neuroendocrine systems (a system whereby hormones are secreted by neural tissue). The neuroendocrine system is used to send information a lot quicker than the endocrine system. The neuroendocrine response is rapid because the signal is under nervous control mechanism, whereas the endocrine system secretes hormones into the circulation and hormones often act quite a distance away. Together these two systems form a complex functionally integrated system (Gäde & Marco, 2006).

Neuropeptides are responsible for a wide range of important physiological processes such as reproduction, metabolism, growth and development, ion and water balance (Badisco et al., 2007). When comparing peptide sequence data of these neuropeptides it becomes evident that they are conserved across the deuterostomian (characterised by the anus developing before the mouth in the embryo) and the protostomian barrier (characterised by the mouth developing before the anus in the embryo) (Vanden Broeck, 2001). Neuropeptides are grouped into large families based on their mature peptide structure and, although they may have similar structures, neuropeptides can have different (sometimes opposing) functions in other organisms. Two neuropeptide families that control a range of physiological processes are the adipokinetic hormone (AKH)/ red pigment-concentrating hormone (RPCH) family and the parsin family of peptides.

1.2 Animals used in the study

According to a review by Lampert (2006) the genus *Daphnia* (commonly known as the waterflea) come from an ancient crustacean order called Cladocera. It has adapted to many fresh water habitats and has become one of the keystone species in ecosystems.

The species adopts a central role in the aquatic ecosystem by consuming primary producers and being consumed by predators. The species has become a model organism used for ecological and evolutionary studies, because of the numerous amounts of species in the genus and because of how wide spread they are. Some *Daphnia* species, like *D. magna*, are even used to test for ecotoxicological problems in the environment. *Daphnia* reproduces rapidly by sexual but mostly asexual reproduction. This is called cyclic parthenogenesis i.e. a process whereby embryos develop without being fertilised; sexual reproduction is also possible under certain environmental conditions. Eggs produced via sexual reproduction are encased in a durable ephippium that can survive harsh conditions (Lampert, 2006).

N. viridula, commonly known as “the southern green stink bug”, belongs to the suborder Heteroptera or “true bugs” from the insect order Hemiptera. It is a common pest found all over the world. *N. viridula* is an economically important pest because it feeds on legumes like soybeans and beans, but also on macadamia nuts, mangos, orchids, peppers, potatoes, tomatoes and many other plant species. *N. viridula* is a highly mobile insect, which can travel great distance due to flight and in this way can easily reach new agricultural land and damage even more crops. *N. viridula* is also resistant to certain pesticides due to its feeding mode. The mouthparts of *N. viridula* are developed so that they pierce the crop and suck out nutrients; in this way most of the surface sprayed pesticides are avoided.

J. lalandii commonly known as the “West Coast rock lobster”; is a marine crustacean which belongs to a phylogenetically advanced order of crustaceans called Decapoda. *J. lalandii* is of local economic importance in the fishing industry of South Africa. In the wild *J. lalandii* commonly feed on mussels and to a lesser extent kelp and fish.

J. lalandii and *N. viridula* were chosen because they both contain the same mature peptide sequence for the AKH/RPCH hormone, called Panbo-RPCH. This sequence is conserved in all decapod crustaceans analysed thus far; only the primitive crustacean *D. pulex* has been shown to contain a different RPCH, called Dappu-RPCH. It is interesting to investigate whether the AKH/RPCH gene of the ancient crustacean *D. pulex* would

cluster with the AKH/RPCH gene of the insects or the higher crustaceans. The AKH/RPCH gene family will be discussed in more detail below.

1.3 Adipokinetic hormone / Red pigment-concentrating hormone family

In insects and crustaceans one neuropeptide family which has been extensively studied is the AKH/ RPCH family (Gäde, 2004; Gäde & Marco, 2009). Below is a detailed account of the initial discovery of RPCH and later the AKH for better clarity into the origins of the AKH/ RPCH family.

1.3.1 The crustacean RPCH

RPCH was the first member of the AKH/RPCH family of neuropeptides discovered from crustaceans by Fernlund & Josefsson (1972). Later the AKH was discovered in insects; initially in the locust species *Schistocerca gregaria* and *Locusta migratoria* (see 1.3.2). RPCH was isolated from the eyestalks of the shrimp, *Pandalus borealis* and is known now under the acronym Panbo-RPCH. The isolation of the mature peptide was achieved with the use of a gel filtration system and water/butanol mixtures. Elucidation of the structure was achieved by use of high resolution mass spectrometry and Edman-dansyl sequencing (Fernlund & Josefsson, 1972). The primary amino acid structure of Panbo-RPCH is an octapeptide consisting of pGlu-Leu-Asn-Phe-Ser-Pro-Gly-Trp-NH₂ (Fernlund & Josefsson, 1972). The name RPCH stems from its ability to concentrate pigment granules within the integument, allowing the crustacean to change its colour, specifically, turning lighter in colour (Rao, 2001). Colour change is attributed to photomechanical changes within the compound eye, whereby retinal screening of pigment movement in the compound eye allows the crustacean to detect the amount of light in its environment (Rao, 2001). Using this attribute of the eye to screen between varying degrees of light and dark, it is able to adjust its colour to blend into the environment (Rao, 2001). RPCH has also been shown to control the motor output of the swimmeret system (also known as pleopods which are used for swimming movements) (Sherff & Mulloney, 1991) and it has been claimed (but not conclusively proven) to be involved in the mobilisation of energy stores within the isopod crustacean *Porcellio*

scaber (Zralá et al., 2010), as well as stimulate the release of methyl farnesoate from mandibular organs (Rao, 2001).

Since its initial discovery, RPCH has been discovered in many other decapod crustacean species; the amino acid sequence was always identical to that of the mature Panbo-RPCH (Gäde, 2009). A novel RPCH structure is found in the waterflea (order Cladocera) *D. pulex* and *Daphnia magna* (<http://wfleabase.org>), which had its RPCH sequence identified with the use of whole genome sequencing (Christie et al., 2008). The cladoceran mature neuropeptide is called Dappu-RPCH (Figure 1) (Dircksen et al., 2011). The primary structure of Panbo-RPCH is not limited to the crustaceans, it was initially also identified in the hemipteran insect, *N. viridula*, but later in many other insects belonging to the insect suborder Heteroptera (Gäde et al., 2003; Kodrík et al., 2010).

Panbo-RPCH	pGlu	Leu	Asn	Phe	Ser	Pro	Gly	Trp	amide
Dappu-RPCH	pGlu	Val	Asn	Phe	Ser	Thr	Ser	Trp	amide

Figure 1: Comparison between the primary structure of Panbo-RPCH (accession: P08939.2) and Dappu-RPCH (accession: ACD75498.1), the grey highlight represents conserved residues.

Work done by Marco & Gäde (2010) showed that bioassays conducted with Dappu-RPCH in the shrimp *Palaemon pacificus* and the stinkbug *N. viridula*. The authors report that Dappu-RPCH had no effect in concentrating the red, brown, yellow and blue pigments in the epithelial cells of the shrimp at physiological doses. They showed that Dappu-RPCH was, however, able to mobilise lipids in *N. viridula*. Also, they observed that when Val² is changed to Leu² of Dappu-RPCH there was a substantial loss in lipid mobilisation in *N. viridula*. Their work showed that Dappu-RPCH had no effects on a phylogenetically closely related crustacean like the shrimp compared to being able to induce the lipid mobilisation effect of the insect AKH.

1.3.2 The insect AKH

The first RPCH-related neuropeptide discovered in insects came from the migratory locust, *L. migratoria* and is known as Locmi-AKH (Stone et al., 1976). The AKH was so named because it resulted in an increase in the amount of lipids in the haemolymph when released during locust flight (Stone et al., 1976). Locmi-AKH was one of the first insect neuropeptides to be fully characterised and was isolated from the corpora cardiaca of *S. gregaria* and *L. migratoria* by means of mass spectrometric methods and confirmed to a certain extent by Edman-dansyl sequencing (Stone et al., 1976). The primary amino acid structure of Locmi-AKH is a decapeptide consisting of pGlu-Leu-Asn-Phe-Thr-Pro-Asn-Trp-Gly-Thr-NH₂ (Stone et al., 1976). In insects, the AKH is a pleiotropic neuropeptide, because AKH modulates fundamental physiological processes involved in sugar homeostasis, lipid metabolism and reproduction (Lindemans et al., 2009; Michitsch & Steele, 2008).

Many different functions for AKH have been discovered in insects but one of the main functions is to activate glycogen phosphorylase and lipases. These enzymes activate the process by which metabolites, such as diacylglycerols, trehalose and/or proline are mobilised (Gäde, 2004). AKHs are generically known as adipokinetic hormones because the first function attributed to an AKH peptide was to increase haemolymph lipids, later it was also shown to be a hypertrehalosaemic (HrTH: increases haemolymph sugar trehalose) and hyperprolinaemic hormone (increasing proline levels) in other insects (Gäde, 2009; Stone et al., 1976). AKH was shown to stimulate heartbeat, general locomotion, lipid oxidation for the flight muscles, inhibition of RNA synthesis and increase the carrying capacity of lipoprotein carriers as reviewed by Gäde (2004) and Gäde & Marco (2006). Another function of AKH according to observations by Kodrík (2008), is that AKH is involved in the oxidative stress response in locusts. Earlier work conducted by Goldsworthy et al. (2002) reported that AKH was involved in the immune response of locusts, but Kodrík (2008) suggested that the immune response observed could be a result of the oxidative stress response mechanisms in the locusts. These stress response mechanisms work when the immune response is activated by the prophenoloxidase cascade in response to laminarin, an immunogen from algae which contains β -1,3-glucans common in fungal cell walls. This response is also activated

when locusts are co-injected with bacterial lipopolysaccharide (*Escherichia coli*) and Locmi-AKH-I.

Since the initial discovery of the insect AKH in *Locusta*, there have been considerably more AKHs discovered in insects (Gäde, 2009). When comparing the insect AKH and the decapod crustacean RPCH, Panbo-RPCH is highly conserved in decapods. The insect AKH, however, have considerably more heterogeneity, such as variations in chain length (between eight - ten amino acid residues) and intramolecular residue substitutions (Gäde, 2009). The sequence of the mature AKH follows a common formula which includes: a pyroglutamate (stemming from a Gln amino acid which cyclises into the pGlu residue) at the N-terminus, a Phe or Tyr at position four, Trp at position eight and a carboxyamidation at the C-terminus (Gäde & Marco, 2006). When compared with other neuropeptide families, the AKH family does not have a 'core structure' which is essential for the activity of the peptide (Gäde & Marco, 2006). Despite this, certain residues are important for efficacy, which include the pyroglutamate at the N-terminus, the carboxyamidation at the C-terminus and the aromatic residues at positions four and eight (Gäde & Marco, 2006). The structural components that comprise the AKH preprohormones are signal peptide, immature AKH peptide and the AKH-precursor-related peptide (APRP) (Kaufmann et al., 2009) (see example of the yellow fever mosquitoes in Figure 2). In addition, certain insects (e.g. *L. migratoria*) contain up to three different AKH peptides, originating from three separate peptide precursors (Gäde, 2004; Gäde & Marco, 2006).



Figure 2: Diagram illustrating the position of elements that comprise the open reading frame (5' to 3') for the Aedae-AKH-I preprohormone. Abbreviations: AKH: adipokinetic hormone and APRP: AKH-precursor-related peptide. Diagram taken from Kaufmann et al. (2009).

1.3.3 Locust physiology and AKH

In *L. migratoria* the release of AKH was shown to be induced (*in vitro*) by the peptide tachykinin and inhibited by peptides Schgr-FLRFamide and Schgr-FMRamide (Nässel et al., 1995; Vullings et al., 1998). AKH secretion is increased in response to high energy

demanding activities such as insect flight. Preliminary experiments using biological assays, conducted by Cheeseman and Goldsworthy (1979), showed an increase in the AKH titre in *L. migratoria* haemolymph during flight. At the time of their experiments other homologues of AKH had not been discovered (Locmi-AKH-II and III), which would have influenced their results. Radioimmunoassay experiments conducted by Candy (2002) show that Schgr-AKH-I and II increased during flight by fifteen-fold and six-fold, respectively within the haemolymph compared to the control group.

1.4 Mobilisation of energy stores during insect flight

Energy intensive activities such as flight require abundant and quick release of energy stores due to the rapid decline of energy substrates in the haemolymph and flight muscles (Gäde & Auerswald, 2003). The flight muscles of insects function under totally aerobic conditions, which the insect provides with a sophisticated tracheal system (Gäde & Auerswald, 2003). The rate of oxygen consumption is a very good indicator of how well the muscles perform metabolically during intense activities like flight. Observing an increase of oxygen consumption between 50–100-fold is not uncommon during flight (Gäde & Auerswald, 2003). For example, experiments conducted by Auerswald et al. (1998) show that the beetle *Pachnoda sinuata* had an 80-fold increase in oxygen consumption during 2 minutes of tethered flight. Insects store different types of fuels in varying concentrations in the form of either glycogen and/or triacylglycerols in the fatbody. These fuels are broken down into simpler metabolites like glucose/trehalose and diacylglycerols (DAG)/ free-fatty acids (FFA) respectively; this process occurs when energy stores become limiting within flight muscle tissue (Gäde & Auerswald, 2003; Van der Horst & Rodenburg, 2010). As well as using trehalose and DAG, certain insects use the amino acid proline in combination with trehalose as fuel for flight (Gäde & Auerswald, 2002; Auerswald et al., 1998). These metabolites are either liberated or resynthesized to maintain and prolong flight (Gäde, 1992; Gäde & Auerswald, 1998; Gäde & Auerswald, 2003). Two independent studies done in 1976 by Spencer & Candy and Gäde & Holwerda on locusts, discovered that the first messenger AKH works via the second messenger cyclic AMP (cAMP) in order to mobilise lipids from fatbody cells. However, this is not the case in all insects. In the cockroach

Periplaneta americana, trehalose is the main energy source liberated by Peram-AKH-I, and proline is the main energy source in the beetle *P. sinuata* (Auerswald & Gäde, 2001). Energy metabolism and substrate mobilisation were best studied in the following insect model species: *L. migratoria*, *P. americana*, *P. sinuata*, *Manduca sexta* and *Blaberus discoidalis*. The details of the signal transduction pathways for hypertrehalosaemia, hyperlipaemia and hyperprolinaemia in these insect models in response to flight are shown in Figure 3-Figure 5.

The AKHR is an integral part of the AKH signal transduction pathway. It was initially discovered in the fruit fly *Drosophila melanogaster* and in the silkworm *Bombyx mori* as a class of G protein-coupled receptors (Park et al., 2002; Staubli et al., 2002). The AKHR is a membrane bound protein containing 7 transmembrane segments (Ziegler et al., 2011). AKHR was shown to be expressed in fatbody and in certain nervous tissues (Ziegler et al., 2011; Wicher et al., 2006). The route the signal travels from the first messenger AKH to the second cAMP is via the AKH receptor protein and the coupled G protein either Gs or Gq (Figure 3).

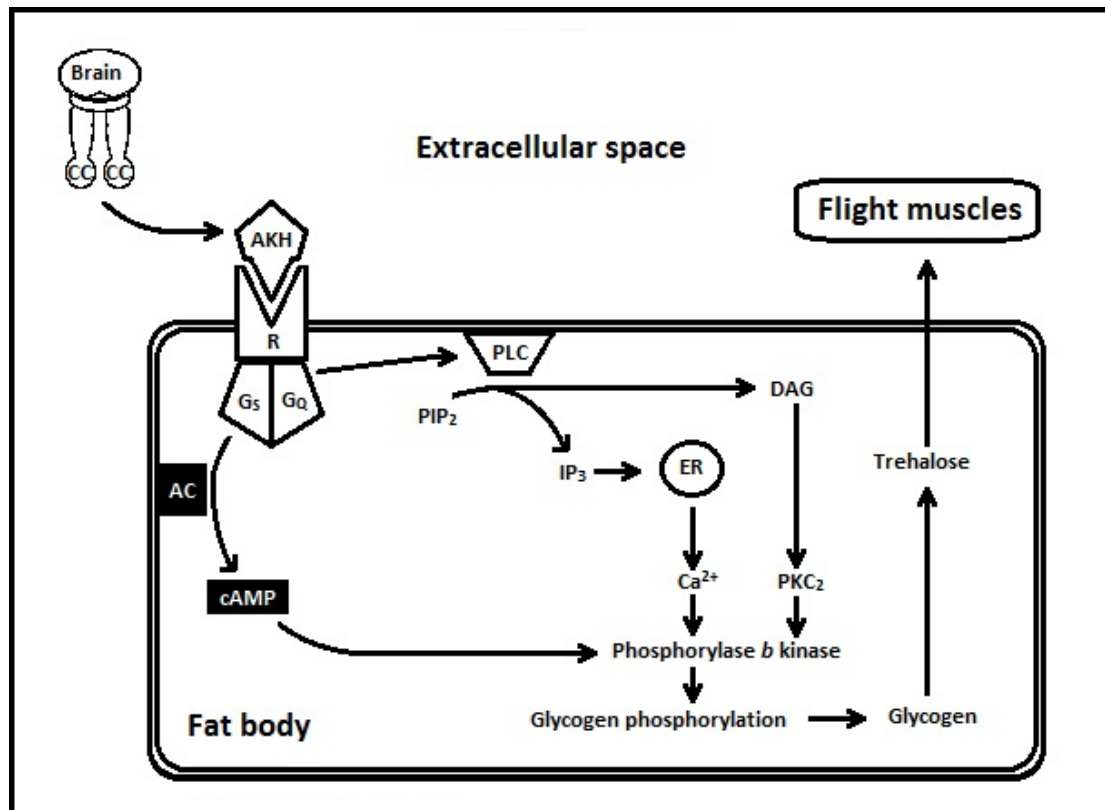


Figure 3: A general overview of AKH controlled mobilisation of trehalose in the fatbody of insects (specifically studied in models *P. americana*, *P. sinuata* and *L. migratoria*) during high energy requiring activities such as flight. Abbreviations - CC: corpus cardiacum; AKH: adipokinetic hormone; R: AKH receptor; G_s/G_q: G-protein; PLC: phospholipase C; PIP₂: phosphatidylinositol bisphosphate; DAG: diacylglycerol; IP₃: inositol triphosphate; ER: endoplasmic reticulum; AC: adenylate cyclase; cAMP: cyclic AMP; PKC₂: protein kinase C and black-boxed AC and cAMP represent a specific pathway in *L. migratoria*. Adapted from Gäde & Auerswald, 2003; Van der Horst & Rodenburg, 2010.

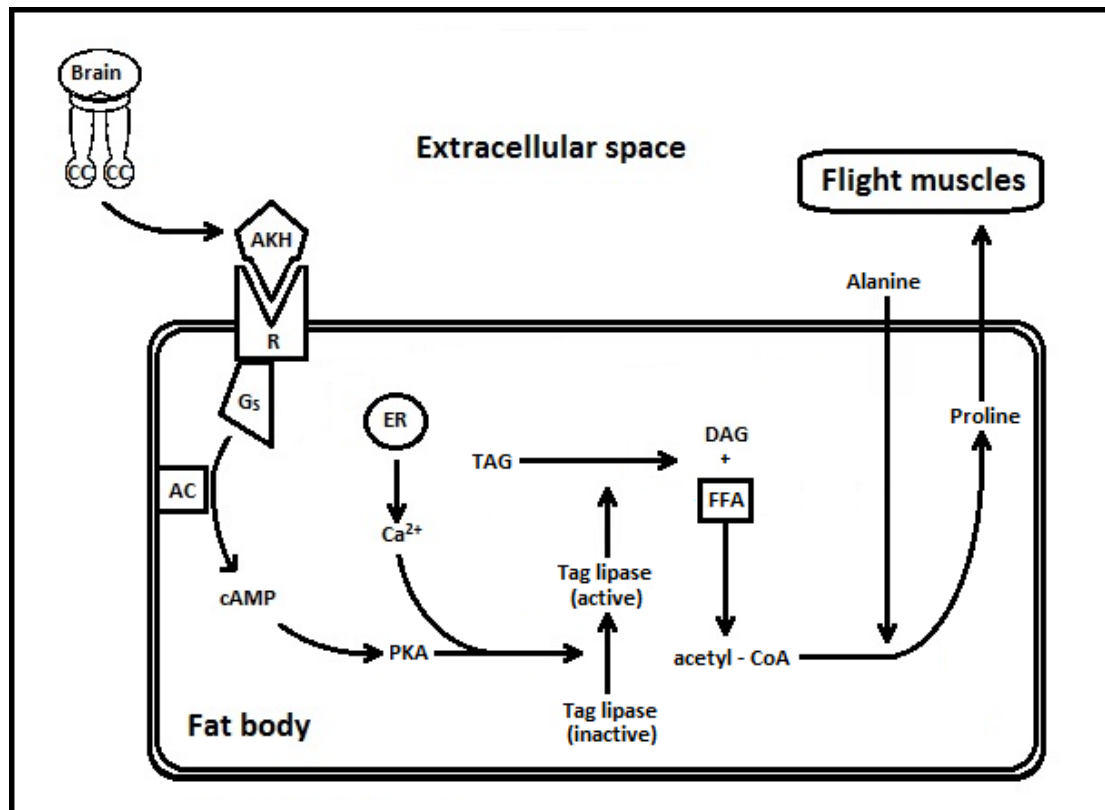


Figure 4: A general overview of AKH controlled mobilisation of proline in the fatbody of the beetle, *P. sinuata* during flight when carbohydrate energy stores are depleted. Abbreviations - CC: corpus cardiacum; AKH: adipokinetic hormone; R: AKH receptor; G_s: G-protein; DAG: diacylglycerol; TAG: triacylglycerol; ER: endoplasmic reticulum; AC: adenylate cyclase; cAMP: cyclic AMP; PKA: protein kinase A. Adapted from Gäde & Auerswald, 2003.

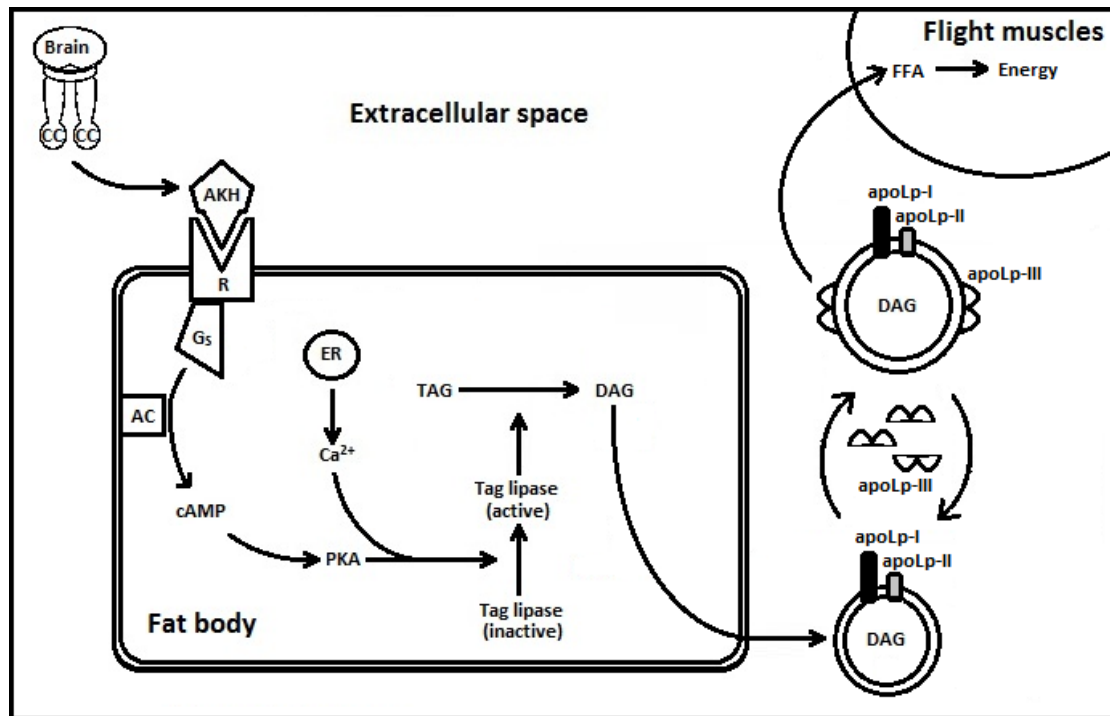


Figure 5: An overview of AKH controlled mobilisation of lipids (DAG) in the fatbody of locusts during flight when carbohydrate energy stores are exhausted. DAG lipids transported to the flight muscle via lipophorins are broken down into FFA by lipophorin lipase. Abbreviations - CC: corpus cardiacum; AKH: adipokinetic hormone; R: AKH receptor; G_s: G-protein; DAG: diacylglycerol; TAG: triacylglycerol; ER: endoplasmic reticulum; AC: adenylate cyclase; cAMP: cyclic AMP; PKA: protein kinase A and apoLp I-III: apolipoprotein I-III. Adapted from Gäde & Auerswald, 2003; Van der Horst & Rodenburg, 2010.

1.5 Introduction to the parsin neuropeptide family

Another neuropeptide family in insect endocrinology, which regulates a variety of functions, is the parsin family of hormones. The parsin neurohormone family, was initially discovered in the pars intercerebralis-corpora cardiaca neurohaemal complex of locusts (Girardie et al., 1989). Parsins were initially associated with insect reproduction, but later it appeared that parsins could also be linked to and considered as markers of locust 'phase transition' (Ayali et al., 1996). Phase transition refers to the phenomenon whereby the desert locust can exist in two extreme phases, i.e. the solitary and the gregarious phase (De Loof et al., 2006; Ayali et al., 1996). The gregarious phase in locust is associated with swarming behaviour and has been the cause for concern in the agricultural sector in Africa and Asia (Badisco et al., 2007). The phase transition between the two states results in behavioural, reproductive and morphological (colour, pattern, size, weight, development) changes (Pener & Yerushalmi, 1998; Bouaïchi et al., 1995; Collett et al., 1998). The parsin neurohormone family is comprised of the following neuropeptides: the insulin-related peptides (IRP - also commonly referred to as 'insulin-like peptides'), the ovary maturing parsin (OMP), pacifastin and neuroparsin (NP) (Badisco et al., 2007). Detailed below is a brief summary of the OMP, IRP and pacifastin members of the parsin family with special emphasis placed on NP, specifically in locust species.

1.5.1 Insulin-related parsin

The first IRP substance discovered in an invertebrate, came from the bivalve mollusc *Mya arenaria*, in 1923 (Collip, 1923). Since the discovery of the first IRP and with the help of bioassay and immunocytochemical detection methods, many more IRPs have been identified (Badisco et al., 2007). The first IRP discovered in insects came from the silk moth *B. mori* and was given the name bombyxin. Bombyxin in another moth species (e.g. *Samia cynthia ricinii*) was isolated because it displayed prothoracicotropic hormone (PTTH) activity, which stimulates the prothoracic gland to produce ecdysone, inducing ecdysis (moulting). Bombyxin in *B. mori*, however, does not display PTTH activity, rather, moulting is controlled by the PTTH hormone itself (Ishizaki & Suzuki,

1994). Bombyxin, like insulin, exist as heterodimers (IRP-A and IRP-B), structural comparison showed that they are homologous and were, thus, considered a member of the insulin super family (Nagasawa et al., 1984). Since the discovery of bombyxin, many more IRPs have been identified in insects and have been shown to be important in a variety of physiological processes, which include reproduction and cell and tissue growth, as reviewed by Claeys et al. (2002).

1.5.2 Ovary maturing parsin

The ovary maturing parsin from *L. migratoria* (Locmi-OMP) was the first gonadotropin discovered in insects (Girardie et al., 1991) and was later also found in the desert locust *S. gregaria* OMP (Schgr-OMP) (Girardie et al., 1998). Locmi-OMP stimulates oocyte development and induces the synthesis of vitellogenin (an egg yolk) in locusts. Thus far, OMPs have only been found in the insect family Acrididae, which contain the genera *Locusta* and *Schistocerca* (Badisco et al., 2007; Richard et al., 1994).

1.5.3 Pacifastin

The pacifastin family of neuropeptides was named according to the crustacean species *Pacifastacus leniusculus* in which it was discovered (Hergenhahn et al., 1987). The *P. leniusculus* pacifastin was shown to be a heterodimeric protein with a heavy (105 kDa) and a light chain (44 kDa) synthesised from two different mRNAs. The heavy chain of the pacifastin contains domains which have homology to transferrin proteins, which have an iron binding function. The light chain is comprised of similar repeating homologous structure of 9 inhibitory cysteine-rich domains (also called the 'Pacifastin light chain domains') (Liang et al., 1997). The members of the pacifastin family are classified as strong inhibitors of serine proteases and are only found in arthropods (Breugelmans et al., 2009).

1.5.4 Neuroparsin

1.5.4.1 Neuroparsin in locusts

A peptide synthesised within the pars intercerebralis and subsequently transported and stored within the CC was first purified from the CCs of the migratory locust *L. migratoria* and called neuroparsin (Girardie et al., 1987). Neuroparsin (NP) from *L. migratoria* was isolated using anion exchange chromatography and sequenced using mass spectrometric methods (Girardie et al., 1989). Locmi-NPA is 85 AA residues long; four isoforms of NPA (namely NPA I-IV, where NPA I = NPA and NPA IV = NPB) are present in the CC (Hietter et al., 1991). Locmi neuroparsin B (NPB) contains 78 AA residues which is identical to a section of NPA AA sequence and is thus believed to be a post-translationally shortened form of NPA (Badisco et al., 2007). It appears that the four isoforms of NPA (NPA I-IV) are intermediary products as NPA is post-translationally modified to form NPB (Badisco et al., 2007). Another NPA was discovered from the CC of the desert locust *S. gregaria* (Schgr-NPA) which is functionally similar to *L. migratoria* NPA (Girardie et al., 1998). Schgr NPA consists of 83 AA and is processed stepwise at the N-terminus which results in the formation of NPB (Girardie et al., 1998). Neuroparsins are cysteine-rich monomeric neurohormones which contain six intramolecular disulphide bridges formed from twelve cysteine residues (Girardie et al., 1989; Girardie et al., 1998; Hietter et al., 1991). When comparing the amino acid sequences of the Locmi-NPA and Schgr-NPA prohormones (Figure 6), 15 AA substitutions can be seen in the alignment. Overall there is 86% sequence identity between the two and when the C-terminus is compared to the N-terminus the sequence homology is higher at the N-terminus, 97% compared to 78% (Girardie et al., 1998). The comparison shows how variable NP is, even between closely related organisms.

Locmi NPA	<u>MKATAALVAATLLLA</u> VTLFHRAERNPISRCEGANCVVDLTRCEYGDVTDFFGRKVCAKG	60
Schgr NPA	<u>MKPAAALAAATLLLA</u> VTLFHRAEANPISRCEGANCVVDLTRCEYGEVTDFFGRKVCAKG	60
Locmi NPA	PGDKCGGPYELHGKCGVGMDCRCGLCSGCSLHNLQCCFFEGGLPSSC	107
Schgr NPA	PGDKCGGPYELHGKCGDGMDCRCGVCSGCSMQSLECCFFEGGAAPNSC	107

Figure 6: An alignment of *L. migratoria* (Locmi-NPA, accession: P10776) and *Schistocerca gregaria* neuroparsin prohormone (Schgr-NPA, accession: Q966Y1). The Alignment was generated using the ClustalW algorithm from www.uniprot.org (all default settings). Mismatched AA are highlighted grey, underlined AA represent the signal peptide and the green highlighted AA are conserved cysteine residues.

Locmi-NPA was found to influence many phenotypic traits and effects in *L. migratoria* such as anti-diuretic (Fournier et al., 1987), anti-gonadotropin, affecting larval development (Girardie et al., 1989), neuritogenic (stimulates neurite outgrowth) (Vanhems et al., 1990), hypertrehalosaemic and hyperprolinaemic effects (Moreau et al., 1988). It was initially discovered, however, due to inhibitory effects observed upon vitellogenesis (egg-yolk protein synthesis for oocyte growth) (Girardie et al., 1998; Girardie et al., 1989). NPs appear to generally inhibit or slow down sexual maturation and reproduction in locusts. This effect is in contrast to the action of juvenile hormone (JH), which accelerates sexual maturation and induces vitellogenesis. The effect of accelerated sexual maturation is identical to the result observed when NP is immunoneutralised by NP antiserum in locusts (Girardie et al., 1987). Also, injection of NP antiserum into locust larvae resulted in premature moulting of the larvae into adult intermediary forms (Girardie et al., 1989). Furthermore, treatment with either NP antiserum or NP did not influence the transcript levels of JH, which suggests that JH works independently of NP. However, treatment with JH or ecdysone increased the transcript levels of NP (Claeys et al., 2006). These results at least support a role of NP in reproduction and development (Veenstra, 2010).

There are numerous papers which suggest that NP has an anti-diuretic effect in locust. However, according to Coast et al. (2002) the bioassays used to determine the anti-diuretic effects of NP did not survive scrutiny because the doses needed to produce this effect were unphysiologically high. Also, apparently NP has no effect on ion transport across the rectal epithelium, which is a requirement for fluid uptake. Vanhems et al. (1990) observed that, at least in locusts, NP appears to act as a neuritogenic factor because treatment of the central nervous systems of embryos with NP, increased neuritic outgrowth significantly. NP was also found to induce an increase in trehalose and lipids in the haemolymph of *L. migratoria* and was suggested to likely work together with AKH to regulate carbohydrate metabolism (Moreau et al., 1988), but this effect has never been confirmed since (Veenstra, 2010).

1.5.5 Neuroparsin homologous genes in *L. migratoria* and *S. gregaria*

Initially, only one type of NP was characterized in *L. migratoria* (Locmi-NPA and its isoforms), later a similar peptide was discovered in *S. gregaria* (Girardie et al., 1998). Furthermore, three distinct NP precursor (NPP) transcripts were identified in *S. gregaria* (Schgr-NPP2-4) and two from *L. migratoria* (Ma et al., 2006; Claeys et al., 2003). This is not to be confused with the intermediary products NPA I-IV, whereby NPA (also known as NPA-I and NPP1) is converted through intermediary products (NPA II-III) into NPB (also known as NPA-IV). Schgr-NPP1 and Schgr-NPP2 seem to be restricted to the brain and fatbody of *S. gregaria* (Claeys et al., 2003; Janssen et al., 2001); whereas Schgr-NPP3 and Schgr-NPP4 were detected via Northern blot analysis in the brain, ventral nerve cord, fatbody and muscle tissue of both male and female locusts, as well as in the male accessory organs and testes (Claeys et al., 2003). Also, *in silico* searches in locustdb.genomics.org.cn, which contains expressed sequence tags (ESTs) of different body parts from both solitary and gregarious locusts, helped to verify the presence of the different transcripts (Ma et al., 2006).

1.5.6 Neuroparsin-like peptides in other organisms

The presence of neuroparsin-like peptides (NPLP) was also shown in other organisms by using immunocytochemical methods (using anti-NPA serum); NPLPs were detected in a variety of insect orders, which include: Odonata (e.g. dragonflies), Dictyoptera (e.g. termites and cockroaches), Orthoptera (e.g. locusts) and Diptera (e.g. mosquitoes) (Tamarelle & Girardie, 1989; Bourême et al., 1989). Tamarelle & Girardie (1989) demonstrated (via immunostaining) that NPLPs were also present in two annelid species.

1.5.6.1 Ovary ecdysteroidogenic hormone

The first insect gonadotropin was discovered in mosquitoes by Lea (Lea, 1967; Lea, 1972), who described a neurohormone that controlled reproduction and egg maturation in female mosquitoes of the species *Aedes taeniorhynchus*. The reproductive cycle in female mosquitoes begins with the ingestion of a blood meal; this stimulates the

release of gonadotropins from neurosecretory cells in the brain into the haemolymph (Lea, 1967; Lea, 1972). The gonadotropins, in turn, stimulate the ovaries to secrete ecdysteroids (a hormone which induces moulting) and were therefore termed “ovary ecdysteroidogenic hormones” (OEH) (Matsumoto et al., 1989). The OEH precursor consists of 149 AA residues; the signal peptide is 22 AA residues long and is followed by the mature peptide sequence of 83 AA residues (Badisco et al., 2007; Brown et al., 1998). The highest sequence identity of OEH is to Locmi-NPA (29%). Both peptides share eight similarly positioned conserved cysteine residues which are signature properties of the NP family (Figure 7). OEH was identified in *Aedes aegypti* and *Anopheles gambiae*, localised (using anti-OEH) in the cells of the brain, ventral nerve cord and the gut of both sexes (Brown & Cao, 2001). The presence of OEH in the axonal network of the hindgut of *A. gambiae* suggests that the peptide might be associated with the regulation of water balance and perhaps also the movement of food through the gut (Brown & Cao, 2001).

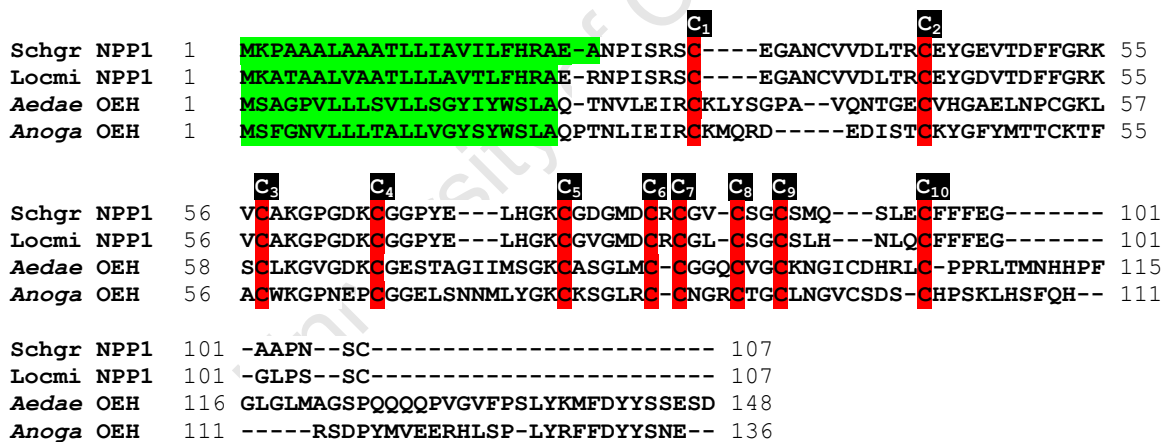


Figure 7: An alignment showing *L. migratoria/ S. gregaria* NP precursor 1 (NPP) vs. *A. aegypti/ A. gambiae* “ovary ecdysteroidogenic hormone (OEH). The green box represents the signal peptide region; the conserved cysteine residues are in red boxes numbered C1-C10.

1.5.6.2 Insulin-like growth factor binding proteins

Amino acid homology searches in vertebrates show that there are sequences which appear to have a similar distribution of cysteine residues as the NPs. These peptides are the “insulin-like growth factor binding proteins” (ILGFbps) (Claeys et al., 2003). Badisco

et al. (2007) suggest that vertebrate ILGFEBPs belong in the NP family of peptides due to similarity at the N-terminus and similarly arranged cysteine residues. It was also shown that in *S. gregaria* NP has an insulin binding function (Badisco et al., 2008). Although there are similarities between NPs and ILGFEBPs, Veenstra (2010) disputes this claim. The author states that the similarities between NP and ILGFEBP are only superficial when comparing the distribution of the cysteine residues. In addition, and perhaps more convincing, is the fact that both the red flour beetle *Tribolium castaneum*, as well as the bee *Apis mellifera*, contain the ILGFEBP-like peptide and the NP peptide. Another argument of Veenstra is that the insulin binding proteins are typically not expressed in neuroendocrine cells as are the NP.

1.5.6.3 Locust Phase transition

Phase transition refers to the two extreme phases/ states that locusts can adopt: the solitary and the gregarious phase (for a list of other locusts which show phase transition refer to review by Pener & Yerushalmi, (1998)). The gregarious phase leads to formation of massive locust swarms at irregular intervals between ten – twenty years (Badisco et al., 2007). Locusts are also able to travel great distances when in the gregarious state and, in the process, destroy many crops. This poses a big threat to the African agricultural sector where crops can be wiped out entirely. The last major locust plague that occurred, lasted for three years (Showler & Potter, 1991). Phase transition from the solitary to gregarious phase is induced by crowding between the first – third generations, when crowded locusts become increasingly stressed (Ayali et al., 1996). Phase transition results in behavioural and physiological changes, e.g. colour, size, pattern, weight, development and reproduction (Pener & Yerushalmi, 1998; Collett et al., 1998; Ayali et al., 1996). Desert locusts in the gregarious state mature sexually much slower than isolated individuals (Girardie et al., 1987). Due to the clear differences in reproductive physiology between the two states and the antagonistic effects that NP has on locust reproduction, it seemed likely that NP plays a part in the phase transition (as mentioned in section 1.5.4.1). Also, because locusts migrate over great distance in their gregarious phase and require large amounts of energy, it seems very likely that AKH plays a major role in maintaining this state. According to De Loof et al. (2006) when looking at reproduction within the different phases, each phase would require specific

energy regulations. From the spatial and temporal differences of the Schgr-NPP transcript levels in gregarious and solitary locusts, there is evidence that Schgr-NPP transcripts are finely tuned during locust adult life in a reproductive phase dependant manner. For example, in the fatbody Schgr-NPP1 and Schgr-NPP2 mRNA transcripts are expressed in the solitary state, whereas in the gregarious state Schgr-NPP3 and Schgr-NPP4 mRNA transcripts are chiefly expressed.

When the energy profiles are compared for locust in the solitary and gregarious state at rest, the gregarious state has a significantly higher haemolymph lipid concentration compared to locust in the solitary state and this is reversed when crowded fledgling locusts are isolated (Ayali & Pener, 1992; Ayali & Pener, 1995). Although phase transition has not been shown in other insects and is not the focus of the study, it does show how NP and AKH are involved in reproduction and energy balance respectively in insects.

1.6 Aims of the study

The two arthropod subphyla, insects and crustaceans are closely related sister groups. Investigation of their neuroendocrine peptides will further our knowledge of how closely related they are. The AKH/RPCH gene family was selected because in insects the AKH gene is quite variable compared to the crustacean RPCH analysed so far. For this reason, it is a good candidate to investigate the similarities and/or differences between the two subphyla and identify whether *D. pulex* is more closely related to the insects or the younger crustaceans based on this gene family. Although other genes (18s rDNA, 16s rDNA, cytochrome c oxidase I, etc.) have been used in the study of arthropods, my study aims to determine how well neuropeptides explain the relatedness between crustaceans and insects. Also, by using more than one neuropeptide to investigate homology between the crustaceans and insects a better picture of this relationship may emerge.

As said previously, all AKH/RPCH peptide sequences obtained thus far from crustaceans are well conserved (Panbo-RPCH), with exceptions from *D. pulex* and *D. magna* (Dappu-RPCH), and is poorly conserved in insects (Fernlund & Josefsson, 1972; Marco & Gäde, 2010). One particular insect, the Hemipteran *N. viridula* contains the same mature AKH/RPCH peptide sequence common in decapod crustaceans (namely Panbo-RPCH). Subsequently it was uncovered that Panbo-RPCH is common to the Pentatomidae family (also known as shield bugs) (Gäde et al., 2003; Kodrık et al., 2010). The comparison of the AKH/RPCH preprohormones of three organisms *D. pulex* (Cladocera), *N. viridula* (Heteroptera) and *J. lalandii* (Decapoda) might reveal whether the phylogenetically ancient crustacean *D. pulex* clusters closer to insects like *N. viridula* or to the evolutionarily advanced crustaceans like *J. lalandii*. This investigation would help in understanding the link between the insects and crustaceans using the neuropeptide AKH/RPCH preprohormone.

The overall aim of this project was to determine whether neuropeptides of the *D. pulex* are more homologous to insects or crustaceans. Specifically, I aimed to identify the AKH/RPCH nucleotide sequences encoding the preprohormone sequence of the decapod crustacean *J. lalandii*, a hemipteran insect *N. viridula* and verify the Dappu-

RPCH preprohormone and Dappu-RPCHR sequence from *D. pulex* (German ecotype). Locally grown Daphnids in the Western Cape, South Africa were sourced because of the ease of attaining animals for future experiments. As a result, the Dappu-RPCH and Dappu-RPCHR of the locally grown daphnids needed to be verified so that it could be determined whether their sequences were similar enough for use in future experiments. In the process of trying to identify the AKH preprohormone sequence, the NP neuropeptide prohormone was fortuitously amplified in *N. viridula* and *J. lalandii*. This finding allowed for comparison of another set of neuropeptides with which to assess the degree of homology between the two subphyla. According to literature, by using northern blot analysis the expression of NP was discovered in the CC, as well as in other non-neural tissues, e.g. locust fatbody, (Claeys et al., 2003; Girardie et al., 1987). We wanted to investigate the spatial distribution of NP in *J. lalandii* and *N. viridula* and to identify whether the spatial expression in the different tissues was conserved as well. As the function of Dappu-RPCH in *D. pulex* is unknown, we aimed to identify the tissues that expressed the Dappu-RPCHR mRNA transcript in order to extrapolate its putative function.

2. Materials and Methods

2.1 Animals

2.1.1 *Nezara viridula* (Southern green stink bug)

Adults and nymphs of *N. viridula* of both sexes and of unspecified age were collected from *Agapanthus africanus* plants found in gardens within the Southern Suburbs of the Western Cape Province of South Africa. Collection took place during the austral summer and autumn months (December to April 2009 to 2012) by Prof G. Gäde, Dr. H.G Marco and I. The bugs were maintained in a constant environmental chamber in the Zoology department of the University of Cape Town (UCT) under the following conditions: 28 °C ± 2 °C, austral summer light conditions of 17 h light : 7 h dark and a relative humidity of approximately 50% (obtained by filling large trays with water).

Eggs and nymphs up to the third instar were reared in petri dishes (10cm diameter). Thereafter, nymphs were moved to glass tanks (30 cm length x 20 cm width x 20 cm height and fitted with a mesh lid). Once stinkbugs reached the adult stage, both males and females were moved to another glass tank with the same dimensions and allowed to breed and lay eggs. Immature adults were defined as one - two days after moulting into the adult form and mature adults as three weeks after moulting into the adult form. Stinkbugs were fed shelled sunflower seeds which were replaced every two days. Bugs in the petri dishes were supplied water via moist cotton wool, while a 50 ml falcon tube filled with water and fitted with a cotton plug provided water to bugs in the glass tanks.

2.1.2 *Jasus lalandii* (South African west coast spiny lobster)

Adult spiny lobsters, both male and female were supplied by the Department of Agriculture, Forestry and Fisheries. They were collected from the Atlantic Ocean along the west coast of South Africa and transported to the Zoology aquarium at UCT by Dr. Marco and Prof. Gäde. The spiny lobsters were held in tanks with dimensions: width x 30 cm, height x 30 cm and either length x 30 cm (small) or 60 cm (medium) or 90 cm (large). Tanks contained circulating seawater (replaced weekly) at temperatures of 12 - 14 °C and aerated by pipes supplying air. The animals were fed twice a week with pilchards and subjected to a photoperiod of 12 h light: 12 h dark.

2.1.3 *Daphnia* spp. (waterflea)

2.1.3.1 *Daphnia pulex*

Adult *D. pulex*, both male and female, were originally collected from streams and river beds in Gräfenhain, Sachsen, Germany. From this ecotype *D. pulex* was reared in the laboratory in Germany by Dr. Bettina Zeiss (Zoology Department, University of Münster, Germany), who kindly supplied Dr. Marco with live specimens. Whole *D. pulex* were immersed in RNAlater® (Applied Bioscience – Ambion®) and transported to the Zoology Department, UCT, where it was stored at -20 °C until use.

2.1.3.2 *Daphnia* spp. (South African ecotype)

Daphnia spp., both male and female, were obtained from Aquaterra a pet shop in the Western Cape Province, South Africa, and transported on ice to the Zoology Department, UCT. Animals were kept in water supplied by Aquaterra at 4 °C for three days until usage.

2.2 Dissections

Before tissues of the different animals were dissected, instruments were cleaned and treated with RNaseZap® (Ambion®) and rinsed in 75% ethanol. The dissected tissues were transferred to a 1.5 ml Eppendorf tube and flash frozen in liquid nitrogen. Tissues were stored at -80 °C until homogenisation in Total RNA Isolation Reagent® (ABgene).

2.2.1 Dissection of *Jasus lalandii*

Spiny lobsters were anaesthetised in a mixture of ice and seawater and eyestalks were ablated. The microscopic X-organ was immediately dissected from eyestalks with the aid of a dissecting microscope (30 fold magnification) and flash frozen. Other tissues that were dissected from *J. lalandii* were: abdominal muscle, heart, gills, hepatopancreas, thoracic ganglion, sub-oesophageal ganglion, brain, suboesophageal ganglion, ovaries and testis.

2.2.2 Dissection of *Nezara viridula*

Different tissues were dissected from mature versus immature males and females for experiments involving NP expression and localisation: suboesophageal ganglion, ventral nerve cord, corpora cardiaca, brain, whole leg, flight muscle, mid- and hindgut, antennae, fatbody, male reproductive organs and female reproductive organs. Tissues were dissected using a Dumont® No. 5 microfine tweezers and ophthalmological microscissor under a dissecting microscope (30 fold magnification).

2.2.3 Dissection of *Daphnia pulex*

For the studies where Dappu-RPCH or the RPCH receptor needed to be amplified the whole animal of the South African ecotype was used. For experiments where the spatial distribution of the RPCH and RPCHR need to be ascertain the following tissues were dissected: antennae, eyes/ brain, gut/ ovaries, muscle/ ovary, body cuticle (Figure 8). Tissues were dissected using a Dumont® No. 5 microfine tweezers and ophthalmological microscissor under a dissecting microscope (30 fold magnification).

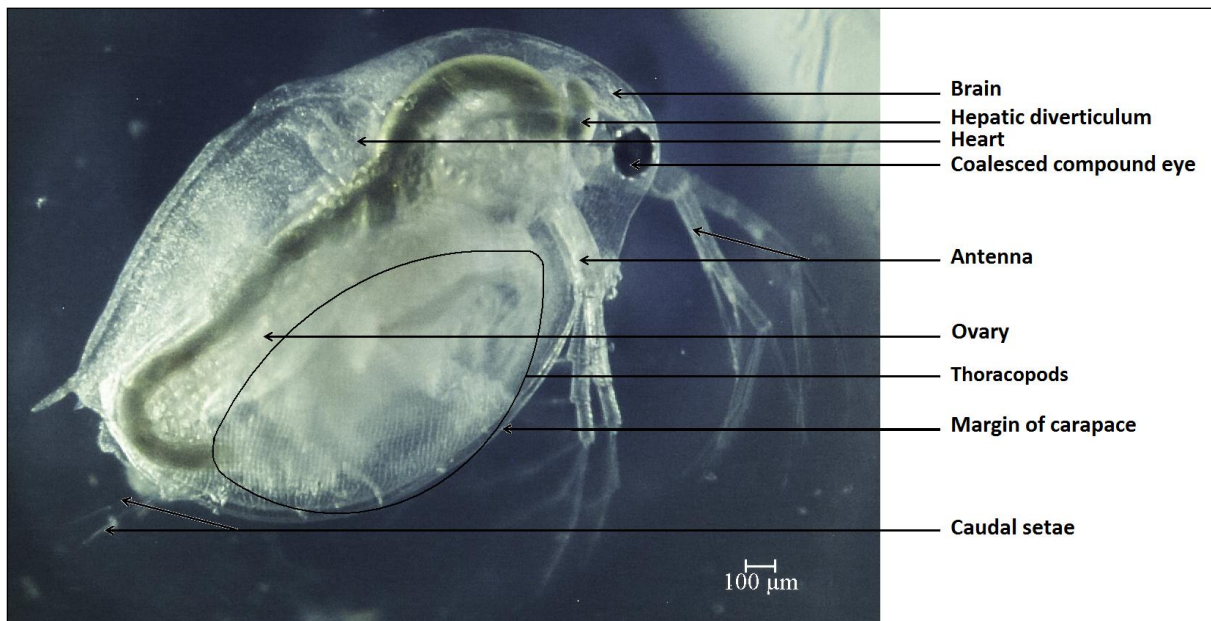


Figure 8: Diagram illustrating certain body parts of *D. pulex* (South African ecotype). This image was generated in the study.

University of Cape Town

2.3 DNA techniques

2.3.1 Genomic DNA extraction method

DNA was extracted for the purposes of isolating the AKH/ RPCH from *N. viridula* and *J. lalandii* gene. A single specimen of *N. viridula* (head and legs) and 10-20 mg *J. lalandii* abdominal muscle tissue were used and placed into 1.5 ml Eppendorf tube. Tissues were manually homogenized in 160 μ l extraction buffer (200 mM TrisCl pH 8, 70 mM EDTA pH 8, 2 M NaCl) using a plastic drill bit. Further disruption and solubilisation of membrane samples was achieved by addition of 40 μ l of 5 % sarkosyl solution and subsequent incubation for 1 h at 60 °C. The cell lysate was centrifuged for 15 min at 16,000 g (all centrifugation steps were done at room temperature). The resulting supernatant was pipetted into another sterile 1.5 ml Eppendorf tube. Phenol (pH 8): chloroform: isoamyl alcohol (50: 24: 1) v/v solution was added to the cell lysate in a 1: 1 ratio. Samples were mixed by inversion and centrifuged at 16,000 g for 5 min. The upper aqueous phase was pipetted into a sterile 1.5 ml Eppendorf tube and sodium acetate (3 M, pH 5.2) was added in a 0.1: 1 ratio to the supernatant. Ethanol was added in a 2: 1 volume ratio to the supernatant and incubated at -70 °C for 30 min. The DNA was precipitated by centrifugation at 16,000 g for 10 min, and the resulting pellet was washed with 70 % ethanol and centrifuged at 16,000 g for 5 min. This step was repeated twice. The pellet was resuspended in 25 μ l of TE buffer (10 mM TrisCl pH 8, 1 mM EDTA). DNA samples were quantified and the quality was checked using the Nanodrop spectrophotometer (NanoDrop Technologies, Wilmington, USA) and stored at -20 °C (Cheung et al., 1993).

2.3.2 Splinkerette PCR system

2.3.2.1 The splinkerette PCR system

The splinkerette polymerase chain reaction (PCR) system described by Devon et al (1995) is an advancement on the vectorette PCR system. A brief explanation about its function is given here. The vectorette system involves using a double stranded linker sequence with a central mismatch and a cohesive sticky end able to ligate with DNA that has been digested with a suitable restriction enzyme (Figure 9A). The problem with a vectorette system is that it produces many non-specific products, which arise from 'end repair' priming. End repair priming occurs when unligated cohesive ends of vectorettes and inserts are filled in, during the first cycle of PCR. These ends have enough stability to initiate priming and amplify non-specific products during PCR. The splinkerette (Figure 9B) is a modified vectorette, whereby the mismatch is replaced with a hairpin loop on the bottom strand (splkecol). This modification reduces end repair and end priming in splinkerettes, which is achieved by the hairpin loop forming during the first cycle of PCR and initiating self-priming. The self-primed (splkecol) structure is stable and is functionally removed from the reaction leaving the top strand in the reaction (splktop). The linker-specific primers (splkO and splkI) are only able to bind to the splktop's complimentary nucleotide sequence after the initial cycle of PCR.

2.3.2.2 The splinkerette PCR system construction and reaction

The splinkerette PCR system was used to probe DNA from *N. viridula* for the Nezvi-AKH gene. Construction of the splinkerette was achieved as follows: 150 ng/ μ l of each oligonucleotide, splnktop and splkecol (Table 1) and 1x splinkerette buffer (10 mM TrisCl, pH 7.4 and 5 mM MgCl₂) to a final reaction volume of 20 μ l. The splinkerette solution was heated to 90 °C for 10-20 min and allowed to cool to room temperature. Genomic DNA between 200-500 ng was digested using 2 U of *Eco*RI (Roche) and 1x sure/cut buffer H (Roche) in a total volume of 20 μ l at 37 °C. Following overnight digestion, *Eco*RI was heat inactivated at 60 °C for 10 min. The ligation of digested DNA and splinkerette was carried out as follows: 2 μ l *Eco*RI digested genomic DNA, 6 μ l of

splinkerette solution, 1x T4 ligase buffer (Fermentas), 1 U T4 DNA ligase (Fermentas) and filled to 20 µl with dH₂O and incubated at 4 °C overnight.

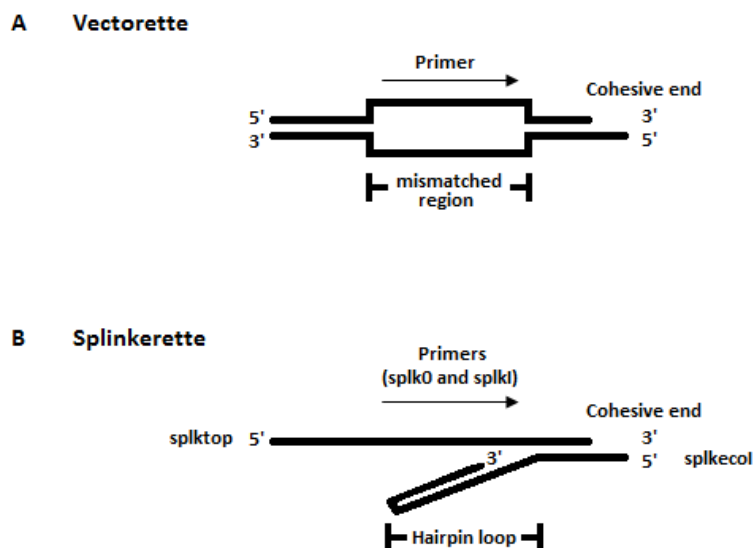


Figure 9: A modified diagram taken from Devon et al. (1995) of the (A) vectorette and (B) splinkerette system, which shows the different attributes associated with each linker system.

Table 1: Oligonucleotides used in splinkerette construction and the splinkerette PCR system to amplify the Nezvi-AKH gene.

Primer name	Oligonucleotides (5' - 3')	Reference	Length (bp)
splktop	CGA ATC GTA ACC GTT CGT ACG AGA ATT CGT ACG AGA ATC GCT GTC CTC TCC AAC GAG CCA AGA	(Devon et al., 1995)	63
splk0	CGA ATC GTA ACC GTT CGT ACG AGA A		25
splk1	TCG TAC GAG AAT CGC TGT CCT CTC C		25
splkecol	AAT TTC TTG GCT CGT TTT TTT TTG CAA AAA		30
Nvakh	TVA MNT TYW CNC CNN VYT GG	MSA	20
NvDeg1	GCN CAG CTV AMV TTY WC		17
NvDeg2	CAR CTB AAY TTY TCW CCB GGM TGG GG		26
28srDNAf	CCC GTC TTG AAA CAC GGA CCA A	(Kavar et al., 2007)	22
28srDNAr	CCA CAG CGC CAG TTC TGC TTA C		22

Primers were designed based on insect AKH multiple sequence alignments

Legend:

N - Either A, C, G, C **V** - Either A, C, G **B** - Either C, G, T **Y** - Either C, T **W** - Either A, T

M - Either A, C

MSA - Derived from a multiple sequence alignment

2.4 Polymerase chain reaction

All PCR amplifications were performed in either the GeneAmp PCR Systems 2700 (Applied Biosystems, Foster City, USA) or the Eppendorf® mastercycler® personal (Perkin-Elmer Corporation).

2.4.1 Splinkerette PCR

The attempted amplification of AKH preprohormone was performed in 20 µl reactions containing: 0.2 µmol splk0 in the first PCR (0.2 µmol of splkI in the second round of nested PCR), 0.4 µmol Nvakh primer (0.2 µmol of NvDeg1 or NvDeg2 in the second round of nested PCR), 1x reaction buffer (Promega), 2 µM dNTPs (Fermentas), 1.5 mM MgCl₂, 1 U Super-therm DNA polymerase (Promega) and 50-100 ng of *N. viridula* template DNA. The splinkerette PCR cycling conditions used were as follows: initial denaturation at 94 °C for 4 min, followed by 15 cycles of (94 °C for 15 s, 70 °C for 1 min decreasing by 1 °C, 72 °C for 1 min), 20 cycles of (94 °C for 15 s, 55 °C for 1 min, 72 °C for 2 min), final elongation at 72 °C for 10 min.

2.4.2 28s ribosomal DNA amplification

Genomic DNA quality was assessed by amplifying a portion of the 28s ribosomal DNA sequence from *N. viridula*; the reaction was performed in 20 µl containing: 0.2 µmol of 28srDNAf and 28srDNAr primers, 1x reaction buffer (Promega), 2 µM dNTPs (Fermentas), 1.5 mM MgCl₂, 1 U Super-therm DNA polymerase (Promega) and 50-100 ng DNA. Cycling conditions were as follows: initial denaturation at 94 °C for 5 min, followed by 35 cycles of (94 °C for 30 s, 60 °C for 30 s and 72 °C for 1 min) and a final elongation step of 72 °C for 10 min (Kavar et al., 2007).

2.4.3 Agarose gel electrophoresis

PCR products were loaded on a 2% w/v agarose gel containing: 100 ml of 1x Tris-acetate (TAE) buffer (40 mM Tris, 20 mM acetic acid and 1 mM EDTA pH 8.0), 0.2 mg/ml ethidium bromide and 2 g of agarose (Lonza, SeaKem®) together with a 1 kb

DNA ladder (O' GeneRuler™ #SM1163 from Fermentas). Agarose gels were run in 1x TAE buffer for approximately 1.5 h, at 120 V and visualized on short wavelength (260 nm) UV transilluminator. PCR bands were excised on a long wavelength (360 nm) UV transilluminator using a surgical blade. DNA was extracted from the agarose gel slice using the Wizard® SV gel and PCR cleanup system (Promega Corporation, Madison, USA) according to the manufacturer's instructions.

2.4.4 Ligations

DNA bands purified from agarose gels were cloned into the sequencing vector pGEM®-T Easy vector system (Promega) following the manufacturer's instructions with modifications. The reaction was carried out with a 3: 1 (insert: vector) ratio, determined using the formula:

$$\frac{25 \text{ ng vector} \times \text{insert length (kb)}}{3 \text{ kb vector}} \times \frac{3 \text{ insert}}{1 \text{ vector}} = \text{Mass DNA required}$$

The ligation mixture contained: 25 ng pGEM®-T Easy vector, 1 x T4 ligase buffer, 1.5 Weiss units T4 ligase and dH₂O to 5 µl total volume and was incubated at room temperature for 1.5 h.

2.4.5 Transformation

The transformation was conducted using 50 µl of competent DH5α *Escherichia coli* cells to 5 µl of the pGEM®-T Easy vector/insert mixture after ligation. The transformation mixture was mixed by inversion and placed on ice for 20 min. The mixture was heat-shocked for 45-50 s at 42 °C and immediately placed on ice thereafter. 500 µl of Luria Bertani broth (LB) (Sambrook et al., 1989) was added to the transformation mixture and incubated at 37 °C for 2 h with shaking. A 100 µl of the transformed cells were spread-plated on LB agar plates containing: 100 µg/ml ampicillin (Amp), 0.5 M isopropyl-β-D-thiogalactopyranoside (IPTG) and 80 µg/ml 5-bromo-4-chloro-2-indolyl-β-D-galactoside (X-gal), and incubated for 16 h at 37 °C. Successfully transformed *E. coli* colonies (white colonies) were selected using the X-gals blue/white selection for subsequent experiments.

2.4.6 Plasmid DNA extraction

White colony forming units (cfu) were inoculated into 5 ml LB and Amp 100 µg/ml and grown at 37 °C for 16 h. The plasmids were purified from the cells using the BioSpin plasmid DNA extraction kit (Bioflux - Bioer Technology Co., Ltd.) and eluted in 40 µl of elution buffer. Extracted plasmids were digested with *EcoRI* (Promega) as follows: 2 µg plasmid DNA, 1x buffer H, 10 U *EcoRI* and dH₂O to 20 µl. The digested plasmids were visualized on a 2% w/v agarose gel described above. Positive recombinant plasmids were sent to Macrogen Incorporated (Korea) for sequencing using M13F and /or M13R primers.

2.4.7 Bioinformatics

Sequence data were analysed using the DNAMAN (Lynnon, Quebec, Canada) and BioEdit bioinformatic tools (Hall, 1999). Homology searches were conducted using the Blast® programs, namely Blastn, Blastp and Blastx from the National Center for Biotechnology Information (<http://blast.ncbi.nlm.nih.gov/>). Phylogenetic trees and molecular evolutionary analyses were inferred using the minimum evolution method in MEGA version 5 (Tamura et al., 2011) and bootstrapping was conducted with 5000 replicates. The substitution method used in MEGA 5 was the Poisson model when designing phylogenetics trees.

2.5 RNA Techniques

2.5.1 RNA extraction

Total RNA was isolated from insect and crustacean tissues using Total RNA Isolation Reagent® (TRIR) (ABgene) following the manufacturer's recommended protocol. Up to 100 mg tissues were briefly homogenized in 1 ml TRIR reagent in Eppendorf tubes with a plastic pestle. Total RNA was resuspended in 40 µl of 0.1% DEPC (diethylpyrocarbonate) treated water and stored at -80 °C.

2.5.2 Total RNA quality and quantity determination

Total RNA was analyzed using the Nanodrop spectrophotometer (Nanodrop Technologies, Wilmington, USA). Total RNA quality was considered of high enough quality for further analyses with a ratio above 1.8 for the absorbance readings 260/280 nm.

2.5.3 DNase treatment of total RNA

DNase treatment was performed following the manufacturer's suggestions and modified as follows: each DNase reaction tube contained 1.5 - 2.0 µg of total RNA, 1x reaction buffer (Fermentas), 1.5 U DNase I (Fermentas) and DEPC treated water to 13 µl total volume and was incubated at 37 °C for 30 min. The treatment was stopped by addition of 2 mM EDTA (Fermentas) and incubated at 65 °C for 10 min. Following DNase treatment, 1 µl of the total RNA mixture was tested for DNA contamination by attempting to amplify the actin gene in a PCR reaction before cDNA synthesis (refer to section 2.5.6.3 for actin PCR).

2.5.4 First strand cDNA synthesis

The cDNA synthesis reaction was performed following the SuperScript™ III (Invitrogen™) protocol with the following modification: a mixture was prepared using 1.5 - 2.0 µg total RNA, 0.5 µg primer (oligo dT, oligo dT-anchor or gene specific reverse primer [GSRP] 1 see Table 2), 1 µM DTT and DEPC treated water up to 12 µl total volume, and incubated at 70 °C for 10 min. Thereafter, the RNA mixture was immediately placed on ice for 2 min. For cDNA synthesis the following reagents were added to the denatured RNA mixture: 1x SuperScript™ III buffer (Invitrogen™), 0.5 mM dNTPs mix, 5 U/µl of SuperScript™ III reverse transcriptase (Invitrogen™), and incubated at 42 °C for 2 h, followed by 10 min incubation at 72 °C. After cDNA synthesis, samples were kept at -20 °C until used to conduct actin PCR (cDNA integrity check) and amplification of target experiment genes.

2.5.5 3' Rapid Amplification of cDNA Ends PCR system

2.5.5.1 A brief explanation

The 3' rapid amplification of cDNA ends (RACE) PCR system involves synthesising cDNA using oligo dT-anchor primer (abridged anchor primer [AAP] or Roche 8 adapter [R8]) which binds to the mRNA polyadenosine-tail (Table 2). Following cDNA synthesis, PCRs involve a gene specific forward primer (GSFP) and the oligo dT-anchor specific primer (Abridged Universal Amplification Primer [AUAP] or Roche 9 [R9]) (Figure 10).

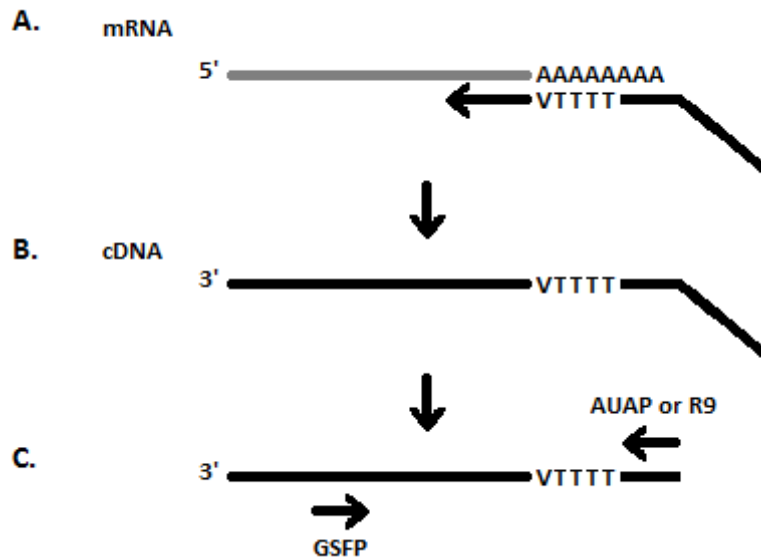


Figure 10: Diagram illustrating the 3' RACE PCR system (A) cDNA synthesis using an oligo dT-anchor primer and the mRNA as grey lines and DNA/cDNA are shown as black lines. (B) Shows the product of cDNA synthesis with attached adapter sequence. (C) PCR amplification of the gene of interest using a gene specific primer (GSFP) and anchor specific primer Abridged Universal Amplification Primer (AUAP) or Roche 9 (R9) (Adapted from the Roche 3'/5' RACE kit manual).

Table 2: Primers used in general and 3'/5' RACE PCR.

Organism	Gene name	Primer name:	Sequence 5' - 3':	Order of use	Ta(°C)*	PCR or RACE method
<i>D. pulex</i>	RPCHR	DpRPCHRr2	TTA AAA TAT ATG TGT GAC GAC AGT TGG	GSRP 1	ND	UI5R
		DpGnRHRr	CGT CAG CAG TGG CAA CCG GTA G	GSRP 2	52	
		DpRPCHRr3	GAG AGC AAA TGA CGG AAC AAG C	GSRP 3		
		DpRPCHRf	ATG TCG ACA GTT CAG TTT GAC G	GSFP 1		UI3R
		DpGnRHRf	ACG ATT TCC AAC GCC TGC GCT T	GSFP 2		
		DpRPCHRr2	TTA AAA TAT ATG TGT GAC GAC AGT TGG	GSRP 2		
	RPCH	Dpr2	CAT CCT GCC ATC TAG CGG TG	GSRP 1		ND
		Dpr3	GAC CAA CAA GAG AGA TTT CAG	GSRP 2	55	
		Dpr4	GGA TAG GCA AGA CTT GGT AGA G	GSRP 3		
<i>J. lalandii</i>	RPCH	JlRPCHf	ACT GCA ACA GGT GAG GCC A	GSFP 1	52	UI3R
		Jlf2	ATG GTT CGT GCC GGC GTC GCC C	GSFP 2		
<i>N. viridula</i>	NP	NvNpr1	TGG GGT CGC ACT CGA TGA TGC	GSRP 1	ND	UI5R
		NvNpr2	CTT GGA GGA TGC GGT CGT TGC TG	GSRP 2	55	
		NvNpr3	GCA CTT GCC GCA CTT GCA GGT C	GSRP 3		
		NvAKH	TVA MNT TYW CNC CNN VYT GG	GSFP 1	ND	UI3R
		NvDeg1	GCN CAG CTV AMV TTY WC	GSFP 2	ND	
		NvDeg2	CAR CTB AAY TTY TCW CCBG GMT GGG G	GSFP 3	ND	
	Adapter	Roche 8	GAC CAC GCG TAT CGA TGT CGA C (T)16 V	Oligo dT-anchor primer	ND	UI3R & UI5R
		AAP	GGC CAC GCG TCG ACT AGT AC (T)17			
	Adapter primers	R9	GAC CAC GCG TAT CGA TGT CGA C	RACE primer	ND	
		AUAP	GGC CAC GCG TCG ACT AGT AC		ND	

Primers were designed based on partial cDNA sequence information obtained from experiments or from genomic databases.

Legend:

* - Annealing temperature for primer set

ND - Not determined

UI5R - Used in 5' Race PCR

UI3R - Used in 3' Race PCR

GSFP - Gene specific forward primer

GSRP - Gene specific reverse primer

2.5.5.2 *Jasus lalandii* RPCH 3' RACE PCR

J. lalandii (X-organ) cDNA was synthesised according to the cDNA synthesis protocol in section 2.5.4 with oligo dT-adaptor (R8). *J. lalandii* RPCH 3' RACE semi-nested PCR one was conducted as follows: 1x MyTaq™ reaction buffer (Bioline) (contains dNTPs and MgCl₂), 0.025 U/μl MyTaq™ DNA polymerase (Bioline), 0.5 μM J1RPCHf and R9, 200 ng *J. lalandii* X-organ and dH₂O filled to 50 μl (for primer sequences refer to Table 2). Cycling conditions were as follows: initial denaturation at 95 °C for 2 min, followed by 35 cycles of 95 °C for 15 s, 51 °C annealing for 30 s and 72 °C for 10 s, concluding with a final elongation step of 72 °C for 10 min.

The second part of the 3' RACE semi-nested PCR was conducted as follows: 1x MyTaq™ reaction buffer (Bioline), 0.025 U/μl MyTaq™ DNA polymerase (Bioline), 0.5 μM J1f2 and R9, 1 μl of a 1/20 dilution of the "3' RACE semi-nested PCR 1" and filled to 50 μl with dH₂O. Cycling conditions were as follows: initial denaturation at 95 °C for 2 min, followed by 35 cycles of 95 °C for 15 s, 55 °C annealing for 30 s and 72 °C for 10 s, and a final elongation step of 72 °C for 10 min.

2.5.5.3 *Daphnia pulex* RPCHR 3' RACE PCR

The amplification of the 3' end of the *D. pulex* RPCHR was conducted in a nested PCR reaction. The first PCR reaction in the nested PCR experiment was conducted as follows: 1x MyTaq™ reaction buffer (Bioline), 0.025 U/μl MyTaq™ DNA polymerase (Bioline), 0.5 μM DpRPCHR and R9, cDNA containing the R8 Roche linker and dH₂O filled to 50ul. The cycling conditions were as follows: initial denaturation at 95 °C for 2 min, followed by 35 cycles of 95 °C for 15 s, 52 °C annealing for 30 s, 72 °C for 10 s and a final elongation step of 72 °C for 10 min. The first nested PCR was cleaned with the Wizard® SV gel and PCR cleanup system (Promega Corporation, Madison, USA) and 1/20 dilution was used as the template DNA in the second nested PCR reaction with the same conditions mentioned for the first nested PCR but with primers DpGnRHRf and DpRPCHRr2.

2.5.6 5' Rapid Amplification of cDNA Ends PCR system

2.5.6.1 A brief explanation

The 5' RACE PCR system involves the use of three gene specific reverse primers (GSRP 1, GSRP 2, GSRP 3) (although two primers can be used but with reduced specificity) to amplify genes of interest from mRNA (Table 2) . As illustrated in Figure 11, the primer GSRP 1 is used to reverse transcribe gene specific mRNA into cDNA, which is column purified after cDNA synthesis. A poly-A tail is added to the 5' end of the cDNA. In the subsequent PCR, cDNA is turned into dsDNA using primers GSRP 2 and an oligo dT-anchor primer (AAP or R8). In the first round of PCR, the oligo dT-anchor primer (AAP or R8) binds to the poly-A tail and acts as a forward primer while the GSRP 2 anneals to the gene specific sequence for the reverse primer. In the final PCR GSRP 3 and the linker specific primer (AUAP or R9) are used to amplify the target gene specifically (Figure 11).

University of Cape Town

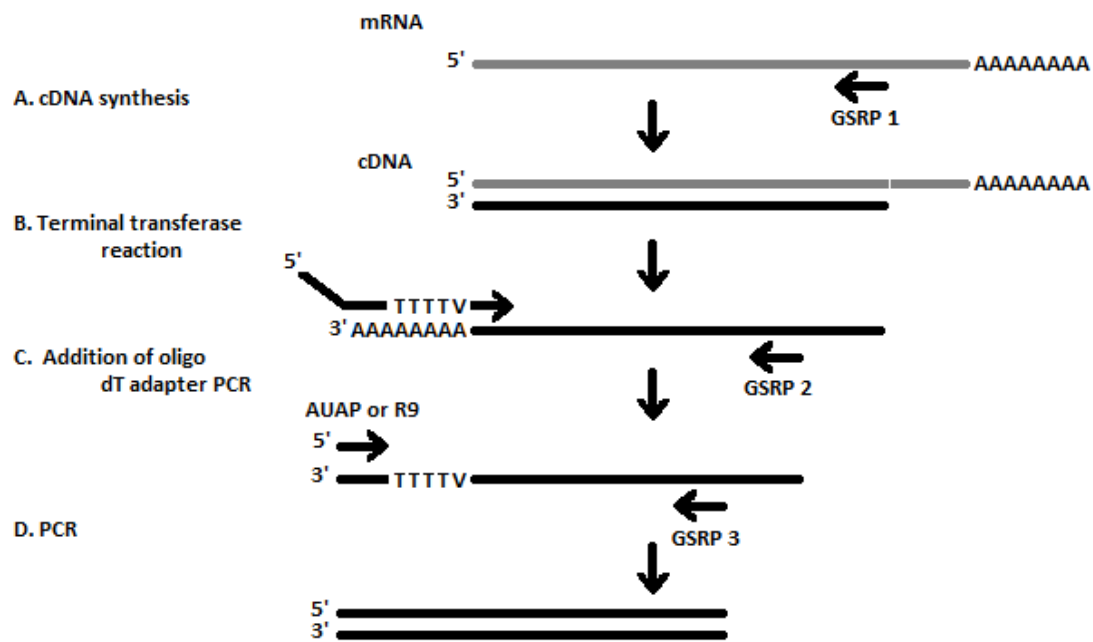


Figure 11: Diagram illustrating 5' RACE PCR, (A) cDNA synthesis from total mRNA (mRNA are represented as grey lines and DNA/cDNA as black lines) using gene specific reverse primer (GSRP) 1, (B) addition of poly-dA tail at 5' end using the terminal transferase reaction and subsequent PCR amplification using the oligo dT-anchor primer (AAP or R8) and GSRP 2, (C) PCR amplification of gene of interest using Abridged Universal Amplification Primer (AUAP) or Roche 9 (R9) with GSRP 3 (Adapted from the Roche 3'/5' RACE kit manual).

2.5.6.2 5' RACE PCR reaction

5' RACE PCR (Roche) first strand cDNA synthesis was conducted as per section 2.5.4 and using the GSRP 1 primer (Table 2). After first strand cDNA synthesis the sample was cleaned through the Wizard® SV PCR clean-up system (Promega Corp, Madison, WI, USA) according to centrifugation speeds and times suggested by the 5' RACE kit instructions (Roche) and cDNA was eluted in 20 µl nuclease free water. The terminal transferase reaction was performed following the 5' RACE kit instructions (Roche).

Addition of the oligo dT-adapter primer (AUAP or R8) to the 5' poly dA-tail cDNA was conducted as follows: 5 µl of the terminal transferase reaction mix, 0.2 µM AAP or R8, 0.5 µM GSRP 2, 1x MyTaq™ buffer (Bioline), 0.025 U/µl MyTaq™ DNA polymerase (Bioline) and dH₂O up to 50 µl. Cycling conditions were as follows: initial denaturation at 95 °C for 2 min, followed by 10 cycles of 95 °C for 15 s, 52 °C for 30 s and 72 °C for 40 s; followed by 25 cycles of 95 °C for 15 s, 52 °C for 30 s and 72 °C for 40 s +20 s/cycle and a final elongation step of 72 °C for 10 min.

Nested PCR was performed to amplify the gene of interest and was conducted as follows: 1 µl of oligo dT-adapter/GSRP 2 PCR (referring to reaction directly above), 0.5 µM AUAP or R9 primers, 0.5 µM GSRP 3, 1x MyTaq™ buffer (Bioline), 0.025 U/µl MyTaq™ DNA polymerase (Bioline) and dH₂O up to 50 µl. Cycling conditions were as follows: initial denaturation at 95 °C for 2 min, followed by 35 cycles of 95 °C for 15 s, 52 °C for 30 s and 72 °C for 40 s and a final elongation step of 72 °C for 10 min. Nested PCR's were processed according to sections: 2.4.3 - 2.4.6.

2.5.6.3 Polymerase chain reaction

PCR amplification of the actin genes (described directly below) from *D. pulex*, *J. lalandii* and *N. viridula* was initially achieved using degenerate primers, Deg actf and Deg actr (Chaty et al., 2004) and yielded a 900 bp fragment, which was analysed according to sections 2.4.3 to 2.4.7. After acquiring the 900 bp actin sequences using the degenerate primers, a new set of specific primers were designed to amplify a smaller portion of the actin sequence (Table 3).

Specific PCR reaction details, particularly the annealing temperature (T_a), elongation time (T_e) and cDNA template used as well as the expected size are listed in Table 3 and the common reaction conditions are listed here. The general reaction conditions are as follows: 1 x MyTaq™ reaction buffer (Bioline), 0.025 U/ μ l MyTaq™ DNA polymerase (Bioline), 0.5 μ M of forward and reverse primers, 0.2 μ M cDNA and dH₂O up to either 20 μ l or 50 μ l. Cycling conditions were as follows, initial denaturation at 95 °C for 2 min, followed by 35 cycles of 95 °C for 15 s, (T_a) °C for 30 s and 72 °C for (T_e) and a final elongation step of 72 °C for 10 min. Genes were further analysed and sequenced following sections 2.4.3 to 2.4.6.

University of Cape Town

Table 3: Primers to amplify target genes in general PCR

Organism	cDNA Template	Gene of interest	Primer names	Primer sequence (5'-3')	Expected size (bp)	Ta (°C)	Te (s)																																																																															
	any tissue	Partial generic actin	Deg actf	GGA YGA YAT GGA GAA RAT CTG G	900	55	30																																																																															
			Deg actr	CCT GYT TGC TGA YCC ACA TCT G				<i>D. pulex</i>	whole body	Actin	Dappu actf	TGT GCG ACG ATG AAG TAG CAG C	452	54	30	Dappu actr	TAC CGG TGG TAC GAC CGG AAG C	RPCH	Dpf	GCT ATT GAT GAT CGG CCT GGC C	435	52	10	Dpr	TAC GGA TGT AAA GCC TAC ATG TCC T	RPCHR	DpGnRHRf	ACG ATT TCC AAC GCC TGC GCT T	1404	52	60	DpRPCHR2	TTA AAA TAT ATG TGT GAC GAC AGT TGG	<i>J. lalandii</i>	X-organ	Actin	Jlact f	TGC TAC TGA CGG AGG CTC CCC TC	458	60	10	Jlact r	GAA CCT CTC GTT GCC GAT GGT GA	cHH-II	JlcHHIIIf	ATG GCC CAC AGA ACG ATG	400	50	10	JlcHHIIr	CTT GCC CAC CAT CTG TAC	RPCH	JIRPCHf	ACT GCA ACA GGT GAG GCC A	197	52	5	JIRPCHr	AAG TTG AGC TGG GCG GAG A	JIRPCHf	ACT GCA ACA GGT GAG GCC A	297	60	10	JIRPCHr5	AGG CCT GGT GAT GGT GGG AG	Neuroparsin	JINPf	CCA GGT TGG ACG CAG CAC CAC	334	60	10	JINpr	GTG CCA AGT GCT GCC ACT GGT AG	<i>N. viridula</i>	corpus cardiacum	Actin	Nvactin f	CTC CCC TCA ACC CTA AAG CCA AC	444	52	30	Nvactin r	GGA ACC TCT CGT TAC CGA TGG TG	Neuroparsin	NvNPf2
<i>D. pulex</i>	whole body	Actin	Dappu actf	TGT GCG ACG ATG AAG TAG CAG C	452	54	30																																																																															
			Dappu actr	TAC CGG TGG TAC GAC CGG AAG C						RPCH	Dpf	GCT ATT GAT GAT CGG CCT GGC C	435	52	10	Dpr	TAC GGA TGT AAA GCC TAC ATG TCC T	RPCHR	DpGnRHRf	ACG ATT TCC AAC GCC TGC GCT T	1404	52	60	DpRPCHR2	TTA AAA TAT ATG TGT GAC GAC AGT TGG	<i>J. lalandii</i>	X-organ	Actin	Jlact f	TGC TAC TGA CGG AGG CTC CCC TC	458	60	10			Jlact r	GAA CCT CTC GTT GCC GAT GGT GA	cHH-II	JlcHHIIIf	ATG GCC CAC AGA ACG ATG	400	50	10	JlcHHIIr	CTT GCC CAC CAT CTG TAC	RPCH	JIRPCHf	ACT GCA ACA GGT GAG GCC A	197	52	5		JIRPCHr	AAG TTG AGC TGG GCG GAG A	JIRPCHf	ACT GCA ACA GGT GAG GCC A	297	60	10	JIRPCHr5	AGG CCT GGT GAT GGT GGG AG	Neuroparsin	JINPf	CCA GGT TGG ACG CAG CAC CAC	334	60	10	JINpr	GTG CCA AGT GCT GCC ACT GGT AG	<i>N. viridula</i>	corpus cardiacum	Actin	Nvactin f	CTC CCC TCA ACC CTA AAG CCA AC			444	52	30	Nvactin r	GGA ACC TCT CGT TAC CGA TGG TG	Neuroparsin	NvNPf2	TCC CAG AGA TCG AGA TGT ACT CC	350	54
		RPCH	Dpf	GCT ATT GAT GAT CGG CCT GGC C	435	52	10																																																																															
			Dpr	TAC GGA TGT AAA GCC TAC ATG TCC T						RPCHR	DpGnRHRf	ACG ATT TCC AAC GCC TGC GCT T	1404	52	60	DpRPCHR2	TTA AAA TAT ATG TGT GAC GAC AGT TGG	<i>J. lalandii</i>	X-organ	Actin	Jlact f	TGC TAC TGA CGG AGG CTC CCC TC	458	60	10			Jlact r	GAA CCT CTC GTT GCC GAT GGT GA	cHH-II	JlcHHIIIf	ATG GCC CAC AGA ACG ATG	400			50	10	JlcHHIIr	CTT GCC CAC CAT CTG TAC	RPCH	JIRPCHf	ACT GCA ACA GGT GAG GCC A	197	52	5		JIRPCHr	AAG TTG AGC TGG GCG GAG A	JIRPCHf	ACT GCA ACA GGT GAG GCC A	297	60	10	JIRPCHr5	AGG CCT GGT GAT GGT GGG AG	Neuroparsin	JINPf	CCA GGT TGG ACG CAG CAC CAC	334	60	10	JINpr	GTG CCA AGT GCT GCC ACT GGT AG	<i>N. viridula</i>	corpus cardiacum	Actin	Nvactin f	CTC CCC TCA ACC CTA AAG CCA AC	444			52	30	Nvactin r	GGA ACC TCT CGT TAC CGA TGG TG	Neuroparsin	NvNPf2	TCC CAG AGA TCG AGA TGT ACT CC	350	54	10	NvNpr4	TCC GGC AGA TTT AAG AGC AGA GC			
		RPCHR	DpGnRHRf	ACG ATT TCC AAC GCC TGC GCT T	1404	52	60																																																																															
			DpRPCHR2	TTA AAA TAT ATG TGT GAC GAC AGT TGG				<i>J. lalandii</i>	X-organ	Actin	Jlact f	TGC TAC TGA CGG AGG CTC CCC TC	458	60	10	Jlact r	GAA CCT CTC GTT GCC GAT GGT GA			cHH-II	JlcHHIIIf	ATG GCC CAC AGA ACG ATG	400	50	10			JlcHHIIr	CTT GCC CAC CAT CTG TAC	RPCH	JIRPCHf	ACT GCA ACA GGT GAG GCC A	197			52	5	JIRPCHr	AAG TTG AGC TGG GCG GAG A		JIRPCHf	ACT GCA ACA GGT GAG GCC A	297	60	10	JIRPCHr5	AGG CCT GGT GAT GGT GGG AG	Neuroparsin	JINPf	CCA GGT TGG ACG CAG CAC CAC	334	60	10	JINpr	GTG CCA AGT GCT GCC ACT GGT AG	<i>N. viridula</i>	corpus cardiacum	Actin	Nvactin f	CTC CCC TCA ACC CTA AAG CCA AC	444	52	30			Nvactin r	GGA ACC TCT CGT TAC CGA TGG TG	Neuroparsin	NvNPf2	TCC CAG AGA TCG AGA TGT ACT CC	350	54	10	NvNpr4	TCC GGC AGA TTT AAG AGC AGA GC											
<i>J. lalandii</i>	X-organ	Actin	Jlact f	TGC TAC TGA CGG AGG CTC CCC TC	458	60	10																																																																															
			Jlact r	GAA CCT CTC GTT GCC GAT GGT GA						cHH-II	JlcHHIIIf	ATG GCC CAC AGA ACG ATG	400	50	10	JlcHHIIr	CTT GCC CAC CAT CTG TAC			RPCH	JIRPCHf	ACT GCA ACA GGT GAG GCC A	197	52	5			JIRPCHr	AAG TTG AGC TGG GCG GAG A		JIRPCHf	ACT GCA ACA GGT GAG GCC A	297	60	10	JIRPCHr5	AGG CCT GGT GAT GGT GGG AG	Neuroparsin	JINPf	CCA GGT TGG ACG CAG CAC CAC	334	60	10	JINpr	GTG CCA AGT GCT GCC ACT GGT AG	<i>N. viridula</i>	corpus cardiacum	Actin	Nvactin f	CTC CCC TCA ACC CTA AAG CCA AC	444	52	30	Nvactin r	GGA ACC TCT CGT TAC CGA TGG TG			Neuroparsin	NvNPf2	TCC CAG AGA TCG AGA TGT ACT CC	350	54	10	NvNpr4	TCC GGC AGA TTT AAG AGC AGA GC																					
		cHH-II	JlcHHIIIf	ATG GCC CAC AGA ACG ATG	400	50	10																																																																															
			JlcHHIIr	CTT GCC CAC CAT CTG TAC						RPCH	JIRPCHf	ACT GCA ACA GGT GAG GCC A	197	52	5	JIRPCHr	AAG TTG AGC TGG GCG GAG A				JIRPCHf	ACT GCA ACA GGT GAG GCC A	297	60	10	JIRPCHr5	AGG CCT GGT GAT GGT GGG AG	Neuroparsin	JINPf	CCA GGT TGG ACG CAG CAC CAC	334	60	10	JINpr	GTG CCA AGT GCT GCC ACT GGT AG	<i>N. viridula</i>	corpus cardiacum	Actin	Nvactin f	CTC CCC TCA ACC CTA AAG CCA AC	444	52	30	Nvactin r	GGA ACC TCT CGT TAC CGA TGG TG			Neuroparsin	NvNPf2	TCC CAG AGA TCG AGA TGT ACT CC	350	54	10	NvNpr4	TCC GGC AGA TTT AAG AGC AGA GC																															
		RPCH	JIRPCHf	ACT GCA ACA GGT GAG GCC A	197	52	5																																																																															
			JIRPCHr	AAG TTG AGC TGG GCG GAG A							JIRPCHf	ACT GCA ACA GGT GAG GCC A	297	60	10	JIRPCHr5	AGG CCT GGT GAT GGT GGG AG	Neuroparsin	JINPf	CCA GGT TGG ACG CAG CAC CAC	334	60	10	JINpr	GTG CCA AGT GCT GCC ACT GGT AG	<i>N. viridula</i>	corpus cardiacum	Actin	Nvactin f	CTC CCC TCA ACC CTA AAG CCA AC	444	52	30	Nvactin r	GGA ACC TCT CGT TAC CGA TGG TG			Neuroparsin	NvNPf2	TCC CAG AGA TCG AGA TGT ACT CC	350	54	10	NvNpr4	TCC GGC AGA TTT AAG AGC AGA GC																																									
			JIRPCHf	ACT GCA ACA GGT GAG GCC A	297	60	10																																																																															
			JIRPCHr5	AGG CCT GGT GAT GGT GGG AG				Neuroparsin	JINPf	CCA GGT TGG ACG CAG CAC CAC	334	60	10	JINpr	GTG CCA AGT GCT GCC ACT GGT AG	<i>N. viridula</i>	corpus cardiacum	Actin	Nvactin f	CTC CCC TCA ACC CTA AAG CCA AC	444	52	30	Nvactin r	GGA ACC TCT CGT TAC CGA TGG TG			Neuroparsin	NvNPf2	TCC CAG AGA TCG AGA TGT ACT CC	350	54	10	NvNpr4	TCC GGC AGA TTT AAG AGC AGA GC																																																			
Neuroparsin	JINPf	CCA GGT TGG ACG CAG CAC CAC	334	60	10																																																																																	
	JINpr	GTG CCA AGT GCT GCC ACT GGT AG				<i>N. viridula</i>	corpus cardiacum	Actin	Nvactin f	CTC CCC TCA ACC CTA AAG CCA AC	444	52	30	Nvactin r	GGA ACC TCT CGT TAC CGA TGG TG			Neuroparsin	NvNPf2	TCC CAG AGA TCG AGA TGT ACT CC	350	54	10	NvNpr4	TCC GGC AGA TTT AAG AGC AGA GC																																																													
<i>N. viridula</i>	corpus cardiacum	Actin	Nvactin f	CTC CCC TCA ACC CTA AAG CCA AC	444				52	30																																																																												
			Nvactin r	GGA ACC TCT CGT TAC CGA TGG TG				Neuroparsin			NvNPf2	TCC CAG AGA TCG AGA TGT ACT CC	350	54	10	NvNpr4	TCC GGC AGA TTT AAG AGC AGA GC																																																																					
		Neuroparsin	NvNPf2	TCC CAG AGA TCG AGA TGT ACT CC	350				54	10																																																																												
			NvNpr4	TCC GGC AGA TTT AAG AGC AGA GC																																																																																		

Legend:**Ta - Annealing temperature****Te - Elongation time**

2.6 Whole mount *in situ* hybridization for the detection of mRNA transcripts

2.6.1 *Nezara viridula* and *Daphnia pulex* tissues

Whole mount *in situ* hybridizations were performed by removing animal tissues by dissection. The following tissues were dissected: *N. viridula* male and female head, thorax and abdomen were open so that the tissue would be attached to half of the insect carapace and for *D. pulex* whole animals were used. These tissues were subsequently fixed in MEMFA described below.

2.6.2 MEMFA fixation

For whole mount *in situ* hybridisation, tissues were fixed in MEMFA. For *N. viridula*, half of the exoskeleton was removed so that the insects' organs would not fall out and become damaged when placed in a 1.5 ml Eppendorf tube. The parts that were collected were the head, the thorax and the abdomen; these parts were rinsed in dH₂O and place in MEMFA fixative (100 mM MOPS, 2 mM EGTA, 1 mM MgCl₂·7H₂O, 3.7% formaldehyde) for 2 h at room temperature. 1 volume of methanol was added and incubated at room temperature for 3 min. The methanol-MEMFA solution was discarded, and 80% methanol was added and left at room temperature for 3 min. The 80% methanol was discarded, and the sample was placed in 100% methanol and stored at -20 °C until further use.

2.6.3 Prehybridisation treatment

Post-fixation of the tissues was conducted as follows: (unless otherwise stated samples were incubated at room temperature) samples were washed in 1 ml phosphate buffered saline Tween 20 (PBT: 130 mM NaCl; 10 mM sodium phosphate, pH 7.2; 0.1% (v/v) Tween 20) 3x 5 min, post fixed in 1 ml PBT with 4% formaldehyde for 15 min and washed in 1 ml PBT for 5x 5 min. Samples were subsequently incubated with 1 ml proteinase K 10 µg/ml

(Roche) made in PBT for 25 min. The proteinase K digestion was stopped by adding 2 mg/ml glycine made in PBT and incubated for 2min. Samples were refixed in 1 ml PBT containing 4% formaldehyde for 20 min and thereafter washed in 1 ml PBT for 5x 3 min.

The pGEM®-T Easy vector system (Promega) was used as the vector for the DNA insert fragments because it contained the SP6 and T7 transcription initiation sites required to synthesise the probe. The inserts were ligated into the pGEM®-T Easy vector and grown in competent DH5α cells according to the manufacturer's protocol. The single stranded DIG labelled RNA probe was synthesised according to the DIG RNA labelling kit (SP6/T7) (Roche) instructions and stored at -20 °C until used.

2.6.4 Hybridisation procedure

Samples were washed in a 1 ml solution of a 1:1 (v/v) ratio hybridisation solution (Hybsol: 750 mM NaCl; 75 mM Na-citrate, pH 7.0; 50% (v/v) formamide; 0.1% (v/v) Tween 20; 50 µg/ml heparin; 50 µg/ml sonicated salmon sperm DNA) to PBT for 5 min. The samples were thereafter washed in 1 ml hybsol for 10 min. Samples were prehybridised in 0.5 ml hybsol for 30 min in a water bath set at 60 °C. After prehybridisation most of the hybsol liquid was removed, leaving behind just enough hybsol to cover 2 mm of the sample surface.

The RNA probe (called probe from henceforth) is synthesised from a plasmid carrying the gene of interest. Reverse transcription starts at either the T7 or SP6 transcription start sites which flank the gene of interest. For more detail about which regions for Dappu-RPCH and Dappu-RPCHR were used to synthesise probes refer to section 3.1.4 on page 60. Probes were prepared by adding 2 µl of probe solution synthesised previously to 5 µl of a solution containing 2 mg/ml sonicated salmon sperm DNA (SIGMA-ALDRICH®), which was denatured at 95 °C for 3 mins and rapidly cooled on ice thereafter. The probe solution was added directly to the samples and incubated at 60 °C overnight.

2.6.5 Washing and detection

Samples were removed from the water bath after overnight hybridisation and washed in 0.5 ml hybsol at hybridisation temperature for 2x 30 min. The samples were washed in a serial dilution of 1 ml hybsol: PBT (4:1; 3:2; 3:2 and 1:4) for 10 min thereafter the sample was washed in 1 ml PBT for 2x 10 min. The anti-DIG antibody conjugate was diluted 1:2000 in 0.5 ml PBT and incubated for 1 h. Samples were subsequently washed in 1 ml PBT for 3x 20 min and then washed in 1 ml staining buffer (0.1 M Tris-HCl, pH 9.5; 0.1 M NaCl) 3x 5 min. Samples were stained with 1ml staining buffer containing 20 µl NBT/BCIP stock solution (18.75 mg/ml NBT and 9.4 mg/ml BCIP) (Roche) and incubated between 30-60 min. At the appropriate time when the colour change is optimal, the staining reaction was stopped with 1 ml PBT.

University of Cape Town

3. Results

3.1 PCR-based cloning and sequence analysis of *Daphnia pulex* (German and South African ecotypes)

3.1.1 Actin

The whole genome of the cladoceran crustacean *D. pulex* was recently sequenced (2007) and was in the process of being annotated. A house keeping gene was required to check whether the quality of *D. pulex* mRNA for the German ecotype was good enough to be used. The Dappu-actin gene was selected, because actin is expressed in large amounts as well as ubiquitously within most organisms. The actin gene was identified by using the “Basic Local Alignment Search Tool” (BLAST) in the *D. pulex* genome database (wfleabase.org). The specific programme used to identify the actin gene was the nucleotide BLAST (nBLAST) programme search, where the “Expressed Sequence Tag” (EST) was selected in the “feature type” within the nBLAST program at wfleabase.org. Primers (Dappu actf and Dappu actr) were designed to amplify a 450 bp partial actin sequence with accession WFes0173874 in wfleabase.org. Using PCR the 450 bp DNA fragment was amplified and sequenced, this sequence was identified as actin when aligned to WFes0173874 and BLASTed at NCBI. Using the same primers the actin gene from the South African ecotype was PCR amplified and sequenced (Figure 12). The resulting sequence aligned to actin sequence of the German ecotype 100%, therefore, has the same partial actin mRNA sequence.

Dappu actf →

1 TGTGCGACGATGAAGTAGCAGC GTTGGTTGTTGACAACGGATCCGGTATGTGCAAGGCTG
1 C D D E V A A L V V D N G S G M C K A

61 GATTCCGCCGAGATGACGCCCTCGCGCCGTCTCCCATCCATTGTCGGCCGCCCGTT
20 G F A G D D A P R A V F P S I V G R P R

121 ACCAGGGCATCATGGTCGGTATGGGTCAGAAGGACTCGTACGTCGGTGACGAGGCTCAAT
40 Y Q G I M V G M G Q K D S Y V G D E A Q

181 CCAAACGTGGTATCTTGACCGTCAAGTACCCGATCGAGCACGGCATCGTCACCAACTGGG
60 S K R G I L T V K Y P I E H G I V T N W

241 ATGACATGGAAAAGATCTGGCACCACACCTTCTACAATGAGTTGCGTGTGCGCCCAGAGG
80 D D M E K I W H H T F Y N E L R V A P E

301 AACACCCCGTCCTCCTCACTGAAGCTCCCCTCAACCCCAAGGCTAACCGTGAGAAGATGA
100 E H P V L L T E A P L N P K A N R E K M

361 CCCAGATCATGTTTCGAGACTTCAACACACCCGCCATGTACGTTGCCATCCAAGCCGTGC
120 T Q I M F E T F N T P A M Y V A I Q A V

← Dappu actr

421 TCTCCCTGTACGCTTCCGGTCTACCACCGGTA
140 L S L Y A S G R T T G

Figure 12: Represents the highly conserved partial actin cDNA and deduced amino acid sequence of *D. pulex* (both the German and South African ecotype). Purple boxes indicate primer binding sites and arrows indicating their direction.

3.1.2 Verification of the red pigment-concentrating hormone

Using the nBLAST programme at wleabase.org the RPCH preprohormone was identified. Initially I searched for the Dappu-RPCH gene in the genomic nucleotide scaffolds and then later in the “Expressed Sequence Tag” (EST), which are selection criteria features within the nBLAST program at wleabase.org. When comparing the RPCH preprohormone nucleotide sequence obtained from the genome and that from the EST, there was a difference in size. The genomic derived sequence contains a start codon further upstream from the EST start codon (Figure 13) which is the only difference between the two sequences. Primers (Dpf and Dpr) were first designed, based on the genomic scaffolding nucleotide sequence, to amplify the whole sequence of the EST RPCH preprohormone. In PCR, a 425 bp DNA product was amplified (result not shown) from (whole animal) German *D. pulex* ecotype cDNA. PCR with the same primers and cDNA from additionally sourced *D. pulex* grown locally in Cape Town, South Africa, amplified a 452 bp DNA product (result not shown). The nucleotide and the deduced amino acid sequence from these amplicons are shown in Figure 14. Both sequences were reconfirmed by PCR-amplification with a high-fidelity *Taq* polymerase and sequencing. The RPCH prohormone derived from the German ecotype sequence contains 336 bp, 111 amino acids, while RPCH prohormone sequence derived from the South African ecotype contains 348 bp and 115 amino acids. Note that there are 2 codon deletions in the South African ecotype sequence and that it is extended by 6 amino acids before the stop codon in comparison with the German ecotype sequence (indicated red box in Figure 14). Later after my initial sequencing of the Dappu-RPCH sequences a paper by Dircksen et al. (2011) showed a Dappu-RPCH preprohormone sequence (annotated but unverified via PCR amplification and sequencing, attained by *in silico* search) that contained a start methionine even further upstream from both the previously published RPCH EST and the genomic scaffold sequence. Using 5' RACE PCR, I attempted to verify this start codon upstream of the putative start codon that we had annotated and amplified (Figure 14). The positive control produced amplicons using the same template cDNA and the primers Dpr and Dpf. All attempts to amplify the extra amino acids using 5' RACE PCR and reverse primers Dpr2, Dpr3 and Dpr4, were unsuccessful and no amplicons were obtained (results not shown).

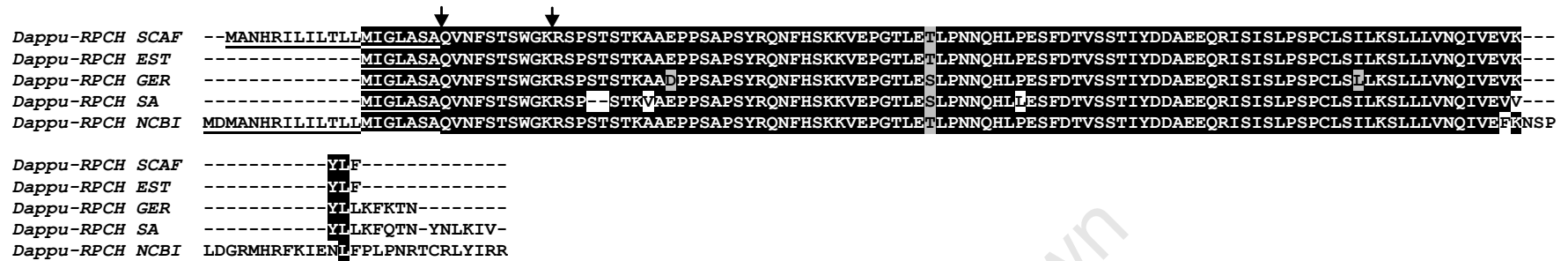


Figure 13: An alignment of the *D. pulex* preprohormone RPCH sequences showing the variation that exists between the genomic scaffolding sequence (SCAF), the expressed sequence tag (EST), the German ecotype (GER), the South African ecotype and the sequence that Dircksen et al. (2011) added to the NCBI database (NCBI). The black box represents amino acids that are 100% identical, the grey boxes represent amino acids that are similar in charge, cleavage sites are indicated by arrows (↓) and sequences that are underlined are predicted signal peptides. The alignment was performed in the bioinformatics tool BioEdit and gaps (-) have been introduced to optimise the alignment (Hall, 1999), “DAPPU RPCH SCAF” taken from Colbourne et al. (2011) and “DAPPU RPCH NCBI” from Dircksen et al. (2011).

A)

Dpf →
GCTATG

ATGATCGGCCTGGCTGGCGCTCAGGTGAATTTCTCGACCAGCTGGGGCAAGCGCTCACCT
M I G L A G A Q V N F S T S W G **K R** S P

TCGACTTCGACGAAAGCAGCGGATCCACCTTCAGCTCCATCTTACCGCCAAAATTTCCAT
 S T **S T** K **A** A **D** P P S A P S Y R Q N F H

TCGAAAAAGGTCGAGCCAGGGACGCTGGAATCGCTCCCCAATAATCAGCATCTTCCCGAA
 S K K V E P G T L E S L P N N Q H L **E** E

TCGTTTCGACACCGTTTCTTCGACGATTTACGACGACGCCGAAGAGCAACGAATCAGCATA
 S F D T V S S T I Y D D A E E Q R I S I

TCTCTACCAAGTCCTTGCCATCCCTTCTGAAATCTCTTGTGGTCAATCAAATAGTC
 S L P S P C L S **L** L K S L L L V N Q I V

GAGTAAAATATCTTTTAAAATTTAAAACAAAT**TAA**
 E V **K** Y L L K F **K** T N +

AATCTCAAATAGTTAAAAATTCACCGCTAGATGGCAGGATGCATAGGTTAAAAATAGAA
 ← Dpr

AATCTTTTCCCGTTGCCAATAGGACATGTAGGCTTTACATCCGTA

B)

Dpf →
GCTATG

ATGATCGGCCTGGCTAGCGCTCAGGTGAATTTCTCGACCAGCTGGGGCAAGCGCTCACCT
M I G L A S A Q V N F S T S W G **K R** S P

TCGACGAAAGTAGCGGAACCACCTTCAGCTCCATCTTACCGCCAAAATTTCCACTCGAAA
 S T K **V** A **E** P P S A P S Y R Q N F H S K

AAGGTCGAGCCAGGGACGCTGGAATCGCTCCCCAATAATCAGCATCTTCTCGAATCGTTC
 K V E P G T L E S L P N N Q H L **L** E S F

GACACCGTTTCTTCGACGATTTACGACGACGCCGAAGAGCAACGAATCAGCATACTCTA
 D T V S S T I Y D D A E E Q R I S I S L

CCAAGTCCTTGCCATCAATTCTGAAATCTCTTGTGGTCAATCAAATAGTTGAGGTA
 P S P C L S **I** L K S L L L V N Q I V E V

GTATATCTTTTAAAATTTCAAACAAATTACAATCTCAAATAGTCT**TAA**
V Y L L K F **Q** T N **Y N L K I V** +

AAATTCACCGCCAGATGGCAGGATGCATAGGTTAAAAATAGAAAATCTTTCCCAT
 ← Dpr

GCCAATAGGACATGTAGGCTTTACATCCGTA

Figure 14: The RPCH preprohormone cDNA and deduced amino acid sequence of *D. pulex* from A) German and B) South African ecotypes. Underlined sequences represent the 5' and 3' untranslated region (UTR), the green box represents the signal peptide region, the grey box the mature RPCH peptide sequence, the blue box the di-basic cleavage sites, the amino acid sequence in italics the "RPCH precursor-related peptide" (RPRP), the purple box the regions where the forward and reverse primer binds, the black boxes are amino acids that are different between the two ecotypes, the red box a substitution/deletion between the two ecotypes and the letters in bold represent the start and stop codons.

3.1.3 Verification of the red pigment-concentrating hormone receptor

The German ecotype *D. pulex* RPCHR was identified in the waterflea genomic database (wfleabase.org) using an *in silico* search. Primers were designed to amplify the RPCHR gene using 5' and 3' RACE PCR methods (for primer information, refer to Table 2, page 37). 3' RACE PCR yielded a DNA fragment of ~1408bp and 5' RACE a DNA fragment of ~872 bp. After sequencing both 5' RACE and 3' RACE fragments shown in Figure 15, which shows the overlapping amplicons. When attempting to amplify the entire Dappu-RPCHR, no amplicons could be amplified. Therefore, each 3' and 5' RACE fragment was verified in a nested PCR reaction with a high-fidelity *Taq* polymerase and sequencing. The consensus sequence was verified as the Dappu-RPCHR sequence when nBLAST tool was used at <http://blast.ncbi.nlm.nih.gov/Blast.cgi>, which gave a hit to Dappu-RPCHR (Accession: EU503126.1). These methods led to the verification of the *D. pulex* RPCHR sequence. The *D. pulex* RPCHRcDNA sequence contains a 1350 bp ORF which encodes a protein of 451 amino acids, has an isoelectric point of 8.75 and a molecular weight of 50.52 kg/mol (Figure 16, page 58). Dappu-RPCHR contains seven putative transmembrane (TM) regions, which was predicted by TMpred (Stoffel & Hofmann, 1993) (the web based tool can be located at ExpASy bioinformatics portal <http://expasy.org/tools/>). The web tool TMpred predicts that the N-terminus is located outside of the cell and the C-terminus is located inside. The TM regions vary in length, whereby the initial TM regions at the N-terminus are larger and become smaller at the C-terminus i.e.: first – 26, second – 26, third – 22, fourth – 21, fifth – 20, sixth – 19 and seventh – 20.

PCR amplification of the entire Dappu-RPCHR sequence was also attempted in the South African ecotype. The full Dappu-RPCHR sequence could also not be amplified directly. Instead, a portion of Dappu-RPCHR was amplified using the primers DpRPCHRf3 and DpRPCHRr3, produced a 773 bp amplicon (Figure 17A, page 59) which was sent for sequencing (Figure 17B, page 59). This partial RPCHR sequence would later be used to produce a RNA probe for *in situ* hybridisation experiments. The two sequences were aligned to view any differences between the two ecotype RPCHRs (Figure 18, page 59). Only two differences were observed, changes from Asp³⁴ to Gly³⁴ and Tyr¹²⁷ to Ala¹²⁷.

```

          10      20      30      40      50      60      70      80      90     100
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DpRPCHR seq#1 3'R AGAGTGTGAGGCTTATACCGTTCCGCCCGCAGCCGAACGACCGAGGCGCAGCGAGTCAGGGAGGCGAGGAAGCGAAGAGCGCCCAATACGCAACCGCCT
DpRPCHR seq#2 3'R -----
DpRPCHR seq#2 3'R -----
DpRPCHR seq#1f 5'R -----
DpRPCHR seq#1r 5'R -----
DpRPCHR seq#2f 5'R -----
DpRPCHR seq#2r 5'R -----
DpRPCHR scaffold -----
DpRPCHR Acc: EU503126.1 -----

          110     120     130     140     150     160     170     180     190     200
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DpRPCHR seq#1 3'R CTCCCCGCCGCGTGCCGATTCATTAATGCAGCTGCACGACAGGTTCCCGACTGAAAGCGGCAGTGAGCGCAACGCAATATGTGAGTAGCTCACTCATAG
DpRPCHR seq#2 3'R -----
DpRPCHR seq#2 3'R -----
DpRPCHR seq#1f 5'R -----
DpRPCHR seq#1r 5'R -----
DpRPCHR seq#2f 5'R -----
DpRPCHR seq#2r 5'R -----
DpRPCHR scaffold -----
DpRPCHR Acc: EU503126.1 -----

          210     220     230     240     250     260     270     280     290     300
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DpRPCHR seq#1 3'R GCACCCAGGCTTTACACTTTATGCTCCGGCTCGTATGTGTGTAATGTGAGCGGATACATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCT
DpRPCHR seq#2 3'R -----
DpRPCHR seq#2 3'R -----
DpRPCHR seq#1f 5'R -----
DpRPCHR seq#1r 5'R -----
DpRPCHR seq#2f 5'R -----
DpRPCHR seq#2r 5'R -----
DpRPCHR scaffold -----
DpRPCHR Acc: EU503126.1 -----

          310     320     330     340     350     360     370     380     390     400
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DpRPCHR seq#1 3'R ATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTGGGAGCTCTCCCATATGGTCGACTGCAGGCGGCCGGAATTCAGTAGTGAT-TC
DpRPCHR seq#2 3'R -----
DpRPCHR seq#2 3'R -----GGCTACCGGTTGCCACTGCTGACGGAAAGCCACGGTAGTGAAGATGCTCTGTTCGCTGCC
DpRPCHR seq#1f 5'R -----
DpRPCHR seq#1r 5'R -----
DpRPCHR seq#2f 5'R -----
DpRPCHR seq#2r 5'R -----
DpRPCHR scaffold -----
DpRPCHR Acc: EU503126.1 -----

```

Figure 15: An alignment of the different amplicons in the elucidation of the Dappu-RPCHR gene. Abbreviations: seq#1 – sequence number 1, “f” or “r” – 5'-3' sequenced amplicon and 3'-5' sequenced amplicons respectfully, 3'R/5'R – sequence elucidated in 3' RACE and 5' RACE respectfully and EU503126.1 is the DpRPCHR sequence annotated in NCBI. Sequences that are in black boxes are identical and arrows point to codons with attached label. This figure continues on 54 and ends on page 57.

```

      410      420      430      440      450      460      470      480      490      500
DpRPCHR seq#1 3'R CTCCAGTCTCAGTCAGCATTCGAACGCTCCACTCTCGTCATTCAGCATTTAGCACAACTCTTTTGATTCCCAAGTGAAAACCGTGCGACAGCTGCACCATTT
DpRPCHR seq#2 3'R CTCCAGTCTCAGTCAGCATTCGAACGCTCCACTCTCGTCGTTTCAGCATTTAGCACAACTCTTTTGATTCCCAAGTGAAAACCGTGCGACAGCTGCACCATTT
DpRPCHR seq#2 3'R -----
DpRPCHR seq#1f 5'R -----
DpRPCHR seq#1r 5'R -----
DpRPCHR seq#2f 5'R -----
DpRPCHR seq#2r 5'R -----
DpRPCHR scaffold -----
DpRPCHR Acc: EU503126.1 -----

```

```

      510      520      530      540      550      560      570      580      590      600
DpRPCHR seq#1 3'R CCCGTCCATCCAGTGGAAATTTGACGATTTCCAACGCCTGCGCTTTTCAGAAGCCGTTACTCCGTTTCAGCGAAATGTGTTCCAACGACAGCAGCAGCCTCA
DpRPCHR seq#2 3'R CCCGTCCATCCAGTGGAAATTTGACGATTTCCAACGCCTGCGCTTTTCAGAAGCCGTTACTCCGTTTCAGCGAAATGTGTTCCAACGACAGCAGCAGCCTCA
DpRPCHR seq#2 3'R -----
DpRPCHR seq#1f 5'R -----ACGATTTCCAACGCCTGCGCTTTTCAGAAGCCGTTACTCCGTTTCAGCGAAATGTGTTCCAACGACAGCAGCAGCCTCA
DpRPCHR seq#1r 5'R -----
DpRPCHR seq#2f 5'R -----ACGATTTCCAACGCCTGCGCTTTTCAGAAGCCGTTACTCCGTTTCAGCGAAATGTGTTCCAACGACAGCAGCAGCCTCA
DpRPCHR seq#2r 5'R -----
DpRPCHR scaffold --ATGTCGACAGTTCAGTTTGACGATTTCCAACGCCTGCGCTTTTCAGAAGCCGTTACTCCGTTTCAGCGAAATGTGTTCCAACGACAGCAGCAGCCTCA
DpRPCHR Acc: EU503126.1 -----

```

```

      ↑ ↑ Start codon (false)      ↑ ↑ Start codon (true)
      610      620      630      640      650      660      670      680      690      700

```

```

DpRPCHR seq#1 3'R GCGGCAATATGTGCATGATGACTGAACGGGATGCGTCGTTTAGCGGTGAGAGCACCGTCAACCCAGAGTTTGACAGTGGCAGTAGTATTAGTGGCGGAAG
DpRPCHR seq#2 3'R GCGGCAATATGTGCATGATGACTGAACGGGATGCGTCGTTTAGCGGTGAGAGCACCGTCAACCCAGAGTTTGACAGTGGCAGTAGTATTAGTGGCGGAAG
DpRPCHR seq#2 3'R -----
DpRPCHR seq#1f 5'R GCGGCAATATGTGCATGATGACTGAACGGGATGCGTCGTTTAGCGGTGAGAGCACCGTCAACCCAGAGTTTGACAGTGGCAGTAGTATTAGTGGCGGAAG
DpRPCHR seq#1r 5'R -----
DpRPCHR seq#2f 5'R GCGGCAATATGTGCATGATGACTGAACGGGATGCGTCGTTTAGCGGTGAGAGCACCGTCAACCCAGAGTTTGACAGTGGCAGTAGTATTAGTGGCGGAAG
DpRPCHR seq#2r 5'R -----
DpRPCHR scaffold GCGGCAATATGTGATGATGACTGAACGGGATGCGTCGTTTAGCGGTGAGAGCACCGTCAACCCAGTGTGTTGACAGTGGCAGTAGTATTAGTGGCGGAG
DpRPCHR Acc: EU503126.1 -----

```

```

      710      720      730      740      750      760      770      780      790      800
DpRPCHR seq#1 3'R CAGCAGCA-----CCACTGCGGTGCGACTTATCCATGTTGCCGATCGACATGACTTTCAACGACGGCCATATTGTGTCCATCGC-CACCTACAGTGTCTTT
DpRPCHR seq#2 3'R CAGCAGCA-----CCACTGCGGTGCGACTTATCCATGTTGCCGATCGACATGACTTTCAACGACGGCCATATTGTGTCCATCGC-CACCTACAGTGTCTTT
DpRPCHR seq#2 3'R -----
DpRPCHR seq#1f 5'R CAGCAGCA-----CCACTGCGGTGCGACTTATCCATGTTGCCGATCGACATGACTTTCAACGACGGCCATATTGTGTCCATCGC-CACCTACAGTGTCTTT
DpRPCHR seq#1r 5'R -----
DpRPCHR seq#2f 5'R CAGCAGCA-----CCACTGCGGTGCGACTTATCCATGTTGCCGATCGACATGACTTTCAACGACGGCCATATTGTGTCCATCGC-CACCTACAGTGTCTTT
DpRPCHR seq#2r 5'R -----
DpRPCHR scaffold CAGCAGCAGCACTACTACTACGGTGCAGTATCCATGTTGCCGATCGACATGACTTTCAACGACGGCCATATTGTGTCCATCGC-CACCTACAGTGTCTTT
DpRPCHR Acc: EU503126.1 -----

```

```

810      820      830      840      850      860      870      880      890      900
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DpRPCHR seq#1 3'R GCTTATCATCTCAG--TCTGCGGCAACATCACCGTCCT--AGTAAATCT--GATCAAAGGCGACACATCAGCAACCCGC--GTGTCAATATTATG--C
DpRPCHR seq#2 3'R GCTTATCATCTCAG--TCTGCGGCAACATCACCGTCCT--AGTAAATCT--GATCAAAGGCGACACATCAGCAACCCGC--GTGTCAATATTATG--C
DpRPCHR seq#2 3'R GCTTATCATCTCAG--TCTGCGGCAACATCACCGTCCT--AGTAAATCT--GATCAAAGGCGACACATCAGCAACCCGC--GTGTCAATATTATG--C
DpRPCHR seq#1f 5'R GCTTATCATCTCAG--TCTGCGGCAACATCACCGTCCT--AGTAAATCT--GATCAAAGGCGACACATCAGCAACCCGC--GTGTCAATATTATG--C
DpRPCHR seq#1r 5'R -----GTCGC--CGGGCAA--AACATCACCGTCCCTAAGTAAATCTGTAATCCAAAGTCGAC--CATCAGCA--CCCGCCCGGTGTCAATATTATG--C
DpRPCHR seq#2f 5'R GCTTATCATCTCAG--TCTGCGGCAACATCACCGTCCT--AGTAAATCT--GATCAAAGGCGACACATCAGCAACCCGC--GTGTCAATATTATG--C
DpRPCHR seq#2r 5'R ACTATACACTCCAGTTGGAGGCAACATCACCGTCTTT--AGTAAATCT--GATCAAAGGCAACCATTA--CAGACAACCC--CGGGATCAA--ATATAGC
DpRPCHR scaffold GCTTATCATCTCAG--TCTGCGGCAACATCACCGTCCT--AGTAAATCT--GATCAAAGGCGACACATCAGCAACCCGC--GTGTCAATATTATG--C
DpRPCHR Acc: EU503126.1 GCTTATCATCTCAG--TCTGCGGCAACATCACCGTCCT--AGTAAATCT--GATCAAAGGCGACACATCAGCAACCCGC--GTGTCAATATTATG--C

```

```

910      920      930      940      950      960      970      980      990      1000
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DpRPCHR seq#1 3'R TGACTCATCTAGCCA--TCGCCGACCT--GCTGGTTACGTTGCTATTGATGCCGATAGAGATTGGATGGGCGCTACGGT--CCAGTGGAGAGCAGGTGAT
DpRPCHR seq#2 3'R TGACTCATCTAGCCA--TCGCCGACCT--GCTGGTTACGTTGCTATTGATGCCGATAGAGATTGGATGGGCGCTACGGC--CCAGTGGAGAGCAGGTGAT
DpRPCHR seq#2 3'R TGACTCATCTAGCCA--TCGCCGACCT--GCTGGTTACGTTGCTATTGATGCCGATAGAGATTGGATGGGCGCTACGGT--CCAGTGGAGAGCAGGTGAT
DpRPCHR seq#1f 5'R TGACTCATCTAGCCA--TCGCCGACCT--GCTGGTTACGTTGCTATTGATGCCGATAGAGATTGGATGGGCGCTACGGT--CCAGTGGAGAGCAGGTGAT
DpRPCHR seq#1r 5'R TGACTCATCTAGCCA--TCGC--GA--ACT--GCTG--TACGT--GCTAT--GATGCCGATAGAGAT--GATGG--CCGCTACGGT--CCAGTG--AGAGCAGGTGAT
DpRPCHR seq#2f 5'R TGACTCATCTAGCCA--TCGCCGACCT--GCTGGTTACGTTGCTATTGATGCCGATAGAGATTGGATGGGCGCTACGGT--CCAGTGGAGAGCAGGTGAT
DpRPCHR seq#2r 5'R TGACTCATCTAGCCA--ATTCGCCGACCTTGCTGGTTACGTT--GCTAT--GATGC--GATAGAGAT--GATGGGCCACTACCGTCCAGTGGAGAGCAG--TGAT
DpRPCHR scaffold TGACTCATCTAGCCA--TCGCCGACCT--GCTGGTTACGTTGCTATTGATGCCAATCGAGATTGGATGGGCGCTACGGT--CCAGTGGAGAGCAGGTGAT
DpRPCHR Acc: EU503126.1 TGACTCATCTAGCCA--TCGCCGACCT--GCTGGTTACGTTGCTATTGATGCCAATCGAGATTGGATGGGCGCTACGGT--CCAGTGGAGAGCAGGTGAT

```

```

1010     1020     1030     1040     1050     1060     1070     1080     1090     1100
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DpRPCHR seq#1 3'R TTTTCTGCCGGATTCTGGCCTTTTTCC--GGACGTTCCGGCTTGTTCCTCTCTTCATTTGTTTTGGTTTGCATCAGCATTGATAGATTCCGGAGCTATTTTG
DpRPCHR seq#2 3'R TTTTCTGCCGGATTCTGGCCTTTTTCC--GGACGTTCCGGCTTGTTCCTCTCTTCATTTGTTTTGGTTTGCATCAGCATTGATAGATTCCGGAGCTATTTTG
DpRPCHR seq#2 3'R TTTTCTGCCGGATTCTGGCCTTTTTCC--GGACGTTCCGGCTTGTTCCTCTCTTCATTTGTTTTGGTTTGCATCAGCATTGATAGATTCCGGAGCTATTTTG
DpRPCHR seq#1f 5'R TTTTCTGCCGGATTCTGGCCTTTTTCC--GGACGTTCCGGCTTGTTCCTCTCTTCATTTGTTTTGGTTTGCATCAGCATTGATAGATTCCGGAGCTATTTTG
DpRPCHR seq#1r 5'R TTTTCTGCCGGATTCTGGCCTTTTTCC--GGACGTTCCGGCTTGTTCCTCTCTTCATTTGTTTTGGTTTGCATCAGCATTGATAGATTCCGGAGCTATTTTG
DpRPCHR seq#2f 5'R TTTTCTGCCGGATTCTGGCCTTTTTCC--GGACGTTCCGGCTTGTTCCTCTCTTCATTTGTTTTGGTTTGCATCAGCATTGATAGATTCCGGAGCTATTTTG
DpRPCHR seq#2r 5'R TTTTCTGCCGGATTCTGGCCTTTTTCC--GGACGTTCCGGCTTGTTCCTCTCTTCATTTGTTTTGGTTTGCATCAGCATTGATAGATTCCGGAGCTATTTTG
DpRPCHR scaffold TTTTCTGCCGAATCTGGCCTTTTTCC--GGACGTTCCGGCTTGTTCCTCTCTTCATTTGTTTTGGTTTGCATCAGCATTGATAGATTCCGGAGCTATTTTG
DpRPCHR Acc: EU503126.1 TTTTCTGCCGAATCTGGCCTTTTTCC--GGACGTTCCGGCTTGTTCCTCTCTTCATTTGTTTTGGTTTGCATCAGCATTGATAGATTCCGGAGCTATTTTG

```

```

1110     1120     1130     1140     1150     1160     1170     1180     1190     1200
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DpRPCHR seq#1 3'R CAACCCATGAAA--TTAGATTATTGGAACGTCGTGGGCGTTTCATGTTAGCCATTGCTTGGACTTGTTCCTGCTCCTC-----
DpRPCHR seq#2 3'R CAACCCATGAAA--TTAGATTATTGGAACGTCGTGGGCGTTTCATGTTAGCCATTGCTTGGGCTTGTTCCTGCTCCTC-----
DpRPCHR seq#2 3'R CAACCCATGAAA--TTAGATTATTGGAACGTCGTGGGCGTTTCATGTTAGCCATTGCTTGGGCTTGTTCCTGCTCCTC-----
DpRPCHR seq#1f 5'R CAACCCATGAAA--TTAGATTATTGGAACGTCGTGGGCGTTTCATGTTAGCCATTGCTTGGGCTTGTTCCTGCTCCTCAGGTTTTTCGTG
DpRPCHR seq#1r 5'R CAACCCATGAAA--TTAGATTATTGGAACGTCGTGGGCGTTTCATGTTAGCCATTGCTTGGGCTTGTTCCTGCTCCTCAGGTTTTTCGTG
DpRPCHR seq#2f 5'R CAACCCATGAAA--TTAGATTATTGGAACGTCGTGGGCGTTTCATGTTAGCCATTGCTTGGGCTTGTTCCTGCTCCTCAGGTTTTTCGTG
DpRPCHR seq#2r 5'R CAACCCATGAAA--TTAGATTATTGGAACGTCGTGGGCGTTTCATGTTAGCCATTGCTTGGGCTTGTTCCTGCTCCTCAGGTTTTTCGTG
DpRPCHR scaffold CAACCCATGAAA--TTAGATTATTGGAACGTCGTGGGCGTTTCATGTTAGCCATTGCTTGGGCTTGTTCCTGCTCCTCAGGTTTTTCGTG
DpRPCHR Acc: EU503126.1 CAACCCATGAAA--TTAGATTATTGGAACGTCGTGGGCGTTTCATGTTAGCCATTGCTTGGGCTTGTTCCTGCTCCTCAGGTTTTTCGTG

```

1210 1220 1230 1240 1250 1260 1270 1280 1290 1300
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

DpRPCHR seq#1 3'R
DpRPCHR seq#2 3'R
DpRPCHR seq#2 3'R
DpRPCHR seq#1f 5'R
DpRPCHR seq#1r 5'R
DpRPCHR seq#2f 5'R
DpRPCHR seq#2r 5'R
DpRPCHR scaffold
DpRPCHR Acc: EU503126.1
TTCCACGTCAAAGCTCATCCGGAATACCCCTTGGTATGAGCAGTGGCTCACTTTTGATTTCGTTTCCGACCAAGGCTCACGAGATTTCTTACGCTGCCTTC
TTCCACGTCAAAGCTCATCCGGAATACCCCTTGGTATGAGCAGTGGCTCACTTTTGATTTCGTTTCCGACCAAGGCTCACGAGATTTCTTACGCTGCCTTC
TTCCACGTCAAAGCTCATCCGGAATACCCCTTGGTATGAGCAGTGGCTCACTTTTGATTTCGTTTCCGACCAAGGCTCACGAGATTTCTTACGCTGCCTTC
TTCCACGTCAAAGCTCATCCGGAATACCCCTTGGTATGAGCAGTGGCTCACTTTTGATTTCGTTTCCGACCAAGGCTCACGAGATTTCTTACGCTGCCTTC
TTCCACGTCAAAGCTCATCCGGAATACCCCTTGGTATGAGCAGTGGCTCACTTTTGATTTCGTTTCCGACCAAGGCTCACGAGATTTCTTACGCTGCCTTC

1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

DpRPCHR seq#1 3'R
DpRPCHR seq#2 3'R
DpRPCHR seq#2 3'R
DpRPCHR seq#1f 5'R
DpRPCHR seq#1r 5'R
DpRPCHR seq#2f 5'R
DpRPCHR seq#2r 5'R
DpRPCHR scaffold
DpRPCHR Acc: EU503126.1
GGCATGATGATGATGATGATGTTCTTCCACTCGCCGTTTTCGTCCTTCACTTATAGTTTCGATTCTTTGCGAGATCAGCAGAGCAGCAAGGAAG---TCGGAC
GGCATGATGATGATGATGATGTTCTTCCACTCGCCGTTTTCGTCCTTCACTTATAGTTTCGATTCTTTGCGAGATCAGCAGGAGGAGCAAGGAAG---TCGGAC
GGCATGATGATGATGATGATGTTCTTCCACTCGCCGTTTTCGTCCTTCACTTATAGTTTCGATTCTTTGCGAGATCAGCAGAGCAGCAAGGAAG---TCGGAC
GGCATGATGATGATGATGATGTTCTTCCACTCGCCGTTTTCGTCCTTCACTTATAGTTTCGATTCTTTGCGAGATCAGCAGGAGGAGCAAGGAAG---TCGGAC
GGCATGATGATGATGATGATGTTCTTCCACTCGCCGTTTTCGTCCTTCACTTATAGTTTCGATTCTTTGCGAGATCAGCAGGAGGAGCAAGGAAG---TCGGAC
GGCATGATGATGATGATGATGTTCTTCCACTCGCCGTTTTCGTCCTTCACTTATAGTTTCGATTCTTTGCGAGATCAGCAGGAGGAGCAAGGAAG---TCGGAC

1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

DpRPCHR seq#1 3'R
DpRPCHR seq#2 3'R
DpRPCHR seq#2 3'R
DpRPCHR seq#1f 5'R
DpRPCHR seq#1r 5'R
DpRPCHR seq#2f 5'R
DpRPCHR seq#2r 5'R
DpRPCHR scaffold
DpRPCHR Acc: EU503126.1
AAGAGGAGGGTATCCGAAGGG-----
AAGAGGAGGGTATCCGAAGGGTACTGTCGGTACGCTTGGACGAGCAGGTATAAAAACCGTCAAATGACTCTAGTGATCATTTCCGTTTTTCATTTTCTG
AAGAGGAGGGTATCCGAAGGGTACTGTCGGTACGCTTGGACGAGCAGGTATAAAAACCGTCAAATGACTCTAGTGATCATTTCCGTTTTTCATTTTCTG
AAGAGGAGGGTATCCGAAGGGTACTGTCGGTACGCTTGGACGAGCAGGTATAAAAACCGTCAAATGACTCTAGTGATCATTTCCGTTTTTCATTTTCTG
---AGGAGGGTATCCGAAGGGTACTGTCGGTACGCTTGGACGAGCAGGTATAAAAACCGTCAAATGACTCTAGTGATCATTTCCGTTTTTCATTTTCTG
AAGAGGAGGGTATCCGAAGGGTACTGTCGGTACGCTTGGACGAGCAGGTATAAAAACCGTCAAATGACTCTAGTGATCATTTCCGTTTTTCATTTTCTG

1510 1520 1530 1540 1550 1560 1570 1580 1590 1600
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

DpRPCHR seq#1 3'R
DpRPCHR seq#2 3'R
DpRPCHR seq#2 3'R
DpRPCHR seq#1f 5'R
DpRPCHR seq#1r 5'R
DpRPCHR seq#2f 5'R
DpRPCHR seq#2r 5'R
DpRPCHR scaffold
DpRPCHR Acc: EU503126.1
CTGGACCCGTACAACATCATGAGCATCTGGTTCTGGTGTGACAGAGATTTCGGCGCTTCAAGTTGACCAACGAATCCAAAAGGCCTTTTTCTTTTTCG
CTGGACCCGTACAACATCATGAGCATCTGGTTCTGGTGTGACAGAGATTTCGGCGCTTCAAGTTGACCAACGAATCCAAAAGGCCTTTTTCTTTTTCG
CTGGACCCGTACAACATCATGAGCATCTGGTTCTGGTGTGACAGAGATTTCGGCGCTTCAAGTTGACCAACGAATCCAAAAGGCCTTTTTCTTTTTCG
CTGGACCCGTACAACATCATGAGCATCTGGTTCTGGTGTGACAGAGATTTCGGCGCTTCAAGTTGACCAACGAATCCAAAAGGCCTTTTTCTTTTTCG
CTGGACCCGTACAACATCATGAGCATCTGGTTCTGGTGTGACAGAGATTTCGGCGCTTCAAGTTGACCAACGAATCCAAAAGGCCTTTTTCTTTTTCG

AGAGTGTGAGGCTTATACCGTTCCGCCCGCAGCCGAACGACCGAGGCGCAGCGAGTCAGGGAGGCGAGGAAGCGAAGAGCGCCCAATAC
 GCAACCGCCTCTCCCGCCCGCGTGGCGATTCATTAATGCAGCTGCACGACAGGTTCCCGACTGAAAGCGGCAGTGAGCGCAACGCAATA
 TGTGAGTAGCTCACTCATAGGCACCCAGGCTTTACACTTTATGCTCCGGCTCGTATGTGTGAATGTGAGCGGATACATTTACACAGG
 AAACAGCTATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCAT
 ATGGTCGACTGCAGGCGGCCGGAATTCAGTGTGATT **CCTCCAGTCTCAGTCAGCATTTCG** AACGTCCTCTCGTCATTC **AGCATTTCAG**
 → **CACAACCTCTT** TTGATTCCCAAGTGAACCCGTGCGACAGCTGCACCATTCCTCCATCCAGTGGAAATTTGACGATTTCCAACGCCTGGC
 CTTTCAGAAGCCGTTACTCCGTTTCAGCGAA
 ATGTTGTTCCAACGACAGCAGCAGCAGCCTCAGCGGCAATATGTGCATGATGACTGAACGGGA
 M C S N D S S S L S G N M C M M T E R D
 TGGCTCGTTTAGCGGTGAGAGCACCGTCAACCCAGAGTTTACAGTGGCAGTAGTATTAGTGGCGAAGCAGCAGCACCCTGCGGTGCA
 A S F S G E S T V N P E F D S G S S I S G G S S S T T A V D
 CTTATCCATGTTGCCGATCGACATGACTTTCAACGACGGCCATATGTGTCCATCGCCACCTACAGTGTCTTGCTTATCATCTCAGTCTG
 L S M L P I D M T F N D G H **I V S I A T Y S V L L I I S V C**
 CGGCAACATCACCGTCTAGTAAATCTGATCAAAGGCGACACATCAGCAA **CCCGCTGTCAATATATGCTGAC** TCATCTAGCCATCGC
G N I T V L V N L I K R R H I S N P R V N **I M L T H L A I A**
 CGACCTGCTGGTTACGTTGCTATTGATGCCGATAGAGATTGGATGGGCCGTACCGTCCAGTGGAGAGCAGGTGATTTTCCCTGCCGGAT
D L L V T L L L M P I E I G W A A T V Q W R A G D F S C R **L**
 TCTGGCCTTTTTCCGGACGTTCCGGCTTGTCTCTCTTCATTTGTTTTGGTTTGCATCAGCATTGATAGATTCCGAGCTATTTTGC AAC
L A F F R T F G L F L S S F V L V C I S I D R F G A I L Q P
 CATGAAATTAGATTATTGAAACGTCGTGGCGCTTTCATGTTAGCCATTGCTTG **GGCTTGTCCGTCATTGCTCTC** TGCCCTCAGGTTTT
 M K L D Y W K R R G R **F M L A I A W A C S V I C S L P Q V F**
 CGTGTCCACGTCAAAGCTCATCCGGAATACCCTTGGTATGAGCAGTGCCTCACTTTTGGTTTCCGACCAAGGCTCACGAGATTTT
V E H V K A H P E Y P W Y E Q C V T F D S F P T K A H E **I S**
 TTACGCTGCCTTCGGCATGATGATGATGATGATGTTCTTCCACTCGCGGTTTTCTGCTTCACTTATAGTTCGATTCTTTGCGAGATCAGCAG
Y A A F G M M M Y V L P L A V F V F T Y S S I L C E I S R
 AGGAGCAAAGGAAGTCGGACAAGAGGAGGTATCCGAAGGGTGAAGTGTCCGTCAGCTTGGACGAGCAGTATAAAAACCGTCAAAATGAC
 G A K E V G Q E E G I R V T V G T L G R A R I K T V K **M T**
 TCTAGTATCATTTCCGTTTTTCAATTTCTGCTGGAGCGGTACAACATCATGAGCATCTGGTTCTGGTGTGACAGAGATTCGCGCCTTCA
L V I I S V F I F C W T P Y N I M S I W F W C D R D S A L Q
 AGTTGACCAACGAATCCAAAAGGCTTTTTCTTTTTGCGTGTACCAATTCCTGCTTCAACCCATGGTGTACGGTTACTTTTCGCGACG
 V D Q R I Q K **G L F L F A C T N S C F N P M V Y G Y F** S R R
 AACGGTTCCGAGATCTCACCCAGAACTCCACAGAAAGGTGGTCTATCATCCAGTGAATGGTATCAAGGGCTTGGTCCCTCACTCAG
 T V R R S H E L H R K V V Y H P S R M V S R A L G P S L R
 AGTAGATTGCAATTCTAAATCATTGGAAGGAGCAGTAACCTGAACCGTGTGGCACAGCAGCAT **CTACCGGTTGCGACTGCCGACG** ATCC
 V D C N S K S L E G A V T E P C L P Q Q H L P V A T A D D P
 TTCATGTCAGTTGGGACAGCAGAACCATTTCACCGCCTCTCTATCGCCGTAGCTGTTAAAAGAAGTAATAGCTGGCTGCTGAATCA
 S C Q L G Q A E P I F T A S S I A V A V K R S N S W L L N H
 TCGATCGAAT **CCAACCTGTCGTCACACATATATTTAA**
 R S N P T V V T H I F +

Figure 16: Represents the consensus nucleotide sequence for the Dappu-RPCHR and its deduced peptide sequence from the *D. pulex* (German ecotype). Underlined sequences represent the 5' and 3' UTR, the black and grey boxes represent the transmembrane regions (TM I-VII). Primers used in 5'/ 3' RACE are in purple boxes with arrows indicating their direction and the letters in bold represent the start and stop codons. Transmembrane regions were predicted by the web prediction tool TMpred (Stoffel & Hofmann, 1993).

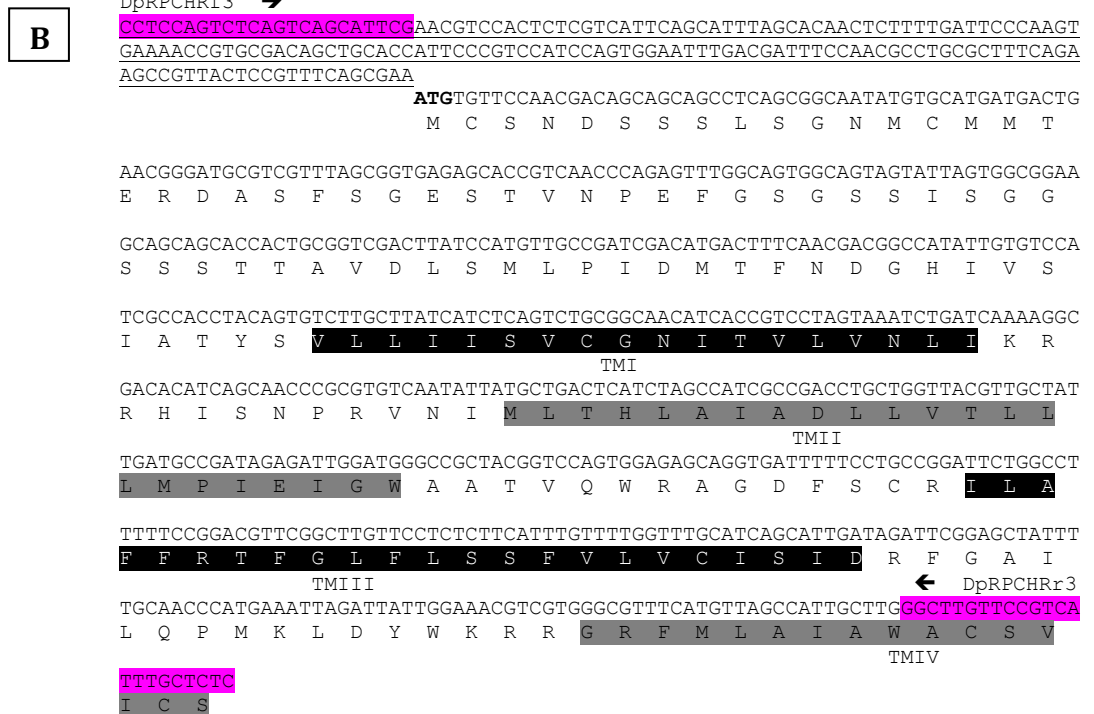
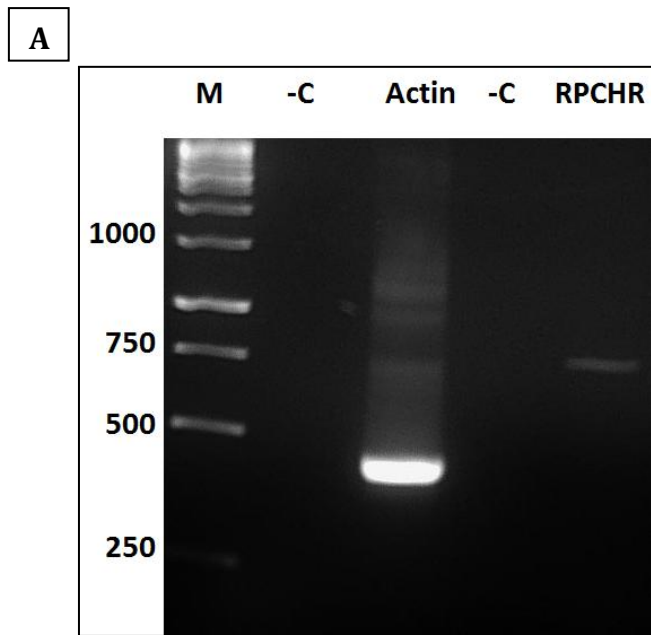


Figure 17: (A) Depicts an agarose electrophoresis gel showing the actin and the partial (~750 bp) amplification of the Dappu-RPCHR gene (RPCHR) and (B) represents the RPCHR cDNA and its deduced peptide sequence from *D. pulex* (South African ecotype). Legend: (A) -C: no template control, DpRPCHR: Dappu-RPCHR and M: marker with sizes indicated on the left side in base pairs and (B) the underlined sequences represent the 5' UTR; the black and grey boxes represent the transmembrane regions (TM I-IV), primer sequences are indicated in purple boxes with arrows showing their direction and the letters in bold represent the start codon.

Dappu RPCHR GER	MCSNDSSLSGNMCMTERDASFSGESTVNPEFD SGSSISGGSSSTTAVDLSMLPIDMTFNDGHIVSIATYSVLLIISVCGNITVLVNLIKRRHISNPRVNI MLTHLAIADLLVTL LLMP
Dappu RPCHR SA	MCSNDSSLSGNMCMTERDASFSGESTVNPEFD SGSSISGGSSSTTAVDLSMLPIDMTFNDGHIVSIATYSVLLIISVCGNITVLVNLIKRRHISNPRVNI MLTHLAIADLLVTL LLMP
Dappu RPCHR GER	IEIGWATVQWRAGDFSCRILAFFRTFGLFLSSEFVLVCISIDRFGAILQPMKLDYWKRRGRFMLAIAWACSVICS SLPQVVFVHVKAHPEYPWYEQCVTFDSFPTKAHEISYAAF GMMMMY
Dappu RPCHR SA	IEIGWATVQWRAGDFSCRILAFFRTFGLFLSSEFVLVCISIDRFGAILQPMKLDYWKRRGRFMLAIAWACSVICS -----
Dappu RPCHR GER	VLPLAVFVFTYSSILCEIRQRSKEVGQEEGIRRVTVGTLGRARIKTVKMTLVIISVFI FCWT PYNIMSIWFWCDRDSALQVDQRIQKGLFL FACTNSCFNPMVYGYFSRRTVRRSHHELH
Dappu RPCHR SA	-----
Dappu RPCHR GER	RKVYVYHPSRMVSRALGPSLRVDCNSKSLEGAVTEPCLPQOHLPVATADD PSCQLGQAEPI FTASSIAVAVKRRNGIKGLGSLTQ SRLQF
Dappu RPCHR SA	-----

Figure 18: An alignment of the complete preprohormone RPCHR peptide sequence amplified from the German ecotype (GER) *D. pulex* and partial sequence from the South African ecotype. The black box represents amino acids that are 100% identical. The alignment was performed in the bioinformatics tool BioEdit.

3.1.4 Localisation of RPCH and RPCHR in *Daphnia pulex* tissues

To analyse where the Dappu-RPCH and the Dappu-RPCHR were localised within the South African ecotype of *D. pulex*, preliminary *in situ* hybridisation experiments were conducted on whole *D. pulex* waterfleas. The probes that were used to locate the mRNA transcripts were the whole Dappu-RPCH (located between primers Dpf and Dpr see Figure 14, page 51) and a ~750 bp Dappu-RPCHR region between primers DpRPCHRf3 and DpRPCHRr3 (see Figure 18). The control probe used as a negative control was the Neo RNA from the DIG/RNA labelling kit. Figure 19 shows the localisation of the Dappu-RPCH and the Dappu-RPCHR probes. This experiment was performed and optimised in three attempts. Although still not fully optimised (due to time constraints), a few observations could be made based on the *in situ* hybridisation images. Note that the region in Figure 19B contains the NBT/BCIP purple colour, which encompasses the ovaries, gut and thoracopods thus suggesting that the RPCH is expressed here. In Figure 19C, the colour is part of something that appears to be part of the thoracopods more than the ovaries, this suggests that the receptor may be expressed here (but this needs to be verified). Also, note that NBT/BCIP accumulated around areas where the animal was damaged.

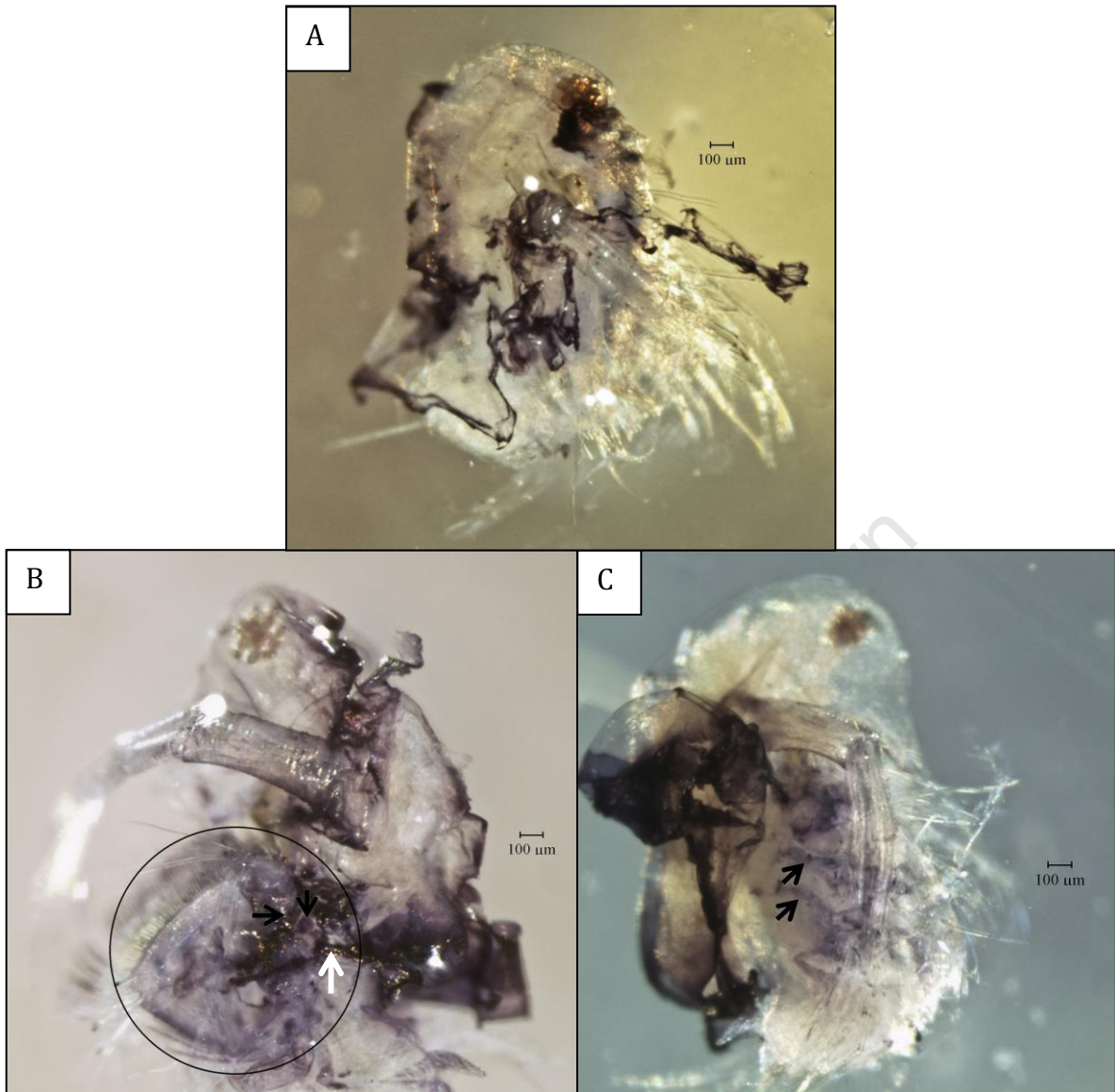


Figure 19: Images from the *in situ* hybridisation experiments of control (A), Dappu-RPCH (B) and Dappu-RPCHR (C). The black arrows indicate cells that contain the NBT/BCIP colour, the white arrows indicates a granular appearance associated with NBT/BCIP and the circle a region of interest.

3.2 PCR-based cloning and sequence analysis of *Nezara viridula*

3.2.1 Splinkerette PCR

The initial attempted amplification of the *N. viridula* AKH preprohormone sequence was performed by the splinkerette system; however, no specific PCR products could be obtained. To verify that the DNA from *N. viridula* tissue was of good enough quality the primers 28srDNAf and 28srDNAr were used to amplify a partial 28s rDNA gene (Kavar et al., 2007). The 28srDNA gene was amplified in PCR, which yielded a ~545 bp DNA fragment and verified as the 28srDNA gene via sequencing and a BLASTx search.

Degenerate primers were used to amplify the AKH preprohormone in a nested PCR. The primers used in the nested PCR were a combination of 2 degenerate primers (NvAKH, NvDeg1 or NvDeg2) used along with splinkerette primers Splk O then Splk I. Using these degenerate primers and the splinkerette primers, the splinkerette system generated amplicons where sizes were not predictable. The splinkerette PCRs generated either no products or smears when electrophoresed on agarose (result not shown). In this system the annealing temperature was changed in different PCR reactions from 60 °C to 50 °C, the MgCl₂ concentrations (1.5 – 3 mM) was also changed in the attempt to amplify the AKH preprohormone. The final attempt at trying to amplify the AKH preprohormone was to use a touch-down PCR method whereby the temperature was progressively decreased from 60 °C – 48 °C. The splinkerette PCR system in the end was not successful at specifically amplifying the AKH preprohormone, and thus was abandoned.

3.2.2 Amplification of NP in 3' RACE and 5' RACE PCR

The next step in attempting to amplify the AKH preprohormone was to use 3' RACE PCR. The APP oligo dT-adapter was used to make cDNA from *N. viridula* CC tissue. The same degenerate primers (Table 2, page 37) in the splinkerette PCRs were used and also in combination with the AUAP primer to perform a nested PCR reaction (the AUAP primer was used in both the first and second PCR reaction). Using the AUAP primer in both the first and second nested PCR reactions resulted in the enrichment of non-target amplicons. Non-target amplicons were defined as amplicons outside the range of ~200-

500 bp, which was based on the average size of other AKH prohormones and starting from the codons which code for the mature peptide where the primer binds and ending with the 3' UTR. One particular 3' RACE amplicon of 392 bp was sent for sequencing (Figure 20). Using BLASTx the 392 bp query had 61% homology to the NP peptide sequence of another hemipteran *Rhodnius prolixus*. In Figure 21 the AUAP primer is shown bound in the 5' to 3' direction to a region between the signal peptide and mature peptide. The partial NP sequence was verified with non-degenerate primers NvNpf and NvNpr1, as well as subsequent sequencing.

Further 5' and 3' RACE PCR experiments in *N. viridula* were conducted using the Roche 5'/3' RACE kit, due to the non-target amplification of the NP amplicon using the AUAP/APP RACE system. The 5' end of the NP sequences was amplified using 5' RACE PCR and primers NvNpr1 – 3; a ~377 bp product was amplified and sent for sequencing. The 3' and 5' RACE sequences were aligned and a consensus region generated. Verification of the NP prohormone region (nucleotide sequence encoding the signal peptide and mature peptide sequence) was achieved by using a high fidelity *Taq* polymerase and primers NvNpf2 and NvNpr4. The NP cDNA contains 321 bp ORF and encodes prohormone of 106 amino acid residues (Figure 21). The NP prohormone includes a signal peptide and a mature peptide sequence made up of 14 cysteine residues which are located at conserved positions homologous to other insect NPs. (the reader is referred to the multiple sequence alignment of NP in Figure 37 (page 89 in subchapter 4.4), where it is further analysed in the discussion).

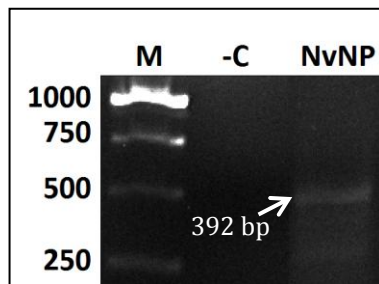


Figure 20: shows an agarose electrophoresis gel with first partial amplification of the NP gene in 3' RACE. Abbreviations: -C: no template control, NvNP: *N. viridula* NP and M: marker. Numbers on the left are the sizes of the marker in base pairs and the arrow points to partial NP band.

```

GACTGAAGACTATATCCAATCCGAAGATGCAAATGCTCCGATAATGAACTGTCAATCAGC
CCTGGAAGGGCCGATATAAGGGCCCTGCAGCCCTGAAGAAGCCATCAGTCTCCAACTCC
NvNpf2 → ****-
CAGAGATCGAGATGTACTCCCTGTTGCTCCTCCTCACAGCAGTGACCTCCGCCATGGCCT
M Y S L L L L L T A V T S A M A
AUAP →
---*-*****-
TCAGCTTCGACTACTACCTTGCAAACCTGTAGGGGAGAAGAGTGCTACGTAGAACCGG
F S F D Y Y P C K P C R G E E C Y V E P
AGGGCTGCAATTACGGCATCACCCGGGACCTTGCGGAAGGATGCAGTGATGGCCGGAC
E G C N Y G I T R D P C G R M Q C M A G
CTGGACAACGATGCGGAGGACGGGACAACCATCTTGGGAAGTGCGGCGACGGAATGACCT
P G Q R C G G R D N H L G K C G D G M T
GCAAGTGCGGCAAGTGCAGGGCTGCAGCAACGACCGCATCTCCAAGGCATCATCGAGT
C K C G K C R G C S N D R I L Q G I I E
← NvNpr4
GCGACCCCAACACCACGCTGCTCTGCTCTTAA
C D P N T T L L C S +
ATCTGCCGACCGCTAAAGTAGTAGCTTCACCTAGACTGTATTCTGGAATAAACTAATTT
TGATACATCTAAAAAAAAAAAAAAAAAAAA

```

Figure 21: Nucleotide sequence of the Hemipteran insect *N. viridula* NP prohormone cDNA and its deduced amino acid sequence. Underlined nucleotide sequences represent the 5' and 3' UTR, the green box represents the predicted signal peptide, the red box the conserved cysteine residues, the yellow box represents the putative polyadenylation signal, the purple boxes primers used to amplify NP and letters in bold represent the start and stop codons. The stars (*) show nucleotide bases that are identical to AUAP, while the hashes (-) show bases that are not identical. The signal peptide for the amino acid sequence was predicted by SIG-Pred (http://bmbpcu36.leeds.ac.uk/prot_analysis/Signal.html).

3.2.3 Actin

For verification of cDNA quality and for studies involving the expression of NP prohormone in different tissues, a house keeping gene was required. The actin gene was selected because it is expressed ubiquitously and at high levels. Degenerate primers designed by Chaty et al. (2004) were used to PCR amplify a ~900 bp actin amplicon from CC tissue. The actin amplicon was verified via sequencing and BLASTx which revealed 99% homology to many insect actin genes like the hemipteran, a pea aphid: *A. pisum*. New non-degenerate primers (Nvactin f and Nvactin r) were designed to amplify a smaller portion of the ~900 bp actin nucleotide sequence. PCR amplification with the new primers yielded a 444bp actin band, which was verified as actin when sequenced (Figure 22).

```

GGATGATATGGAGAAGATCTGGCATCACACCTTCTACAACGAGCTGCGAGTCGC
TCCAGAGGAACACCCCATCCTCCTCACCGAGG
Nvactin f →
CTCCCCTCAACCCTAAAGCCAACAGGGAGAAGATGACCCAGATCATGTTTGAGACCTTCA
P L N P K A N R E K M T Q I M F E T F

ACACCCAGCCATGTATGTCGCCATCCAGGCTGTACTCTCCCTCTATGCCTCCGGTCGTA
N T P A M Y V A I Q A V L S L Y A S G R

CCACCGGTATTGTA CTGACTCAGGAGATGGTGTATCCCACACCGTACCCATCTATGAAG
T T G I V L D S G D G V S H T V P I Y E

GTTATGCCCTTCCCCACGCCATCCTCCGTCTGGATCTTGCTGGACGTGACTTGACTGACT
G Y A L P H A I L R L D L A G R D L T D

ACCTTATGAAGATCCTCACCGAGCGTGGTTACAGTTTCACAACCACCGCTGAAAGGGAAA
Y L M K I L T E R G Y S F T T T A E R E

TCGTCAGGACATCAAGGAAAACTGTGCTATGTTGCCCTGGACTTTGAACAGGAAATGG
I V R D I K E K L C Y V A L D F E Q E M

CCACCGTGCTGCCTCCACCTCCCTCGAGAAGTCTTATGAACTTCCTGACGGTCAGGTCA
A T A A A S T S L E K S Y E L P D G Q V
← Nvactin r
TCACCATCGGTAACGAGAGGTTCC
I T I G N E R F

GTTGCCAGAGGCTCTCTCCAGCCTTCCTTCTTGGGTATGGAATCTTGGCGTA
TCCATGAGACTGTCTACAACCTCCATCATGAAGTGCATGTTGACATCAGGAAGG
ACTTGTACGCCAACACCGTCTCTCCGGAGGTACCACCATGCACCCAGGTATTG
CTGACAGGATGCAGAAGGAAATCACCGCCCTCGCTCCTTCAACCATCAAGATCA
AGATCATCGCTCCCCAGAAAGGAAGTACTCCGTATGGATCGGTGGTTCCATCT
TGGCTTCCCTGTCCACCTTCCAGCAGATGTGGATCAGCAAACAGG

```

Figure 22: Represents the ~900bp actin sequence with the primers Nvactin f and Nvactin r (purple box) indicating the shorter 444 bp actin region used in subsequent PCR experiments for the hemipteran *N. viridula*. The deduced amino acid sequence for the shorter 444 bp sequence is shown below the nucleotide sequence, purple boxes indicate primer binding sites and arrows indicate their direction

3.2.4 Localisation of NP in *Nezara viridula* tissues

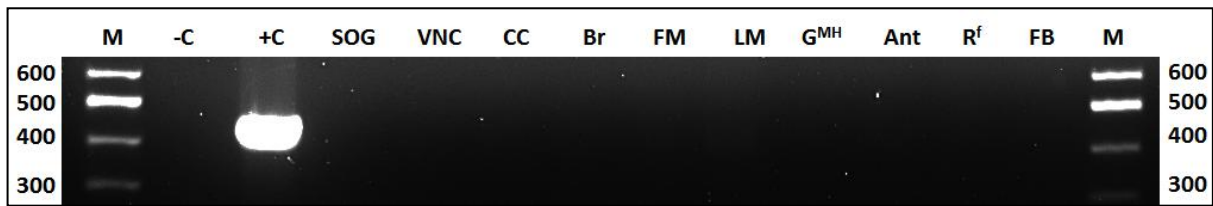
To ascertain whether there was a difference in temporal and spatial expression profile of NP with regards to sexual maturation in *N. viridula* two different time points were selected. The presence of *N. viridula* NP transcripts was analysed using reverse transcription PCR. Different tissues were selected from males and females at two time points i.e. 1-2 days into adulthood and > 3 weeks into adulthood. For the study mature males, mature females, immature males and immature females (n= ten individuals in each category) were dissected and the individual tissues collected and pooled for each group. The sample number n = ten was selected because the neural tissue of the insects, although possibly enriched with the transcript, are very small, and the concern was that the tissues might not produced enough RNA to synthesis cDNA especially for amplification and visualisation of the actin gene. The selected tissues were: the sub-oesophageal ganglion (SOG), ventral nerve cord (VNC), corpora cardiaca (CC), brain (Br), Leg muscle (LM), flight muscle (FM), Gut - mid and hind (G^{MH}), antennae (Ant), fatbody (FB), male reproductive organs (R^m) and female reproductive organs (R^f). To make sure that the amplification of NP was from cDNA only, all tissue total RNA was DNase treated before cDNA was synthesised. The total RNA was tested for amplification of the actin gene in a PCR reaction, the results of which were negative (Figure 23). This insures that the amplification from DNA sources was non-existent and that amplification was due to cDNA transcripts. The controls used in the PCR amplification were positive control (+C – a previous CC cDNA sample shown to amplify the target genes) to show that the PCR was successful, a no template control (-C) to show that there was no non-specific amplification due to contaminating DNA and an actin control to show that there were no problems associated with the synthesis of cDNA. The expression profile for the NP (350 bp) and actin control (444 bp) genes in *N. viridula*, male and female, mature and immature tissues are shown in Figure 24 to Figure 27.

Note that in the +C - NP lane of Figure 24 there was spill over from the +C – Act lane which can be seen in line with the actin amplicons at the 450 bp mark. In the mature female (Figure 24), the NP transcript was detected in all the tissues except G^{MH} where the actin gene was also not detected. In the mature males (Figure 25), the NP transcript was detected in all tissues except the VNC and the SOG. In the immature females (Figure

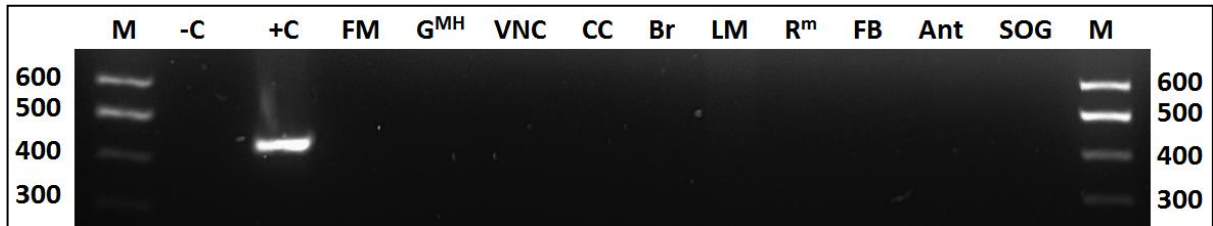
26), the NP transcript was detected in all tissue except the VNC. In the immature males (Figure 27), the NP transcript was detected in all tissues except the VNC, SOG and FB and actin transcripts could not be detected for SOG and FB samples. Preliminary studies involving the optimisation of the localisation experiment on the immature male tissues showed that the actin gene could not be amplified from G^{MH} (result not shown) and was subsequently omitted from the final experiment. The overall result shows that NP in *N. viridula* is expressed in most tissues in males and females (mature and immature). In some tissues, however like the immature male SOG and FB as well as the mature female G^{MH}, the actin controls could not be amplified, indicating that there were problems associated with either the mRNA or the cDNA synthesis.

University of Cape Town

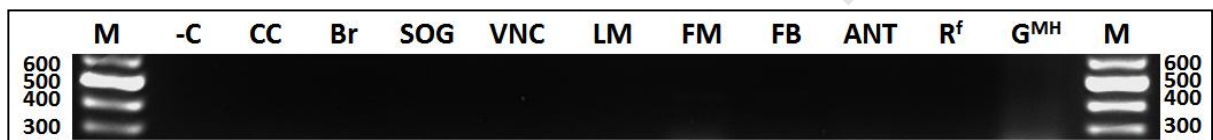
A: Mature female



B: Mature male



C: Immature female



D: Immature male

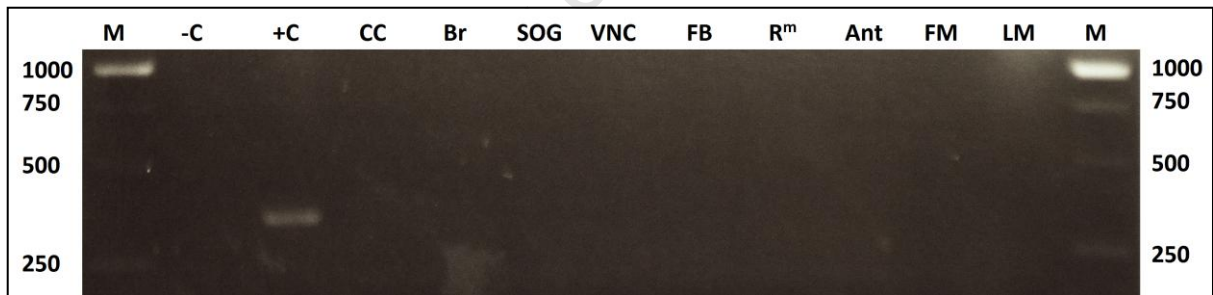


Figure 23: Represents the DNA contamination actin PCRs for various *N. viridula* (A) mature female, (B) mature male, (C) immature female and (D) immature male tissues. Abbreviations: +C - positive control, -C - negative control, SOG - Sub-oesophageal ganglion, VNC - ventral nerve cord, CC - corpora cardiaca, Br - brain, FM - flight muscle, LM - leg muscle, G^{MH} - mid and hind gut, Ant - antennae, R^m - male reproductive organs, R^f - female reproductive organs, FB - fatbody and M - marker. The sizes on the left for each band of the marker are measured in base pairs.

Mature females:

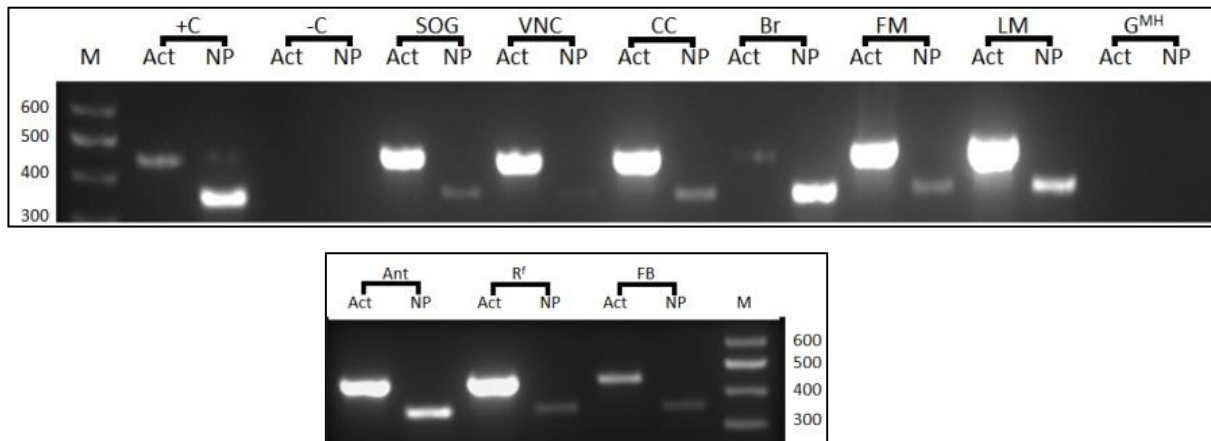
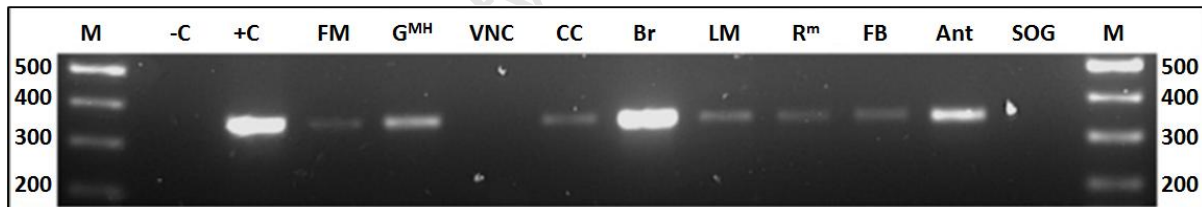


Figure 24: Represents the expression profile for the NP gene in various mature female *N. viridula* tissues. Abbreviations: -C - negative control, +C - positive control, SOG - Sub-oesophageal ganglion, VNC - ventral nerve cord, CC - corpora cardiaca, Br - brain, FM - flight muscle, LM - leg muscle, G^{MH} - mid and hind gut, Ant - antennae, R^f - female reproductive organs, FB - fatbody and M - 100 bp marker. Note that the band at 450 bp in the NP +C is spillage from the actin well. The sizes on the left for each band of the marker are measured in base pairs.

Mature males:

A - NP



B - Actin

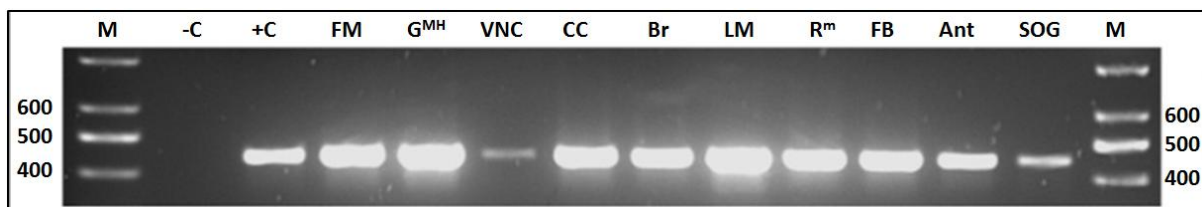


Figure 25: Represents the expression profile for the NP gene in various mature male *N. viridula* tissues (A) and the Actin control (B). Abbreviations: -C - negative control, +C - positive control, FM - flight muscle, G^{MH} - mid and hind gut, VNC - ventral nerve cord, CC - corpora cardiaca, Br - brain, LM - leg muscle, R^m - male reproductive organs, FB - fatbody, Ant - antennae, SOG - Sub-oesophageal ganglion and M - marker. The sizes on the left for each band of the marker are measured in base pairs.

Immature female:

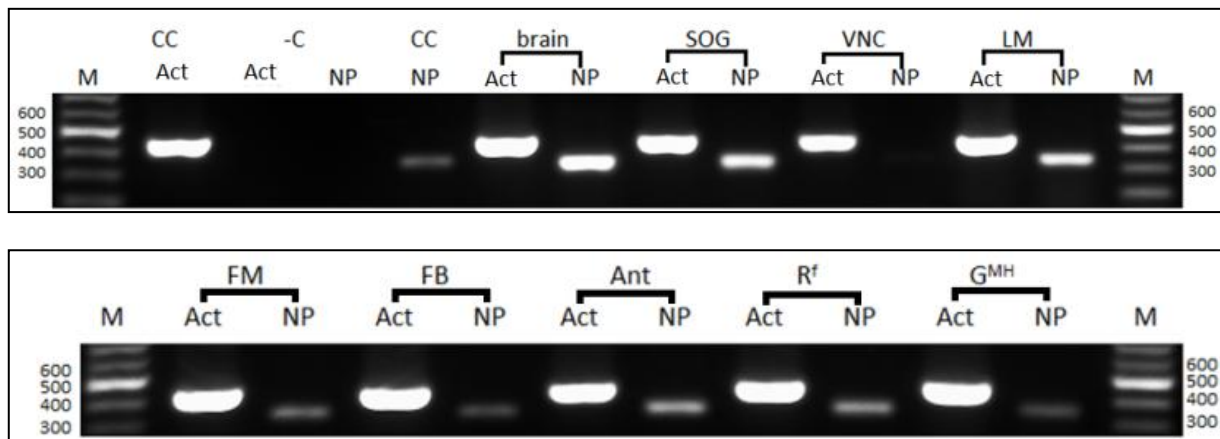


Figure 26: Represents the expression profile for the NP gene in various immature female *N. viridula* tissues. Abbreviations: CC – corpora cardiaca, -C – negative control, Br – brain, SOG - Sub-oesophageal ganglion, VNC – ventral nerve cord, LM – leg muscle, FM – flight muscle, FB – fatbody, Ant – antennae, R^f – female reproductive organs, G^{MH} – mid and hind gut and M – 100 bp marker. Note that the CC – NP and actin are from the same cDNA sample but were loaded on either side of the -C. The sizes on the left for each band of the marker are measured in base pairs.

Immature males:

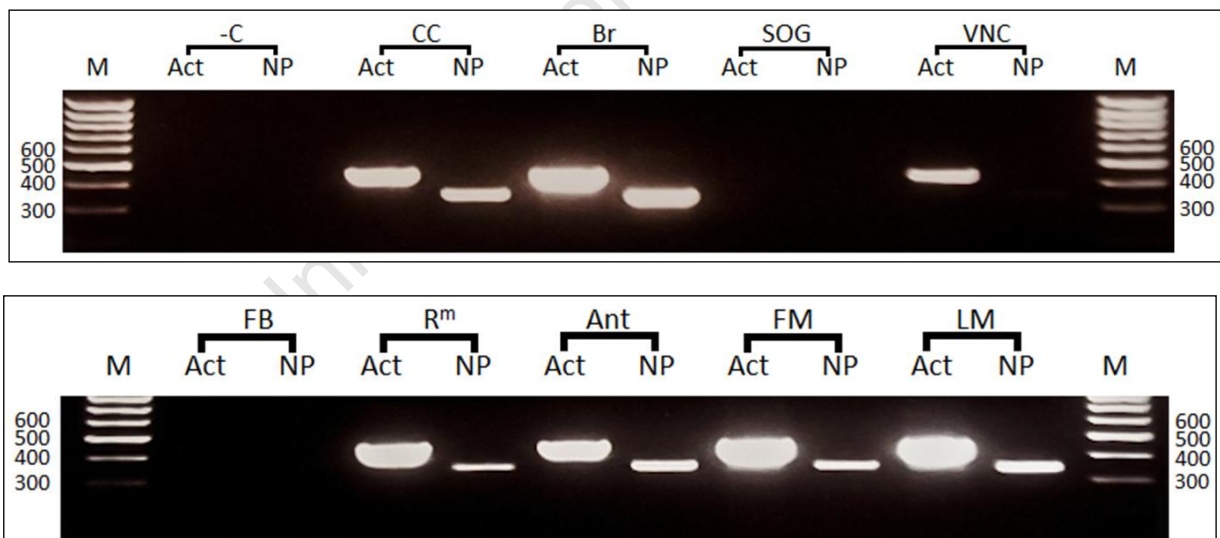


Figure 27: Represents the expression profile for the NP gene in various immature male *N. viridula* tissues. Abbreviations: -C – negative control, CC – corpora cardiaca, Br – brain, SOG - Sub-oesophageal ganglion, VNC – ventral nerve cord, FB – fatbody, R^m – male reproductive organs, Ant – antennae, FM – flight muscle, LM – leg muscle, and M – 100 bp marker. The sizes on the left for each band of the marker are measured in base pairs.

3.3 PCR-based cloning and sequence analysis of *Jasus lalandii*

3.3.1 Actin

A house keeping gene was also required to ascertain whether the mRNA from *J. lalandii* tissues was of good quality when the gene of interest needed to be PCR amplified. Using degenerate primers designed by Chaty et al. (2004), the actin gene was partially amplified from X-organ tissue resulting in a amplicon of ~900 bp, which was sent for sequencing. The sequence of the ~900 bp amplicon had 99% homology to other crustacean actin genes like *Litopenaeus vannamei* when queried in BLASTx. Non-degenerate primers (Jlactinf and Jlactinr) were designed based on this sequence, to amplify a shorter 458 bp fragment of the *J. lalandii* actin. The 458 bp actin amplicon (Figure 28) was verified via sequencing to confirm the identity of the PCR product (Figure 29).

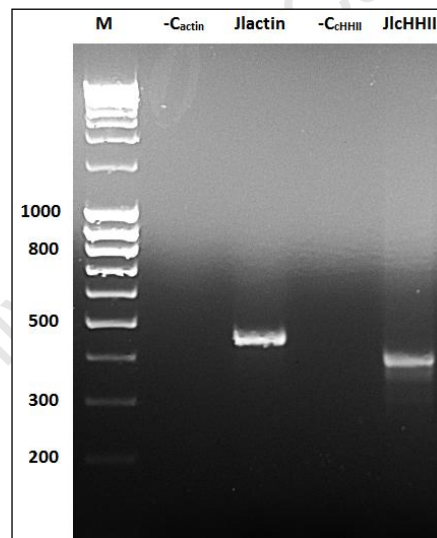


Figure 28: Amplification of the *J. lalandii* actin and cHH-II gene from x-organ tissue. Abbreviations: M - marker, -C - no template control, Jlactin - *J. lalandii* actin gene and *J. lalandii* cHH-II gene. The sizes on the left for each band of the marker are measured in base pairs.

```

Jlactin f →
TGCTACTGACGGAGGCTCCCCTCAACCCCAAGGCCAACCGTGAAAAGATGACCCAGATCA
  L L T E A P L N P K A N R E K M T Q I

TGTTGAAACCTTCAACACCCCGCCATGTACGTTGCTATCCAGGCTGTGCTCTCCCTGT
M F E T F N T P A M Y V A I Q A V L S L

ACGCTTCGGCCGTACCACTGGTATTGTGCTGGACTCTGGTGATGGCGTGTACACACCG
Y A S G R T T G I V L D S G D G V S H T

TTCCCATCTACGAGGGCTACGCTCTTCCTCAGCTATCCTGCGTCTGGATTTGGCTGGAC
V P I Y E G Y A L P H A I L R L D L A G

GTGATCTTACTGACTACCTTATGAAGATCCTTACAGAACGTGGCTACACCTTACAACCA
R D L T D Y L M K I L T E R G Y T F T T

CCGCCGAGCGAGAAATCGTTCGTGACATCAAGGAAAACTTCGCTATGTTGCCCTTGACT
T A E R E I V R D I K E K L R Y V A L D

TCGAACAGGAGATGACTACTGCCGCTTCATCATCATCACTAGAAAAGTCTACGAACCTC
F E Q E M T T A A S S S S L E K S Y E L

← Jlactin r
CCGACGCCAGGTGATCACCATCGGCAACGAGAGGTTTC
P D G Q V I T I G N E R F

```

Figure 29: Represents the cDNA and its deduced amino acid sequence of the decapod crustacean *J. lalandii* partial actin gene. The purple boxes indicate primer binding sites and arrows indicate their direction.

3.3.2 Crustacean hyperglycaemic hormone II

It is known from literature that the RPCH is one of many hormones produced in the X-organ of decapod crustaceans (Marco et al., 2003). Since the X-organ is a microscopic and diffuse piece of neural tissue, it is fairly difficult to identify and dissect. To verify that the X-organ was indeed dissected from the eyestalks of *J. lalandii*, the Crustacean hyperglycaemic hormone II (cHH-II) gene was used as a positive control in the initial PCR reaction of X-organ cDNA. The primers used to amplify the cHH-II gene are JlcHHIIif and JlcHHIIr designed by Marco et al (2003). The expected amplicon size according to Marco et al (2003) was ~400 bp, which was what was amplified in the PCR reaction in Figure 28.

3.3.3 Red-pigment concentrating hormone

A partial RPCH sequence had been amplified by Dr. Marco in 2007 (personal communication). This partial RPCH sequence of ~ 197 bp, which contained the nucleotide sequence encoding the mature Panbo-RPCH amino acid sequence, was amplified using 5' RACE PCR. The partial sequence was verified by PCR amplification, with primers JLRPCHf and JLRPCHr (based on the partial RPCH sequence) and sequencing of the resulting ~200 bp amplicon. The rest of the RPCH 3' end of the preprohormone sequence of the decapod crustacean *J. lalandii* was identified using the 3' RACE PCR (Roche). Identifying the 3' region of the RPCH preprohormone was achieved by using primers JLRPCHf and Jlf2 in a nested 3' RACE PCR reaction. The 3' RACE nested PCR amplified a DNA fragment of ~627 bp. After sequencing, the 3' RACE fragment was aligned with the 5' RACE PCR. The two fragments produced a consensus sequence, which was used as a template to reconfirm via PCR-amplification with a high fidelity *Taq* polymerase and sequencing of the complete open-reading frame using primers JLRPCHf and JLRPCHr5. Using 3' and 5' RACE it was possible to amplify the full *J. lalandii* RPCH preprohormone sequence. The RPCH preprohormone open-reading frame contains 297 bp and encodes a peptide containing 98 amino acids (Figure 30). The peptide contains the 21 amino acid signal peptide (green box), the mature Panbo-RPCH octapeptide (grey box), a dibasic cleavage site (blue box) and a 66 amino acid 'RPCH precursor related peptide' (RPRP in italics).

JlRPC hf →

AAAACACACTGCAACAGGTGAGGCCAGCCAACAGCCAGCCACTCAGCCCCTGCACCACA
GCAGCCGTCACCACCTCCCGTCACCACCCGCCACACCACACTCGCTCGTCCAGCC

Jlf2 →

ATGGTTCGTGCCGGCGTCGCCCTTCTCTTGGTGGTGGTGGTGGTGGCCGCCAGCGTCTCC
M V R A G V A L L L V V L V V A A S V S

← JlRPC hr

GCCCAGCTCAACTTCTCCCAGGTTGGGGCAAGCGGACTGCGGGCCCGCAGGCGGCACC
A Q L N F S P G W G K R T A A A A G G T

GACCCCGCCGAGCCGCCCTCCGCGCCCCAGCAGCCCTAGCCATGGGGCCCTCCTCTCCT
D P A A A A L R A P A A L A M G P S S P

GCCGTCGGCGACGCATGCGGGGCCATCCCAGTCTCCACCGTCATGCACATCTACAGGCTC
A V G D A C G A I P V S T V M H I Y R L

ATCAGGAGCGAAGCGGCGGGCTGGCCCCAGTGTGTCAGGACGAGGAGTACCTGGGGCTAG
I R S E A A R L A Q C Q D E E Y L G +

← JlRPC hr5

ACTCCCACCATCACCAGGCCTCCCACCCACCCCTCAGTCAACCACTGTTCATTGTTCTAA
AATTACTGTGTGATTCTCTGTAACATCGTTAACCGTATGTATGTATCTAGGGCTGAACCG
TACCCTGCAGTCTCGCCTGGTCTCTCCGACGCCCGCCGAGCTCTGGACCTATGTAAC
GTATCATGAGATTCATATTTCTATTTTTATCTCACCTTTATGAAACATGAGGGCTGGAAT
ATCTGGATCAGAGATGCTGCACTTGATTAACTAGTCCATGCAAAAAAAAAAAAAAAAAA

Figure 30: Represents the cDNA and its deduced amino acid sequence of the crustacean *J. lalandii* RPC h preprohormone. The underlined nucleotide sequences represent the 5' and 3' UTR, the green box represents the predicted signal peptide, the grey box the mature RPC h peptide sequence, the blue box the di-basic cleavage site, italicized amino acids the RPRP sequence, purple boxes indicate primer binding sites and arrows indicate their direction. Bold nucleotides represent the start and stop codons. The signal peptide for the amino acid sequence was predicted by SIG-Pred (http://bmbpcu36.leeds.ac.uk/prot_analysis/Signal.html).

3.3.4 Neuroparsin

In the experiments involving the 3' RACE PCR amplification of RPC h preprohormone, the Roche 3'/5' RACE kit was used. The primers R9 and Jlf1 were used in the initial experiments involving the amplification of the 3' end of the RPC h using 3' RACE PCR. The use of the Jlf1 primer resulted in a 1200 bp amplicon. Although other bands were present when electrophoresed on an agarose gel, the 1200 bp band was the brightest and other bands were unclear and not as clearly defined as the 1200 bp amplicon (Figure 31). The 1200 bp was sent for sequencing and the resulting sequence showed a 55% homology to the *S. gregaria* and many other NP sequences. The primer Jlf1 had bound in the 5'-3' direction (forward primer) in the 5' UTR and the R9 primer bound to the linker sequence R8. The Jlf1 primer bound specifically enough to amplify the NP prohormone gene entirely and include a part of the 3' and the whole 5' UTR. Using a high fidelity Taq polymerase and gene specific primers (JlNPf and JlNPr – which do not encompass the region where Jlf1 bound) the *J. lalandii* NP open-reading frame (ORF)

was verified in a PCR reaction and subsequent sequencing. Although the nucleotides encoding the Jlf1 primer were not verified along with the *J. lalandii* NP sequencing, it is marked in Figure 32. The NP ORF was queried in Blastx and showed highest homology to *S. gregaria* (57%) and putative crustacean NPs. The NP ORF contains 312 bp which encode a 103 amino acid peptide containing 12 cysteine residues positioned at conserved positions indicative of the NP family and high sequence homology to other crustacean NP (putative).

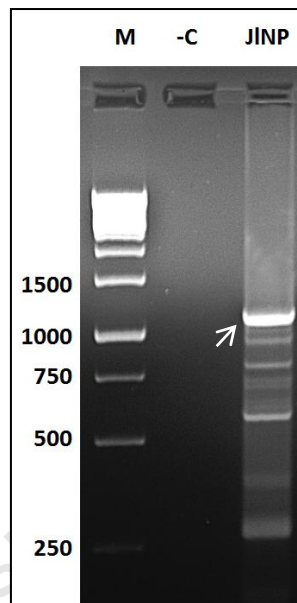


Figure 31: shows a RACE PCR reaction whereby the *J. lalandii* NP was amplified. Abbreviation: M – marker, -C – no template control and JINP - *J. lalandii* neuroparsin. The sizes on the left for each band of the marker is measured in base pairs and the white arrow indicates the 1200 bp amplicon encoding the Jasla-NP gene.

J1f1 → CCGTCACCACTCCCGTCACCACCGCCAGTAGC CCAGGTTGGACGCAGCACCAAACCAC
 CACCATCACCACC
ATGAGGACTTCCTACTCCAGCGGCGTACCTTCCTCCTCGCCTCCTG
M R T S Y S S G V T F L L A S C
 CTCCTCCTCCTCCTCCTGCAATATGCAGCAGCCGGTCTGCCTGTCCACACCGGAATGA
S L L L L L Q Y A A A G P A **C** P H R N E
 AATTGTGCCCGAAGATCTCAGCCAGTGCAAGTACGGCGTTGTGCTAGGCTGGTGCGGCAA
 I V P E D L S Q **C** K Y G V V L G W **C** G N
 CAAGGCTGCGGCAAGGGCCAGGAGACACGTGTGGAGGACGCTGGGAGCAGCACGGGAT
 K A **C** G K G P G D T **C** G G R W E Q H G I
 CTGCGGTGAGGGGATGTACTGCGTGTGTGGCCACTGCGCCGGGTGCTCCAGCAAGCTCAA
C G E G M Y **C** V **C** G H **C** A G **C** S S K L K
 GTGTGCCTTGGGCAGGTTCTGCT**TAG**
C A L G R F **C** +

← J1Npf

TGGTGCATCTGGCTCCCTAGAACCCTACCAGTGGCAGCACTTGGCACTGGCACTCAGCAG
 CTGCTGTGAAATTGCACAGGCTGCTTTAAAACCTTTGCAATGCGTCTCTGACAGCTTTTG
 GTAACATAAATTTCTTTACTACTACTACTACTCCTACTATTGCTTCCATTCATGGAG
 GGACGTTGACCAAAATCGATTTTTTTGACCGTCAGGTCCTTTGTGACAGGTTGCTGTCTGG
 GCCATCAGCAGCAGCAAGCGCCTCCTATACCTACGTACAAGTAATGAGGCTTTCTCTGTA
 TGGTGGAAATTTCTTTCTGTCACGTCCTTCTGCTCCCTACGCTTGCATTTCTTTTTTC
 TTTTATCTGTCATGCTCTTGCTGCCATCACACTGATGAAATGCCAGAGAAGAGGAGGCAG
 TTGCCAAGTGACCACCATGACTATTAATGTTATGCGGGAAGAGCAAAAAACATAAAAAATG
 TTTTCTGGTAAATGTGACACCGACACCTACGGCATGTCCAAGACGGCATCTAAAACCTAC
 TGCTTTGTGATAATTAGTTTGCATAAGAAGCTGGAAGCGAACATAACTCACAGGATCTCC
 AGGCTGGACCGGACCTCCGGGAACCGGACTCTACGGAACCGGCCAGTGATACATATATAC
 GTAAACACATACCTGATGTGTCTCAACAAATCTTTAAGAAATAAATTGAGAAATGTAAAA
 AAAAAAAAAAAAAAAAAA

Figure 32: Nucleotide sequence of the decapod crustacean *J. lalandii* NP prohormone cDNA and its deduced amino acid sequence. The underlined sequences represent the 5' and 3' UTR, the green box represents the predicted signal peptide, the red boxes the conserved cysteine residues, the yellow box the putative polyadenylation signal, the purple boxes the primer binding sites and letters in bold represent the start and stop codons. The predicted signal peptide of the amino acid sequence was determined by SIG-Pred (http://bmbpcu36.leeds.ac.uk/prot_analysis/Signal.html)

3.3.5 Localisation of NP and RPCH in *Jasus lalandii*

To analyse and compare the expression profiles of the NP prohormone between *J. lalandii* and *N. viridula* as well the RPCH preprohormone between *J. lalandii* and *D. pulex*, various tissues (X-organ, brain, sub-oesophageal ganglion, thoracic ganglion, abdominal muscle, heart, hepatopancreas, testes, ovaries, gills and midgut) were dissected. From these tissues, the total RNA was extracted and the total RNA was DNase treated. The total RNA was tested for amplification of the actin gene in a PCR reaction; the results of which are negative (Figure 33). The expected sizes for the *J. lalandii* actin, RPCH and NP amplicons are 450, 427 and 400 bp respectively. A previous DNase treated X-organ sample was used as a positive control (Figure 34). The actin transcript was detected in all tissues (Figure 34), except the gill sample and actin was only faintly detected in the midgut tissue. The RPCH transcript was only detected in the X-organ and the SOG. The NP was detected in the X-organ, brain, SOG, ThG, ovaries and in the heart.



Figure 33: Represents the DNA contamination actin PCRs for various *J. lalandii* tissues. Abbreviations: Br - Brain, SOG - sub-oesophageal ganglion, ThG. - thoracic ganglion, Ova - ovaries, Abm - abdominal muscle, HepP. - hepatopancreas, XO - X-organ M - marker and r - redundant lane. The sizes on the left or right for each band of the marker are measured in base pairs.

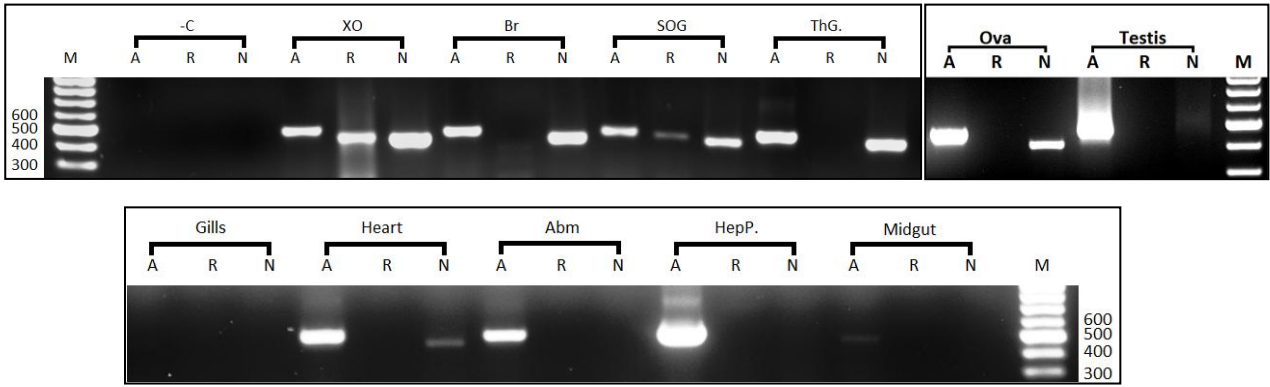


Figure 34: The expression profile of the RPCH preprohormone and NP prohormone for *J. lalandii* in various tissues. Abbreviations: XO - X-organ, Br - Brain, SOG - sub-oesophageal ganglion, ThG. - thoracic ganglion, Abm - abdominal muscle, HepP. - hepatopancreas, A - actin, R - RPCH, N - NP and M -marker.

University of Cape Town

4. Discussion

The main aim of the study was to determine the amount of homology that exists between “higher” and “lower” crustaceans and insects. To achieve this aim, I cloned neuropeptides from two peptide families namely NP and AKH/RPCH, and used the sequence information to identify how closely or distantly related the higher and lower crustaceans are to the insects. Furthermore, I wanted to determine whether these neuropeptides are produced in the same or similar organs in insect and crustaceans and, thus, also investigated the spatial expression profiles of various genes. I used the following animals in this study: *D. pulex*, as an example of a lower or evolutionary older crustacean (a cladoceran); *J. lalandii*, as an example of a higher or evolutionary younger crustacean (a decapod) and *N. viridula* for the insect candidate. In the decapod crustaceans the mature RPCH, Panbo-RPCH, is well conserved; the only other crustacean RPCH variant, thus far, is found in *D. pulex* (Dappu-RPCH). I wanted to investigate whether the *D. pulex* preprohormone RPCH sequence would be similar to or different to other decapod RPCH preprohormones; because the mature region of the peptide varies at some positions (see Figure 35, page 82). The heteropteran insect *N. viridula* was chosen, because it contains the same mature Panbo-RPCH as the decapod crustaceans. Thus, the question to be answered is whether the homology is limited to the mature peptide region or does it extend to the rest of the RPCH preprohormone sequence. With regards to the RPCH sequence data from *N. viridula* and *J. lalandii* it would be possible, perhaps, to ascertain whether the Dappu-RPCH preprohormone sequence would be more homologous and cluster either with the insect AKH or to the decapod crustacean RPCH. During the course of the study, the comparison of the AKH/RPCH was extended to include the NP peptides.

The study focused on using neuropeptides to ascertain whether certain phylogenetic inferences could be made between insects and crustaceans. Although still in its infancy stage neuropeptides have been used to strengthen neuropeptide phylogenetic studies within insects orders (Roth et al., 2009; Predel et al., 2012). This is especially important when one wants to study how neuropeptides change between organisms. Also for the understanding, important structural features which could become conserved between

organisms, it is thus important to understand the evolution of neuropeptide sequence and structure. Understanding why there are so many homologous neuropeptides in certain organisms is also only possible if we can understand and track the lineage of the gene phylogenetically through history. Other proteins (18s rDNA, 16s rDNA, cytochrome c oxidase I, etc.) commonly used in phylogenetic studies were thought to be beyond the scope of the study and were therefore not pursued since the aim was to understand the evolution of neuropeptides within arthropods.

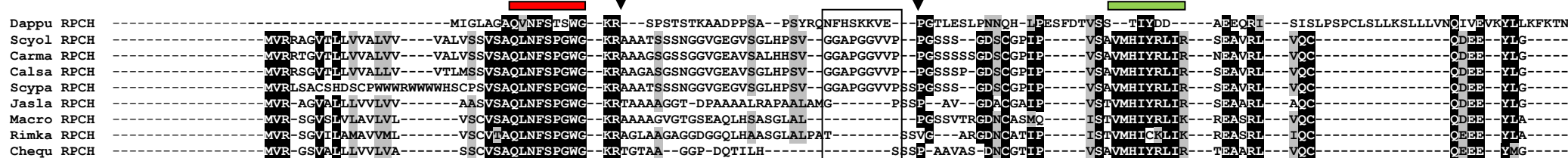
4.1 AKH/ RPCH homology

Initially to obtain the RPCH and AKH preprohormone sequences from *J. lalandii* and *N. viridula* the splinkerette method was used. This method used genomic DNA to amplify the target gene from *N. viridula*. The splinkerette system was used initially because the AKH/RPCH gene would be more abundant in DNA compared to the mRNA transcripts. Also, it would be easier to dissect and work with muscle tissue over the microfine CC. To design primers used in the splinkerette PCRs, an alignment of AKH preprohormone cDNA sequences was generated from various insects. The degenerate primers used in the investigation were designed based on the only conserved region available, i.e. nucleotide sequences encoding the residues of the mature AKH peptides (red box in Figure 35A). The outcome of the splinkerette PCR experiments was that the method and conditions were not effective enough to amplify the AKH gene. Most of the PCR's reactions produced multiple non-target bands and on occasion smearing when electrophoresed on agarose gels. The smearing was likely due to the partial digestion of genomic DNA with restriction endonucleases. It was decided to abandon further work with the splinkerette PCR system and proceed with 3' and 5' RACE PCR experiments instead on tissue where mRNA is targeted. The same degenerate primers that were used in the splinkerette PCR experiments were used in attempts to amplify the preprohormone AKH gene in *N. viridula*. The CC tissue was used to amplify the AKH preprohormone as it is enriched with AKH mRNA, thus it seemed more likely that the transcript would be amplified, if there were greater abundance of similar length transcripts, as is the case with RACE PCRs. Similar to the splinkerette system, however the results from both 3' and 5' RACE PCR were unsuccessful in amplifying the AKH in *N.*

viridula. The likely reasons why these two systems didn't work may be due to the nature of the region that the primer sequences were designed against (red box in Figure 35B). The region that encodes the AKH mature peptide contains one region with four consecutive cytosine residues and another region with four consecutive guanine residues within the coding sequence. Designing primers outside of these two regions would make the primers short and produce primers with low melting temperatures, which would result in increased non-specific amplification or no amplification at all at higher temperatures. The high amount of C and G nucleotides in the primer may have bound to unspecific DNA sites, or very likely formed hairpin loop at the 3' end, or could have formed primer dimers, all of which make the primer unusable (Innis & Gelfand, 1990). In one particular RACE PCR whereby a semi-nested PCR approach was used with two different degenerate primers and the AUAP primer used twice as a reverse primer, PCR bands were generated, when electrophoresed on an agarose gel. When one particular band was sent for sequencing, upon a Blastx search of the nucleotide sequence it was discovered that the band was the partial sequence of the NP prohormone family. The NP story will continue in the NP section of the discussion. When analysing the closely related orders to Hemiptera (for *N. viridula*) one would be hard pressed to find regions which overlap with enough homology to which a primer could be designed. Even when taking the hemipteran Rhopr-AKH sequence into account, it is evident that the peptide sequence does not even match that of the *N. viridula* Panbo-RPCH peptide sequence.

Initially when attempting to amplify the 3' end of the *J. lalandii* RPCH preprohormone in 3' RACE PCR, the template cDNA used came from only two whole *J. lalandii* eyestalks (ommitidia removed). The whole eyestalk was used, because there was greater confidence that the X-organ was indeed still present. The problem with using the whole eyestalk tissue was that the RPCH transcripts were not enriched enough and only after dissecting and using X-organ samples for synthesising cDNA were the PCR amplifications successful. In one particular PCR reaction, using the enriched X-organ cDNA and the Jlf1 and the R9 primers, the NP transcript was amplified from *J. lalandii* (the NP story continues in the NP section 4.3). When experiments using nested 3' RACE PCR with primers JIRPCH/ Jlf2 versus R9 were repeated, the 3' end of the RPCH was discovered.

A.



B.

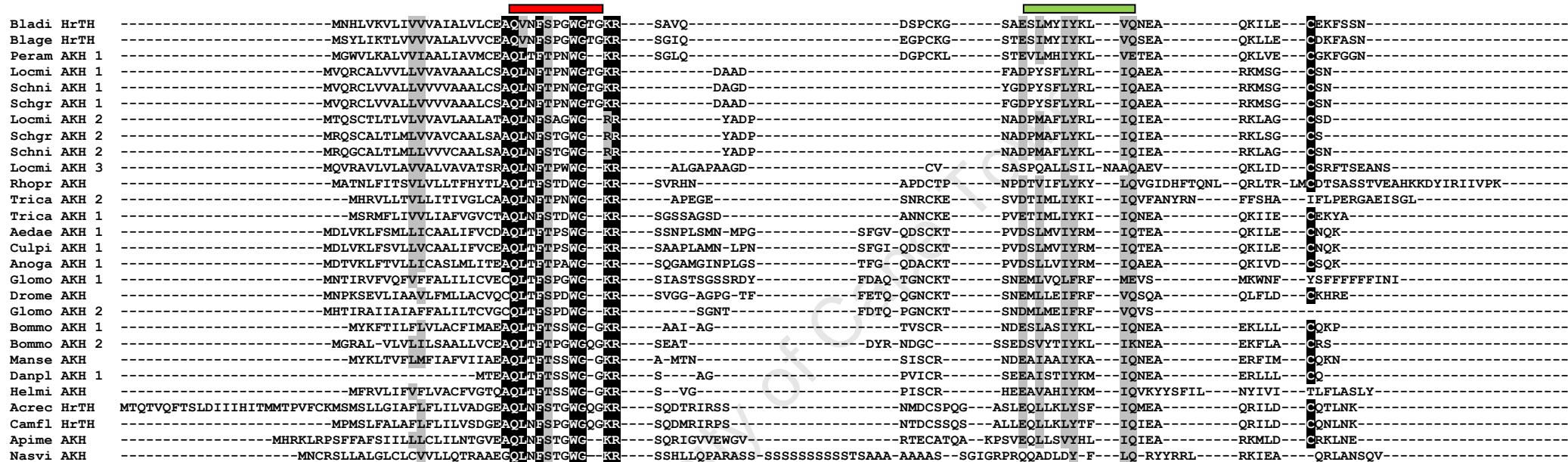


Figure 35: Alignment showing the AKH preprohormone amino acid sequence from *D. pulex* against the crustaceans RPCH preprohormone (A) and the insects AKH preprohormone (B). The letters in black boxes represent amino acids that are identical, the grey boxes amino acids which are similar in type (i.e. their charge and polarity); the red box represents the mature AKH/RPCH peptide regions and the green box a region of interest. The amino acids between the arrows represent a region where the crabs and the shrimp differ in sequence and the region in the clear box a specific region where sequence is missing in the shrimp compared to the crabs. Abbreviations are as follows: Acres: *Acromyrmex echinator*, Aedae: *Aedes aegypti*, Anoga: *Anopheles gambiae*, Apime: *Apis mellifera*, Bladi: *Blaberus discoidalis*, Blage: *Blattella germanica*, Bommo: *Bombyx mori*, Calsa: *Callinectes sapidus*, Camfl: *Camponotus floridanus*, Carma: *Carcinus maenas*, Chequ: *Cherax quadricarinatus*, Culpi: *Culex pipiens*, Danpl: *Danaus plexippus*, Dappu: *Daphnia pulex*, Globo: *Glossina morsitans morsitans*, Helmi: *Heliconius melpomene melpomene*, Jasla: *Jasus lalandii*, Locmi: *Locust migratoria*, Macro: *Macrobrachium rosenbergii*, Manse: *Manduca sexta*, Nasvi: *Nasonia vitripennis*, Peram: *Periplaneta americana*, Rhopr: *Rhodnius prolixus*, Rimka: *Rimicaris kairei*, Schgr: *Schistocerca gregaria*, Schni: *Schistocerca nitans*, Scyol: *Scylla olivacea*, Scypa: *Scylla paramamosain*, Trica: *Tribolium castaneum*. List of accession numbers: Scyol RPCH: E5LDB7, Carma RPCH: Q26324, Calsa RPCH: Q23757, Scypa RPCH: E2JEQ4, Macro RPCH: D8KWQ8, Rimka RPCH: E6YB85, Chequ RPCH: Q4VS21, Bladi HrTH: Q17128, Blage HrTH: F4MJ53, Peram AKH 1: Q5EY02, Locmi AKH 1: P55319, Schni AKH 1: P18830, Schgr AKH 1: P18829, Locmi AKH 2: P08379, Schgr AKH 2: P35808, Schni AKH 2: P35807, Locmi AKH 3: P19872, Rhopr AKH: D0VBL8, Trica AKH 2: D6WNV7, Trica AKH 1: B8K1Y2, Aedae AKH 1: CAY77165.1, Culpi AKH 1: CAY77163.1, Anoga AKH 1: Q5TTQ9, Globo AKH 1: F8SKD3, Drome AKH: NP_523918.1, Globo AKH 2: F8SKD4, Bommo AKH 1: A8WFD9, Bommo AKH 2: B3IUD5, Manse AKH: P67788, Danpl AKH 1: G6DTY9, Helmi AKH: HE669321, Acres HrTH: F4WLC6, Camfl HrTH: E2AMJ9, Apime AKH: G9JXB9, Nasvi AKH: ADM26613.1.

4.2 AKH/ RPCH phylogeny and localisation

There are only a few RPCH preprohormones known and listed in the databases; these are shown in Figure 35A along with that cloned from the South African spiny lobster, *J. lalandii*, from this study. When the RPCH preprohormones are aligned, there is greater conservation visible at the N-terminus in comparison to the C-terminus (Figure 35A). The most variable region occurs before and after the RPCH peptide (between the arrows in Figure 35A). *J. lalandii* RPCH preprohormone appears to sit in the middle between the prawns and crayfish (*Macrobrachium rosenbergii*, *Rimicaris kairei* and *Cherax quadricarinatus*), and the crabs (*Scylla paramamosain*, *Callinectes sapidus*, etc.). It is evident from Figure 35A, that crab sequences share a high degree of homology, even in the region between the arrows. *J. lalandii*, the prawns and crayfish, apparently do not have the same conserved amino acids in the region between the arrows in Figure 35A. Furthermore, a region indicated by the clear box in Figure 35A is absent in *J. lalandii*, the prawns and crayfish compared to crab RPCH sequences. It appears that the *J. lalandii* RPCH preprohormone models the prawns and crayfish more than the crabs.

In contrast to the crustaceans, many more AKH preprohormone sequences have been elucidated in insects, some of which are shown in Figure 35 B. Nearly all the crustacean RPCH preprohormone sequences discovered thus far, come from the decapod order. This is likely the reason for the higher degree of conservation observed in crustaceans compared to the AKH preprohormones of insects (Figure 35). Therefore, in the study the comparison of whether *D. pulex* is more closely related to insects or crustaceans can only make inferences based on the decapod order specifically. When trying to align *D. pulex* RPCH preprohormone with either the insects or crustaceans, only the RPCH amino acids align. It is, thus, difficult to ascertain whether *D. pulex* RPCH preprohormone should be grouped closer to the insects or the decapod crustaceans as there is no clear-cut clustering that can be seen. When comparing *D. pulex* RPCH preprohormone to the insect sequences, it appears that the insect sequences are shorter in length. Also, when comparing the insect against the crustacean sequences, there appears to be a region well conserved in crustaceans and another region partially conserved in the insect alignment (see Figure 35, indicated by the green box). Knowing this and analysing the *D.*

pulex sequence, it is evident that *D. pulex* does not fit in either the conserved region of insects or crustaceans. Knowing this, it is still perhaps possible to answer the original question in spite not being able to amplify and obtaining the AKH preprohormone sequence from *N. viridula*. When looking at the overall construction of the AKH/ RPCH preprohormone, does *D. pulex* cluster more with the crustaceans or with the insect? As is evident with the comparison detailed above, it can be seen that *D. pulex* does not cluster with any conserved domains in both insects and crustaceans and thus one cannot draw definite conclusions whether *D. pulex* clusters more with insects or crustaceans with respect to the AKH/RPCH preprohormone. That said, the amount of information gathered from insects, it is clear that a lot more research on AKH has been carried out in insects and across multiple orders compared to the crustaceans. Considering that RPCH has been investigated mostly in the decapod order, it is harder to make any good assumptions in crustaceans and *D. pulex*.

When comparing how the insects cluster in Figure 36, it is evident that despite the very low bootstrap values between them, the various AKH preprohormone sequences still cluster in their respective orders. Keeping in mind that it is hard to make inferences based on the bootstrap values, Dappu-RPCH in Figure 36 is situated before the node where the insects and crustaceans diverge and does not give possible insights into which way Dappu-RPCH may tend. Marco & Gäde (2010) observed that when the Dappu-RPCH peptide was injected into the insect *N. viridula*, it was able to mobilise lipids in the insect. When the authors repeated the experiment in the shrimp *P. pacificus* they observed that Dappu-RPCH would not concentrate any of the pigments (red, brown, yellow and blue) in the epithelium. Further analysis revealed that when a particular residue (Leu² to Val²) was changed in Panbo-RPCH there was a complete loss of the chromatophorotropic activity in *P. pacificus*. This evidence perhaps points to the possibility that *D. pulex* may cluster closer to the insects in terms of the functionality of the AKH/RPCH peptide since Dappu-RPCH can effect lipid mobilisation in *N. viridula*, but cannot concentrate pigment in *P. pacificus* due to one change in an amino acid residue.

To identify the localisation of Dappu-RPCH within *D. pulex*, whole mount *in situ* hybridisation experiments were conducted using Dappu-RPCH as probe (Figure 19,

page 61). A preliminary study showed that the RPCH probe was located within cells that lie between the thoracopods and the gut. This tissue is likely the ovaries, since gut tissue runs the length of the animal and would otherwise show a colour change from mouth to anus. It is suspected of being the ovaries, because the ovaries lie between the gut and the thoracopods and is the only other tissue which would be located close enough to this region. That said there is substantial damage to the specimen and for now we may only speculate on the type of tissue. If the RPCH is synthesised in the ovaries or neural tissues associated with the ovaries, it is not unlikely that it may be linked to some sort of reproductive function. Since no observable chromatophoric cells could be seen and since it was never observed that *D. pulex* does not change colour, one can assume that Dappu-RPCH is not linked to chromatophore cells like in the higher crustaceans like *J. lalandii*. Since it is known that the AKH in insects are used for energy mobilisation, it is not unlikely that the Dappu-RPCH may still be some linked to an energy mobilising function (like insects) in *D. pulex*. Also, since Dappu-RPCHR seems to be localised in the thoracopods it may be possible that Dappu-RPCH is involved in the locomotive capabilities of the animal with the goal of increasing energy mobilisation closest to the thoracopods.

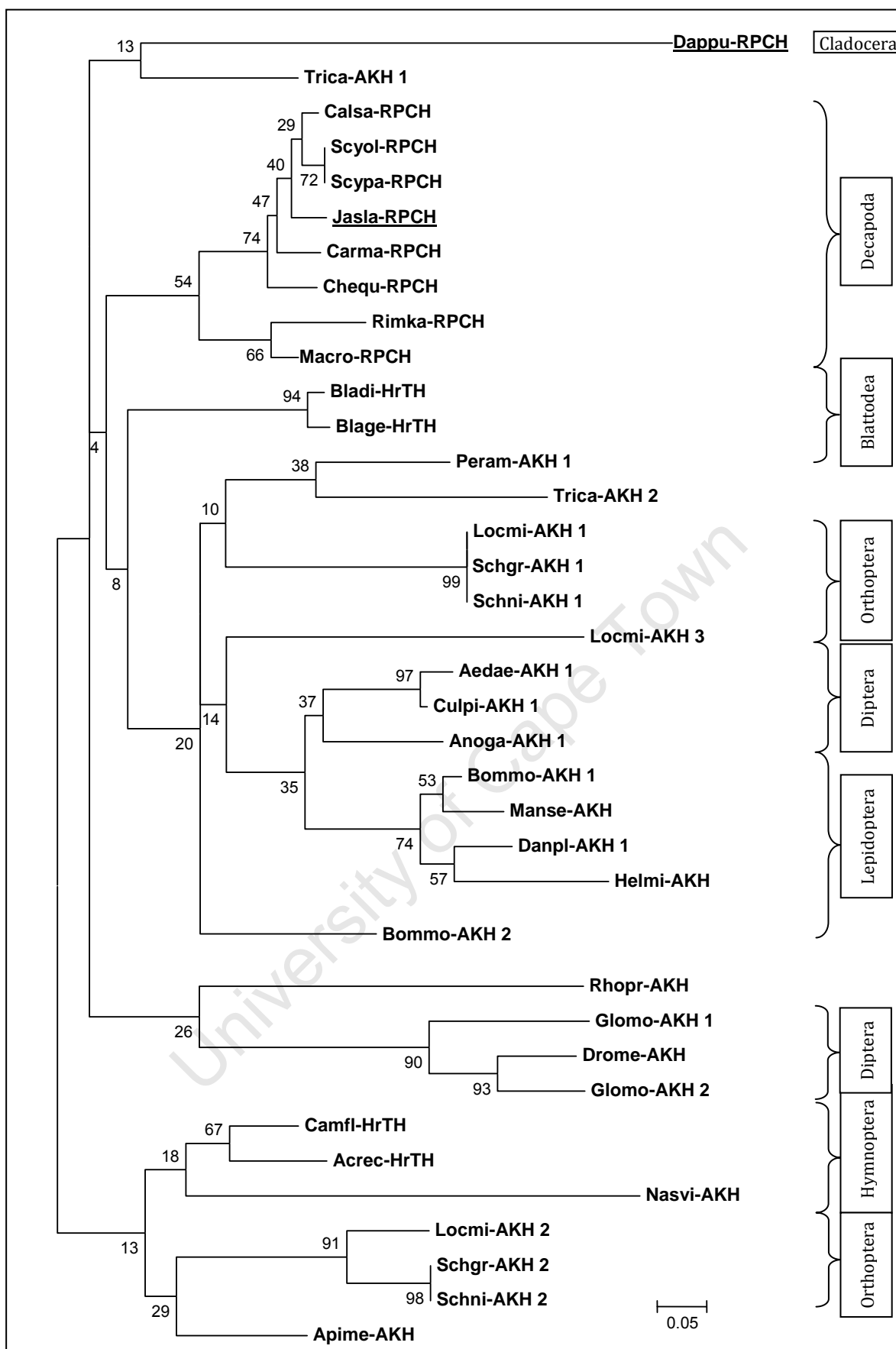


Figure 36: Phylogenetic analysis of the AKH/RPCH preprohormone family in insect and crustaceans. The phylogenetic unrooted tree was constructed using the minimum evolution method. The names of various insect orders are shown on the right-hand side and underlined RPCH occur in animals used in the study. For node support, a bootstrap analysis was performed as described in materials and methods. The number of bootstrap replicate tested was 5000, shown next to the branches. The evolutionary distances are given in units of the number of amino acid substitutions per site. Accession numbers are in Figure 35, page 82.

4.3 NP homology

The *N. viridula* NP gene was initially identified when attempting to amplify the AKH gene cDNA sequence in 3' RACE. In one particular 3' RACE PCR the nested PCR approach was used with the degenerate primers spanning the AKH nucleotide sequence. The annealing temperature for the PCR was decreased gradually in each cycle to 48 °C, which was significant enough to amplify a ~350 bp DNA fragment from CC tissue. After sequencing the 350 bp fragment, it was identified in BLASTX at NCBI as a NP because it contained the same conserved cysteine residues indicative of the NP family. Analysis of the nucleotide sequence revealed that the AUAP primer from Invitrogen™ had bound to both the sense and the antisense strand of NP. By doing 5' RACE, the rest of the 5' region of the NP gene was amplified and identified after sequencing. The *J. lalandii* NP gene was discovered in a similar way in which the *N. viridula* NP gene was identified. The primer Jlf1 was designed and used in the attempt to amplify the RPCH gene from *J. lalandii*. When the primers Jlf1 and R9 (Roche) were used to amplify the 3' end of the RPCH gene from X-organ tissue, the NP gene was amplified instead. When analysed against other crustacean and insect NP prohormone sequences, a similar distribution of cysteine residues was observed indicative of the NP family. Discovery of the *N. viridula* NP prohormone and *J. lalandii* NP prohormone cDNA sequences was fortuitous. It afforded an opportunity to investigate how peptides change and analyse their diversity between the two arthropod sub-phyla. The aim was originally to investigate a neuropeptide (Dappu-RPCH) and its homology between the decapod crustaceans and the insects. Within this context, the NP prohormone peptide sequence for *D. pulex* was also compared to the amount of homology between the decapod crustaceans and insect. From published data on NP in locusts, it seems that NP is strictly speaking, not a neurohormone, because in *S. gregaria* the NP gene is expressed in non-neural tissues like the fatbody and muscle tissue and should therefore just be called a hormone (Claeys et al., 2003). To affirm whether expression of NP occurred in multiple tissues within *N. viridula* and *J. lalandii*, the distribution of the NP transcript was investigated. Therefore, another aim was to investigate the spatial distribution of NP within *J. lalandii* and *N. viridula* and observe whether the spatial distribution of the NP gene is also conserved

between the two sub-phyla. In the study, the *D. pulex* NP gene was not verified, but it was previously PCR amplified and verified by sequencing (Dircksen et al., 2011).

4.4 NP phylogeny and localisation

Analysis of the NP prohormone sequences in a multiple sequence alignment reveals that there are little to no conserved regions within the various insect of different orders (Figure 37, page 89). Even the crustacean NP sequences are very poorly conserved (Figure 37, page 89). That said, when comparing the types of changes that occur between each sequence, more often, amino acids with similar charge and polarity are conserved. The NP gene is highly variable and it would be hard to draw any conclusions or links of how the gene might have evolved between the two subphyla. The *N. viridula* NP gene, when compared to its closest relative, *R. prolixus* only shows 54% homology to Rhopr-NP1 and 37% homology to Rhopr-NP2 and Rhopr-NP3 (Figure 37, page 89). But when comparing the *D. pulex* NP sequence against crustaceans and insects, the *D. pulex* NP sequence has the same conserved 14 cysteine residues as insects, whereas the crustaceans are missing cysteine residues C1 and C3. Even though cysteine residues C1 and C3 are not conserved in all the insect orders, a link can at least be drawn that the NP sequence of *D. pulex* mimics the insect NP more than the crustacean NP in terms of the distribution of cysteine residues.

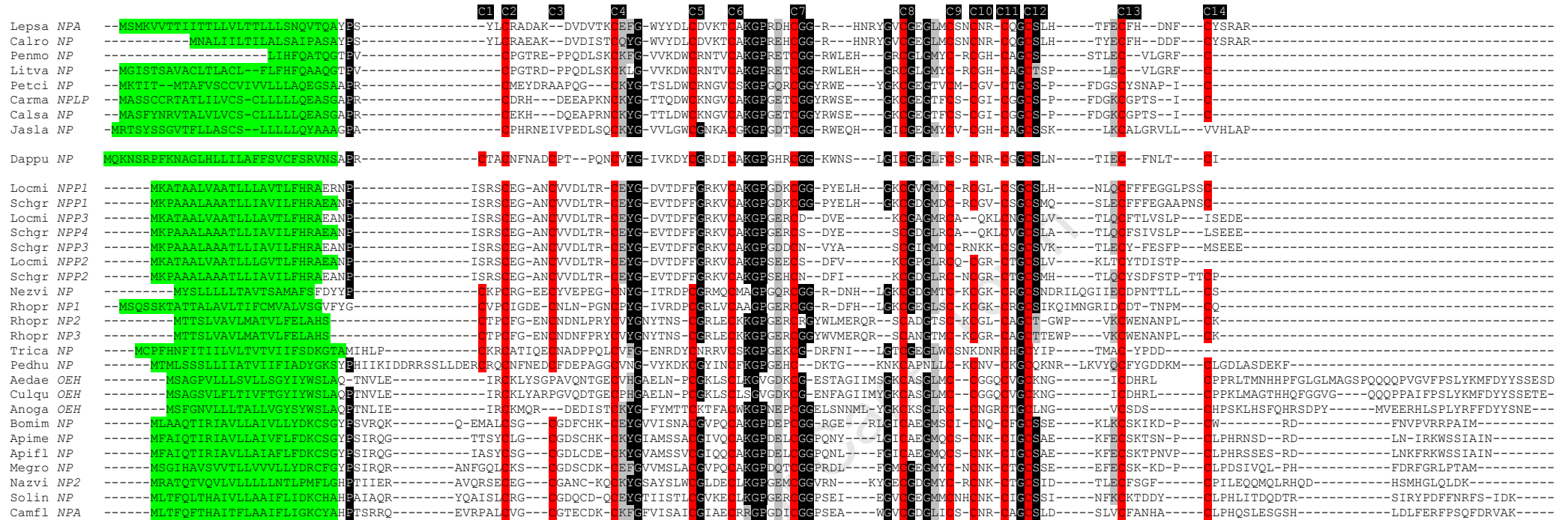


Figure 37: Alignment showing the *D. pulex* NP prohormone amino acid sequence against the crustaceans NP prohormone and the insects NP prohormone. The letters in black boxes represent amino acids that are identical, the grey boxes amino acids which are similar in type, the amino acids in green boxes the signal peptide, the amino acid red boxes the conserved cysteine residues also numbered C1 to C14. The alignment was generated in BioEdit, Gaps (-) have been introduced to maximize the alignment and the signal peptide was predicted at http://bmbpcu36.leeds.ac.uk/prot_analysis/Signal.html. Abbreviations are as follows: Aedae: *Aedes aegypti*; Anoga: *Anopheles gambiae*; Apime: *Apis mellifera*; Apifl: *Apis floreae*; Bomim: *Bombus impatiens*; Camfl: *Camponotus floridanus*; Calro: *Caligus rogercresseyi*; Calsa: *Callinectes sapidus*; Carma: *Carcinus maenus*; Culqu: *Culex quinquefasciatus*; Dappu: *Daphnia pulex*; Jasla: *Jasus lalandii*; Lepsa: *Lepeophtheirus salmonis*; Litva: *Litopenaeus vannamei*; Locmi: *Locusta migratoria*; Megro: *Megachile rotundata*; Nasvi: *Nasonia vitripennis*; Nezvi: *Nezara viridula*; Pedhu: *Pediculus humanus corporis*; Penmo: *Penaeus monodon*; Petci: *Petrolisthes cinctipes*; Rhopr: *Rhodnius prolixus*; Schgr: *Schistocerca gregaria*; Solin: *Solenopsis invicta*; Trica: *Tribolium castaneum*. Aedae OEH: AAD00823.1; Anoga OEH: Q7PRU0; Apime NP: Q1T786; Apifl NP: XP_003696206.1; Bomim NP: XP_003489217.1; Camfl NPA: E1ZXL4; Calro NP: ACO10490.1; Calsa NP: from ESTs CV031136, CV161870, CV223965, CV071064 and C0897289; Carma NPLP: from ESTs DN634570, DN551462, DN635465, DN634851, DV466984 and DV467101; Culqu OEH: XP_001870999.1; Lepsa NPA: D3P110; Litva NP: from ESTs FE056059, FE068047, FE060966 and FE056058; Locmi NPP1: P10776; Locmi NPP2: C0849957.1; Locmi NPP3: from ESTs C0849957, C0821291 and C0832876; Megro NP: XP_003702643.1; Nasvi NP2: NV21878-PA; Pedhu NP: XP_002428081; Penmo NP: from EST GO071574; Petci NP: from ESTs FE812986, FE782448, FE823759, FE799292 and FE766503; Rhopr NP1: D2KEZ9; Rhopr NP2: ACZ96369.1; Rhopr NP3: from EST ACPB01037565; Schgr NPP1: Q966Y1; Schgr NPP2: CAC38870; Schgr NPP4: CAC82521.2; Schgr NPP4: CAC82522.2; Solin NP: SiJWH08BCH2; Trica NP: D6WME2.

Figure 38 shows a phylogenetic tree for the NP prohormones from crustaceans and insects. Although the bootstrap values at the branches are low, it is still interesting to note that the NP prohormones from certain insect orders remain closely linked to other members of that order. The amount of variation that exists in the NP family is very large even within the same order and even species there is significant variability (e.g. locust). The tree represents graphically the amount of variation that exists within this hormone family. When one compares where *J. lalandii* is situated compared to the other crustaceans it is evident that it is separated significantly enough from the rest of the crustaceans. It is interesting to note that Dappu-NP is closely situated to Trica-NP, a result similarly observed in Figure 36 (page Figure 36).

Work done by Claeys et al. (2003) showed that the locust *S. gregaria* expressed 4 different homologs of the NP genes. These authors showed that the desert locust Schgr-NPP3 and Schgr-NPP4 were expressed in neural and non-neural tissues, whereas Schgr-NPP1 and Schgr-NPP2 were expressed exclusively in the brain tissue. To investigate whether this is the case in *N. viridula* and *J. lalandii*, the aim was to identify whether the spatial expression of the NP gene would be similar in *N. viridula* compared to *S. gregaria* and whether this spatial distribution is mimicked in the crustacean *J. lalandii*. To investigate the spatial distribution of the NP gene in *N. viridula* and *J. lalandii*, various tissues were dissected, mRNA was extracted and cDNA was synthesised. PCR amplification was used to detect whether the transcript was being expressed in selected tissues or not. It is important to note that the results are not quantitative, the experiment only shows whether the transcript is expressed or not and does not show the amount of transcript in each tissue. Table 4 (page 92) is a summary of the results obtained from NP PCR amplification in the various *N. viridula* and *J. lalandii* tissues: it is apparent that the NP gene is expressed in a multitude of tissues and not just neural tissues.

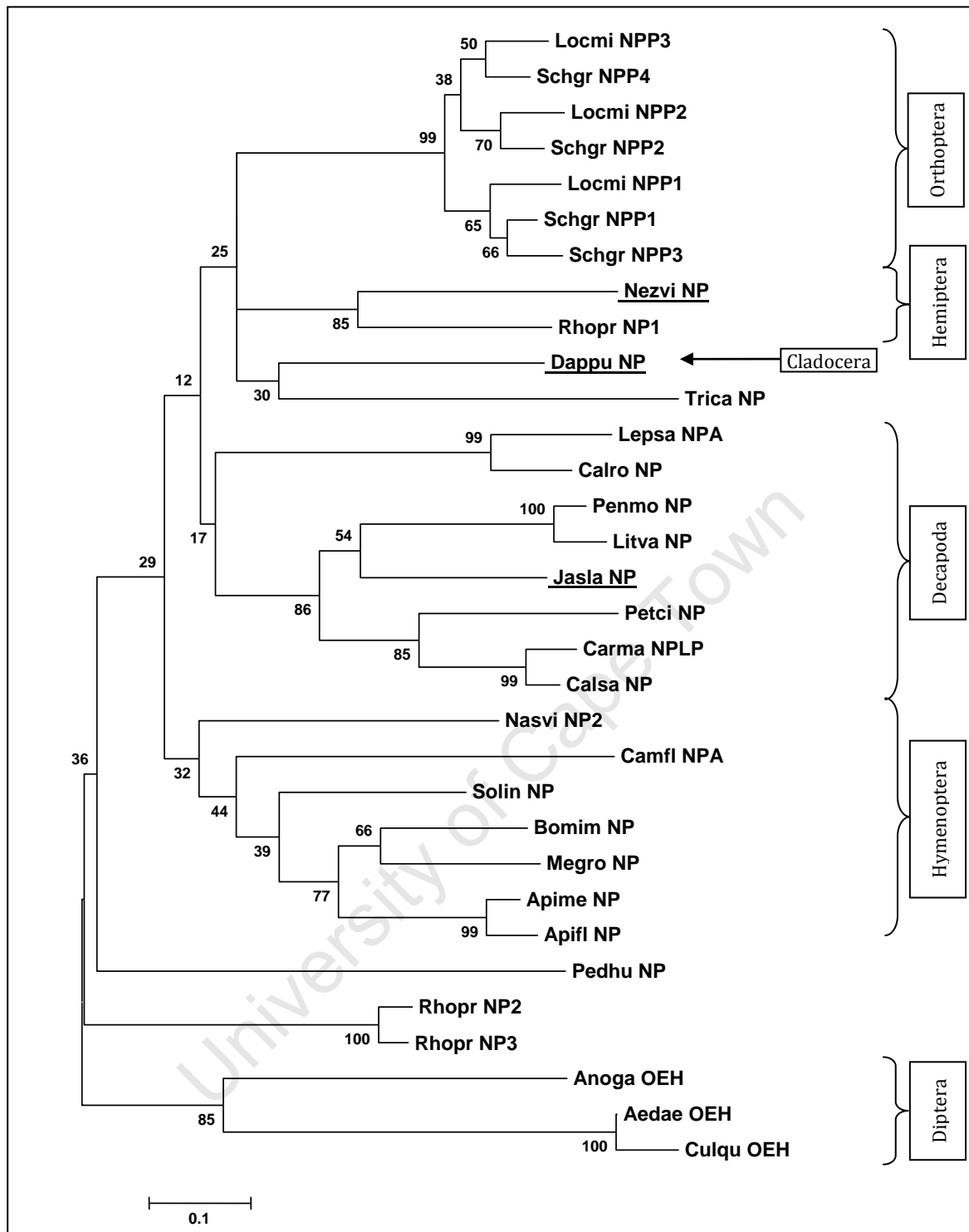


Figure 38: Phylogenetic analysis of the NP gene family in insect and crustaceans. The phylogenetic unrooted tree was constructed using the minimum evolution method. The names of various insect orders are shown on the right-hand side and underlined NP occur in animals used in the study. For node support, a bootstrap analysis was performed as described in materials and methods. The number of bootstrap replicate tested was 5000 and shown next to the branches. The evolutionary distances are given in units of the number of amino acid substitutions per site. For a list of the AKH accession numbers refer to Figure 37, page 89.

Table 4: Expression of NP in various *N. viridula* and *J. lalandii* tissues

<i>N. viridula</i>		Br	SOG	VNC	CC	LM	FM	G ^{MH}	R ^m /R ^f	Ant	FB
Males	Mature	X	N	N	X	X	X	X	X	X	X
	Immature	X	A-	N	X	X	X	N	X	X	A-
Females	Mature	X	X	X	X	X	X	A-	X	X	X
	Immature	X	X	N	X	X	X	X	X	X	X
<i>J. lalandii</i>		Br	SOG	ThG.	XO	Abm	Heart	HepP.	Testis	Ova	Gills
Male and Female		X	X	X	X	N	X	N	N	X	A-

N - No NP transcripts detected in PCR amplification

X - PCR amplified products detected

A- - No amplification of actin transcript

The tissues which did not express the actin gene in *N. viridula* were the SOG and FB in immature males. This is likely due to degradation of the tissue during sample collection or possibly that the amount of starting RNA for cDNA synthesis was way too little for actin transcripts even when the NP transcript is enriched within these tissues. The amplification of the actin transcript for G^{MH} tissue in mature females also consistently produced no amplicons, even with PCR amplification on different cDNA preparations. The actin transcript of gills for *J. lalandii* consistently gave no amplicons, even when different gill tissue and RNA preparations were used. It is unclear why this particular tissue did not amplify the actin transcript.

In the desert locust *S. gregaria*, the transcripts for Schgr-NPP3 and Schgr-NPP4 were detected via Northern analysis in brain, VNC, FB and muscle tissue for both male and female locusts; the NP transcripts were also detected in the male accessory glands and testes, but not in the female ovary nor in gut tissue of either sex (Claeys et al., 2003). In my study, I obtained a similar spatial distribution of the NP mRNA transcript in *N. viridula*, except that NP was detected in both male and female reproductive organs and in the gut tissues (except in immature males). When comparing my result to a phylogenetically more advanced insect group, such as the diptera *D. melanogaster* subgroup, it is evident that the dipteran subgroup has no need for the NP gene, since the gene has been lost entirely in this subgroup (Veenstra, 2010). Therefore, when comparing the absolute loss of NP in the *D. melanogaster* subgroup and the extent that

NP is expressed in locust and in stinkbugs, it was of interest to ascertain whether a comparison of the spatial distribution of NP would be seen in crustaceans. Indeed, there are some commonalities between the less advanced insects phylogenetically and the crustaceans when we compare the spatial expression of NP. The NP gene is expressed in *J. lalandii* neural tissues and, although not detected in the abdominal muscle tissue, it was detected in the heart (cardiac muscle). The possibility also exists that the neurohaemal and pericardial organs were the source of the NP transcript, because these organs are located closely on top of the heart. When comparing the four organisms *J. lalandii*, *N. viridula*, *S. gregaria* and *D. melanogaster* it appears that the spatial distribution of the gene is at least conserved between the phylogenetically less advanced insects (*N. viridula*, *S. gregaria*) and the crustacean (*J. lalandii*). It is also possible that, like locusts, multiple variants of the NP gene exist within *N. viridula*, and that certain variants are expressed exclusively in particular tissues. When compared to the other hemipteran *R. prolixus*, this hypothesis seems likely, since *R. prolixus* contains 3 variants of the NP gene (from shotgun sequence data, Rhopr-NP1: D2KEZ9; Rhopr-NP2: ACZ96369.1; Rhopr-NP3: from EST ACPB01037565).

The function of the NP gene in crustaceans like *J. lalandii* is unknown. Perhaps suggestions made by Veenstra (2010) may shed some light on possible action of the Jasla-NP. The author stated that the loss of the NP gene in *Drosophila melanogaster* and other species of the subgroup *melanogaster* may be linked to the type of development, holometabolous versus hemimetabolous. In larval hemimetabolous insects, such as locust (as well as *N. viridula*) NP has been shown to be necessary for the “gradual progression from a larva into the adult form during development” (Veenstra 2010). In certain holometabolous insects, such as *D. melanogaster*, NP has no relevance anymore. The presence of NP in *J. lalandii* may, thus, also argue for a possible role during development.

4.5 RPCHR homology and localisation

The alignment in figure 39 shows the Dappu-RPCHR, which contains 7 highly conserved transmembrane regions with insect AKHR. The most conserved TM regions are TMII, TMIII and TMVII (homology of roughly 60%) and TMVI (53% homology). The TM regions with the least homology are TMI, TMIV and TMV, which have homology ranging between 30-40%. The regions preceding TMI and proceeding TMVII, are regions with high variability; because they likely interact with either ligand or G-coupled proteins. In *D. pulex* the function of the RPCHR has as not been elucidated yet. It seems unlikely that the main function of the RPCHR would be to signal the concentration of pigment. This is because *D. pulex* does not appear to contain chromatophores. Since the function of Dappu-RPCHR is unknown, possible function may be identified if the target tissue/ organ in which the RPCHR is associating with or synthesised are identified. For example, if the RPCHR was found interacting with fatbody tissue then one could say that it might have a function involved in metabolism. The way we could identify which tissue/ organs the RPCHR is interacting with is to identify where the receptor is expressed in *D. pulex*. Therefore the aim for the work on Dappu-RPCHR was firstly to identify the Dappu-RPCHR sequence, and then to ascertain the spatial distribution of the RPCHR gene. Thus, if the target tissues which express the Dappu-RPCHR can be identified, it might allude to potential functions in that particular tissue.

Whole mount *in situ* hybridisation experiments were conducted to identify the tissues which express the Dappu-RPCHR (Figure 19, page 61). Preliminary results show that the RPCHR transcript was present in tissue that appeared to project outward (horizontally) away from the body. One would be inclined to say that it could be tissue of the thoracopods, because the distribution of the RPCHR probe ran along the length of tissue. The only structure long enough in that area are the thoracopods. Because the animal was damaged in the experiment, there is the chance that the RPCHR probe is interacting with another tissue.

TM I

Dappu RPCHR MCSNDSSSLSGNMCMTERDA-----SFSGESTVNPEFDSGSSISGGSSSTTAVDLSPIDMTFNDGHIVSATYSVLLIISVCGNITVVLNLI-KRR-HIS-NP RVNI
 Peram AKHR -----MALSSCSDQTLN-----CTAVITMATPTTSPSTVL-----MEEYMTDMMKFNDRMSITITYSILMVVSVAGNSTVLITIL-KRR-RTL-RYGNNY
 Blage HTHR -----MTTELP-----REQLTEDMTFGSIHKLCATYCVLMTVSAIGNITVVLNLI-KRR-RNL-RFGNNY
 Grybi AKHR -----MEAS-----AAPPMWVLPNSAD-----GDQLPRDMTFNSGHVVSITACYSAIMVSAAGNITVLT IIL--R-QRGRARS RVNH
 Acypi GnRHR II MEVMSD---DANTVLSVTKG-----VHVGPPVWP-GMDN---VTNASMFDESNLPYDMKFNEGHVVAITVYSILMVVSAIGNITVLT IIL-KRR-RKA-GTR IHA
 Trica AKHR ---MNFSETLWKMKDPMSASET-----VQDHRNLLDWSKTSLDN-----ATEHKLPI SMRFNEGHQLSITIVYSILMVVSAIANTVLT IIL-KRR-RKT-PSRINT
 Drome GnRHR A -----MAKVAEE-----NDHRDLSNWSNVN-----DTNGTIHLTKDMVFNDGHRLSITVYSILFVISTIGNSTVLYLLT-KRRLRG--PLRIDI
 Drome GnRHR B -----MAKVAEE-----NDHRDLSNWSNVN-----DTNGTIHLTKDMVFNDGHRLSITVYSILFVISTIGNSTVLYLLT-KRRLRG--PLRIDI
 Drome GnRHR C -----MAKVAEE-----NDHRDLSNWSNVN-----DTNGTIHLTKDMVFNDGHRLSITVYSILFVISTIGNSTVLYLLT-KRRLRG--PLRIDI
 Glomo AKHR -----MSETEVNGK-----YDHRVLPHWSNVKN-----ETNGSIHYSPIFNAGHRLSITVYSILFVLSLIGNSTVLYLLT-KRRLSSRHTSRIDI
 Anoga AKHR ---MPNTMAAHINQRIE-----DHRNLADWSYANETAGEEY---EMPIDMRFN SGHLSIMVYITLMVFSATGNLTVLSILA-QRKVR--ASSRINI
 Aedae GnRHR ---MSNAILKTERDEVLYNSHSYGENYNDVNTMPYVLSSTSTKTGVFDNETWYG----TNSNWNELPI DMQFNDGHKLOIVVYVLMVISAIGNITVLT IIL-KRRLKS-HS-RIDM
 Danpl AKHR -----MDIDEKVS GPGG-----ASQKNWSHLLHVNT-----YDELPDLMRFN SHMVSMVYVLMVISAIGNITVLYQLV-RRRRAKR-ASRIDI
 Bommo AKHR -----MDIDEKVS GPGG-----ASQKNWT--HLNTS-----YDELPDLMRFN SHMVSMVYVLMVISAIGNITVLYQLV-RRRRAKR-ASRIDI
 Manse AKHR -----MTTA-----LPI DMRFNAGHVSITVYSILMI ISAVGNITVLT IIL-KRRLGAGR-ASRIDI
 Nasvi AKHR -----MESSIK-----IITTTGLENWRVNSN-----YTELPI DMRFN GHIVSITVYVLMVISAIGNITVLT IIL-KRRLGAGR-ASRIDI
 Apime AKHR -----MQ-----LTD--SKESLKND-----TMELPDELRFN SHLVITVYVLMVISAIGNITVLT IIL-KRRLGAGR-ASRIDI
 Polvi AKHR -----MQ-----LTD--SKESLKND-----TMELPDELRFN SHLVITVYVLMVISAIGNITVLT IIL-KRRLGAGR-ASRIDI
 Camfl GnRHR -----MQNAIQ-----LSEVESEESLKND-----TMELPDELRFN SHLVITVYVLMVISAIGNITVLT IIL-KRRLGAGR-ASRIDI

TM II

TM III

TM IV

Dappu RPCHR MLTHLAIADLLVTLMLMPLIEIGWAATVQWRAGDFSCRILAFFRTFGLFLSSFVLCISIDREGAILQPMKL--DYWKRRGRFMLAIWACSVICSLPQVVFHVKAHPEYYPWYEQCVTFD
 Peram AKHR MFMHLAIADLLVTFMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Blage HTHR MFMHLAIADLLVTFMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Grybi AKHR MLMHLAIADLLVTFMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Acypi GnRHR II MLMHLAIADLLVTFMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Trica AKHR MLMHLAIADLLVTFMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Drome GnRHR A MLMHLAIADLMVTLLMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Drome GnRHR B MLMHLAIADLMVTLLMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Drome GnRHR C MLMHLAIADLMVTLLMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Glomo AKHR MLMHLAIADLLVTFMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Anoga AKHR MLTHLAIADLLVTFMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Aedae GnRHR MLTHLAIADLLVTFMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Danpl AKHR ---MHLAVADLMVTFLMMPLEIAGWATVQWRAGDLMCRVMAFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Bommo AKHR LLMHLAVADLMVTFLMMPLEIAGWATVQWRAGDLMCRVMAFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Manse AKHR LLMHLAVADLMVTFLMMPLEIAGWATVQWRAGDLMCRVMAFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Nasvi AKHR MLTHLAIADLLVTFMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Apime AKHR MLMHLAIADLLVTFMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Polvi AKHR LIMHLSIADPLVTFMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN
 Camfl GnRHR LIMHLSIADLLVTFMLMPLIEIANNITVSWKGGDLMCRLMLFFRTFGLFLSSFVLCISLDRCVAILRPMSSKLLNVARRGKMLTVAWILATLCSLPQAVIFHVEPHENVTWYEQCVSFN

TM V

TM VI

Dappu RPCHR SFPTKAHEISYAAFQMMMYVLPLAVFVFTYSSILCEIRORS---KEVQEEGIRRVTVGTIGRARIKTVKMTLVITISVVFIFCWTPYNIMSTIWFWDORDRSALQVDQRIQKGLFLFACTN
 Peram AKHR FFSTKMHFTYRVLGMVMMYGLPLIVIVISYACILGETIRRY---QLS-PDDSFRRSS-LVFLNRAKNRNLTKMAIIIFVVFVFCWTPYYVMCLWYWDIDERSAETVDHRVQKALFLFASTN
 Blage HTHR MFPTQLYQFWRILNMVLVYGFPLLVIFISYACILTEIFRRY---QLS-SDENFRSS-LVFLNRAKNRNLTKMAIIIFVVFVFCWTPYYVMCLWYWDIDERSAETVDHRVQKALFLFASTN
 Grybi AKHR VLKSHRAEVLVYAFISSSSMYGIPLLIIFAYAGSIVAEIFRRS---RRT-GDVFRRSS-LGFLGRAKTRSLKMTLVIVLVVFCWTPYYIMSVWYWFDRSAELVDERIRSGLFIIFACTN
 Acypi GnRHR II AFSSKLEHAYLYFGMFMYYWLPILVILFCYASTIIIEIYRRS---RESICGQGTDNVRLGLIGRAKSRTLKMTIIIVIVFVFCWTPYYIMATWYWDIDERSAELVDERIRSGLFIIFACTN
 Trica AKHR VFPTIYAHLELTYLLFGMVMYALPLAVIIFSYASTILLEIFRRY---RNP-YGDSVTRSS-LAFLGKAKVRTLKMTIIIVLVFVFCWTPYYVMCLWYWDIDERSAKNVDQRIQKALFLFASTN
 Drome GnRHR A SFRSDFDEKLYQAASMCsMYAFPLIMFIYCYGATYLEIYRKSQRVLKDVIAERFRSS-DDVLSRAKKRTLKMTITIVIVFIIICWTPYYTISMWYWDKHSAGKINPLLRKALFIIFASTN
 Drome GnRHR B SFRSDFDEKLYQAASMCsMYAFPLIMFIYCYGATYLEIYRKSQRVLKDVIAERFRSS-DDVLSRAKKRTLKMTITIVIVFIIICWTPYYTISMWYWDKHSAGKINPLLRKALFIIFASTN
 Drome GnRHR C SFRSDFDEKLYQAASMCsMYAFPLIMFIYCYGATYLEIYRKSQRVLKDVIAERFRSS-DDVLSRAKKRTLKMTITIVIVFIIICWTPYYTISMWYWDKHSAGKINPLLRKALFIIFASTN
 Glomo AKHR AL-SQFQSIYFNMTTMCAMYACPLITFIYCYGATYLEIYRKSQRVLKDVIAERFRSS-DDVLSRAKKRTLKMTITIVIVFIIICWTPYYTISMWYWDKHSAGKINPLLRKALFIIFASTN
 Anoga AKHR YFEEIYQIINVLVCMIMYTFPLIVILYCYGSIYYEIFSRT---NPR-NLESFRSS-IDVLSRAKKRTLKMTITIVIVFVFCWTPYYVMSLWYWDKESKTNVDQRIQKGLFLFASTN
 Aedae GnRHR TFSNDNYHTVYINILVMMFMYALPLLTIIICSYASTIYMEIFRHS---RMP-NSEGFRRSS-IDVLSRAKKRTLKMTITIVIVFVFCWTPYYVMSLWYWDKESKTNVDQRIQKGLFLFASTN
 Danpl AKHR SLPTERHEFAYFLVNMALMYVAPLVSTLYCSSAALLEIKRA-----NTANDKMRRSG-VGILGRARARTLKMTVTIVLVFFACWSPPYCYCLWYWDKESVKSLEDPAFQKAMWLFSCFN
 Bommo AKHR SLPTTEKHEFAYFLVNMALMYVAPLVSTLYCSSAALLEIKRA-----NTANDKMRRSG-VGILGRARARTLKMTVTIVLVFFACWSPPYCYCLWYWDKESVKSLEDPAFQKAMWLFSCFN
 Manse AKHR SLPTVHHEFAYFLVNMALMYVAPLVSTLYCSSAALLEIKRA-----NTSNDKMRRSG-VGILGRARARTLKMTVTIVLVFFACWSPPYCYCLWYWDKESVKSLEDPAFQKAMWLFSCFN
 Nasvi AKHR TFPSTHEELTYSLFGMVMYWFPLIVIIYTYTSLAEIYRRS---KDT-TSDRIRRS-LGFLGRAKVRTLKMTIIIVLVFVFCWTPYYVMSLWYWDKESVKSLEDPAFQKAMWLFSCFN
 Apime AKHR AFPPTIYTHEITYSLFGMIMMYWFLPVVIIYTYTSLLEIFRRS---KKS-EDDKIRRS-IGELTRAKVRTLKMTVTIVLVFVFCWTPYYVMSLWYWDKESVKSLEDPAFQKAMWLFSCFN
 Polvi AKHR FFPSHLQETISYNLFSMLIMYWFPLIVIFITYTSSIFMEICRRS---QEK-SEDKIRRS-SGFLSRARVRTLKMTITIVLVFVFCWTPYYVMSLWYWDKESVKSLEDPAFQKAMWLFSCFN
 Camfl GnRHR FFPSHLHEISYTLFGMLMYYWFLFVILYTYSSIFMEICRRS---QEK-NEDKIRRS-SGFLSRARVRTLKMTITIVLVFVFCWTPYYVMSLWYWDKESVKSLEDPAFQKAMWLFSCFN

TM VII

Dappu RPCHR SCFNPMVYGYFNRRTVRRSHH-----ELHRKVYVHPSRMVSRALGPSLRVDCNSKSLGAVTEPCLPQQHLPVATADDPSC-----QLGQAEPIFT
 Peram AKHR SCMNPIVYGYFNRLGRSGYGGATG-----GRVGOQLHHQNVVALSGNSTGLNSRRGNSNSSIYRNNNSQMSWKNLSRSGRRNSRETEHLHPLPHRNSAQAITNNGRDDQHQL
 Blage HTHR SCMNPIVYGYFNFRSGRSGYGGAT-----RPGQQLQHQTALSNSTGVSNSRRGNSNCSSIYRDNNSQMSWN-----RRSSHETE-----MHANNRNDENHL
 Grybi AKHR SCMNPIVYGAFNVRGRG-----RALREYPSSR-CNSQPF SARPRSSAP---EQLQQVAKWR-----NEKNTQLFQ
 Acypi GnRHR II SCMNPIVYGAFNIRTRR-----TLVQATISKSTETRVLDF-----
 Trica AKHR SCMNPIVYGVFNIRARRTGRKVSFRVN---TIKHTSCIPTPNGDSRLPPLEISLKTLE-----
 Drome GnRHR A SCMNPLVYGLYNIIRGRMNNNNPSVN-----NRHTSLNRLDSSNQLMQKQLTNNSLNLRGQVMMAAVSATTKLANVVS LKGTANGNGSAAAAGTVPITPPLTVTIAPLATDDE
 Drome GnRHR B SCMNPLVYGLYNIIRGRMNNNNPSVN-----NRHTSLNRLDSSNQLMQKQLTNNSLNLRGQVMMAAVSATTKLANVVS LKGTANGNGSAAAAGTVPITPPLTVTIAPLATDDE
 Drome GnRHR C SCMNPLVYGLYNIIRGRMNNNNPSVN-----NRHTSLNRLDSSNQLMQKQLTNNSLNLRGQVMMAAVSATTKLANVVS LKGTANGNGSAAAAGTVPITPPLTVTIAPLATDDE
 Glomo AKHR SCMNPIVYGMFNIRGRDNNANSSS-----NRNPSVYQRGDSSSQIPKSLLN-----LSHSESNKIP---PTRTSTNDK-----LVDSPLAKLNK
 Anoga AKHR SCMNPIVYGVFNVRKHKHTKK-----LLK-----TTHEKSCGS-----HLTMA-----
 Aedae GnRHR SCMNPIVYGIYNVKLRKHKKPDGVK-----SGQSSVILRNSAKYTRHSESIRSSSGKIRGYRDLIPSCAIKPV-----
 Danpl AKHR SCANPIVYGVFNRRRWTWRSS-----HTGRCRNG-MRRCRFPFHGDSMELSAATLDAQARQR--SDRHS-----RRDSGFTA
 Bommo AKHR SCANPIVYGVFNRRRWNWRAGK-----FQNGRCRSGSGRKGSRPLPHGESTEISAATLSRARHSNGSDHNG-----RRDSSYAN
 Manse AKHR SCANPIVYGLFNRRRWTWRSG-----HNVRHRSGSMRGRSRLPYGESMELSAATLNRRNSLGS DRNG-----RRDSAFNT
 Nasvi AKHR SCMNPIVYGAFNIRKGNKVTR-----NWDIHTLK-----
 Apime AKHR SCMNPIVYGAFNIRDRNKT-----SARPTTIETRVTPLSLSLKLDD-----
 Polvi AKHR SCMNPIVYGIFNIRKKNKT-----PIRTTTIETRLTPLSLSVKLLD-----
 Camfl GnRHR -----

```

Dappu RPCHR      ASSIAVAVKRRNGIKGLGSLTQSRLQF-----
Peram AKHR       HTNSNKTPNVLEDASNSNSKNALTSVICR-----
Blage HTHR       HPNSAANHNLRTTVSTVSEVPEAR-----
Grybi AKHR       LATNGWDRNGKSSFENIELT-----
Acypi GnRHR II  -----
Trica AKHR       -----
Drome GnRHR A   ANDDSCLSAVTIRCQDQSPIRQK-----
Drome GnRHR B   ANDDSCLSAVTIRCQDQSPIRQK-----
Drome GnRHR C   ANDDSCLSAVTIRCQDQSPIRQKCGDSIELTSVVK
Glomo AKHR      ADT---LPMICIR-----
Anoga AKHR       -----
Aedae GnRHR     -----
Danpl AKHR      KNGSQKH---INNNNVNGIV-----
Bommo AKHR      QNGPQKHWTINNNHVTNGMV-----
Manse AKHR      YNGTQKHWTINNNHVSNGMV-----
Nasvi AKHR      -----
Apime AKHR      -----
Polvi AKHR      -----
Camfl GnRHR     -----

```

Figure 39: Alignment showing the *D. pulex* RPCHR amino acid sequence against insect AKH receptor sequences. The letters in black boxes represent amino acids that are identical, the grey boxes amino acids which are similar in type, trans-membrane segments (TMI - TMVII) are indicated above the alignment. The alignment was generated in BioEdit and gaps (-) have been introduced to maximise the alignment. Abbreviations are as follows: Acypi: *Acyrtosiphon pisum*, Aedae: *Aedes aegypti*; Anoga: *Anopheles gambiae*; Apime: *Apis mellifera* Blage: *Blattella germanica*; Camfl: *Camponotus floridanus*; Dappu: *Daphnia pulex*; Danpl: *Danaus plexippus*, Drome: *Drosophila melanogaster*; Glomo: *Glossina morsitans morsitans*; Grybi: *Gryllus bimaculatus*; Manse: *Manduca sexta* Nasvi: *Nasonia vitripennis*; Peram: *Periplaneta americana* Polvi: *Polyrhachis vicina*; Trica: *Tribolium castaneum*. Acypi GnRHR: XP_003245942.1; Aedae AKHR: CAY77166.1; Anoga AKHR: ABD60146.1; Apime AKHR: NP_001035354.1; Blage HTHR: ADL60118.1; Bommo AKHR: NP_001037049.1; Camfl GnRHR: EFN69256.1; Danpl AKHR: EHJ70223.1; Drome GnRHR A: NP_477387.1; Drome GnRHR B: NP_723206.1; Drome GnRHR C: NP_995639.1; Glomo AKHR: AEH25943; Grybi AKHR: ADZ17179.1; Manse AKHR: ACE00761.1; Nasvi AKHR: NP_001161243; Peram AKHR: AAT00541.1; Polvi AKHR: ADK55068.1; Trica AKHR: ABN79650.1.

5. Future perspectives

The information gathered from my project will be useful for further studies in comparative neuroendocrinology, such as studies involving the comparison of neuropeptides between insects and crustaceans. It will also be useful for studies attempting to design pesticides that are targeted to hinder a specific biological process within the insect pests. By understanding that insects and crustaceans share many common neuropeptides and potentially act on similar biological pathways, this may aid in the design of improved pesticides that do not have broad activity to the crustaceans and closely related insect orders.

The overall result of using the two neuropeptide families AKH/RPCH and NP to investigate homology did not shed any light on whether *D. pulex* clusters closer with the insects or crustaceans. Further research is required within the other crustacean orders as to date neuropeptides such as RPCH have mostly been studied in decapod crustaceans. This will enable better inferences to be made regarding how crustaceans and specifically *D. pulex* is related to the two subphyla.

Future work on localisation studies of RPCH and RPCHR transcripts within *D. pulex* can be verified by doing *in situ* hybridisation on sectioned animal material using RPCH and RPCHR RNA probes so that the results can be clearly analysed. The localisation of NP within *N. viridula* and *J. lalandii* still requires verification from another technique, perhaps also *in situ* hybridisation on section tissue using NP RNA probes.

6. References

- Auerswald, L. & Gäde, G. (2001) Hormonal stimulation of proline synthesis in the fat body of the fruit beetle, *Pachnoda sinuata*, is calcium dependent. *Insect Biochemistry and Molecular Biology*, 32 (1), 23–32.
- Auerswald, L., Schneider, P. & Gäde, G. (1998) Proline powers pre-flight warm-up in the african fruit beetle *Pachnoda sinuata*. *The Journal of Experimental Biology*, 201 (10), 1651–1657.
- Ayali, A. & Pener, M.P. (1992) Density-dependent phase polymorphism affects response to adipokinetic hormone in *Locusta*. *Comparative Biochemistry and Physiology Part A: Physiology*, 101 (3), 549–552.
- Ayali, A. & Pener, M.P. (1995) The relations of adipokinetic response and body lipid content in locusts (*Locusta migratoria migratorioides*) with special reference to phase polymorphism. *Journal of Insect Physiology*, 41 (1), 85–89.
- Ayali, A., Pener, M.P. & Girardie, J. (1996) Comparative study of neuropeptides from the corpora cardiaca of solitary and gregarious *Locusta*. *Archives of Insect Biochemistry and Physiology*, 450, 439–450.
- Badisco, L., Claeys, I., Van Hiel, M., Clynen, E., Huybrechts, J., Vandersmissen, T., Van Soest, S., Vanden Bosch, L., Simonet, G. & Vanden Broeck, J. (2008) Purification and characterization of an insulin-related peptide in the desert locust, *Schistocerca gregaria*: immunolocalization, cDNA cloning, transcript profiling and interaction with neuroparsin. *Journal of Molecular Endocrinology*, 40 (3), 137–150.
- Badisco, L., Claeys, I., Van Loy, T., Van Hiel, M., Franssens, V., Simonet, G. & Vanden Broeck, J. (2007) Neuroparsins, a family of conserved arthropod neuropeptides. *General and Comparative Endocrinology*, 153, 64–71.
- Bouaïchi, A., Roessingh, P. & Simpson, S.J. (1995) An analysis of the behavioural effects of crowding and re-isolation on solitary-reared adult desert locusts (*Schistocerca gregaria*) and their offspring. *Physiological Entomology*, 20 (3), 199–208.
- Bourême, D., Fournier, B., Matz, G. & Girardie, J. (1989) Immunological and functional cross-reactivities between locust neuroparsins and proteins from cockroach corpora cardiaca. *Journal of Insect Physiology*, 35 (4), 265–271.
- Breugelmans, B., Simonet, G., van Hoef, V., Van Soest, S. & Vanden Broeck, J. (2009) Pacifastin-related peptides: structural and functional characteristics of a family of serine peptidase inhibitors. *Peptides*, 30 (3), 622–632.
- Brown, M.R. & Cao, C. (2001) Distribution of ovary ecdysteroidogenic hormone I in the nervous system and gut of mosquitoes. *Journal of insect science*, 1, 3.
- Brown, M.R., Graf, R., Swiderek, K.M., Fendley, D., Stracker, T.H., Champagne, D.E. & Lea, A.O. (1998) Identification of a steroidogenic neurohormone in female mosquitoes. *The Journal of Biological Chemistry*, 273 (7), 3967–3971.
- Candy, D.J. (2002) Adipokinetic hormones concentrations in the haemolymph of *Schistocerca gregaria*, measured by radioimmunoassay. *Insect Biochemistry and Molecular Biology*, 32 (11), 1361–1367.
- Carlisle, D.B. & Knowles, F. (1959) *Endocrine control in crustaceans*. 1st ed. New York, Cambridge University Press: London.
- Chaty, S., Rodius, F. & Vasseur, P. (2004) A comparative study of the expression of CYP1A and CYP4 genes in aquatic invertebrate (freshwater mussel, *Unio tumidus*)

- and vertebrate (rainbow trout, *Oncorhynchus mykiss*). *Aquatic toxicology (Amsterdam, Netherlands)*, 69 (1), 81–94.
- Cheeseman, P. & Goldsworthy, G.J. (1979) The release of adipokinetic hormone during flight and starvation in *Locusta*. *General and Comparative Endocrinology*, 37 (1), 35–43.
- Cheung, W.Y., Hubert, N. & Landry, B.S. (1993) A simple and rapid DNA microextraction method for plant, animal, and insect suitable for RAPD and Other PCR analyses. *PCR Methods and Applications*, 69–70.
- Christie, A.E., Cashman, C.R., Brennan, H.R., Ma, M., Sousa, G.L., Li, L., Stemmler, E.A & Dickinson, P.S. (2008) Identification of putative crustacean neuropeptides using in silico analyses of publicly accessible expressed sequence tags. *General and Comparative Endocrinology*, 156 (2), 246–264.
- Claeys, I., Breugelmans, B., Simonet, G., Franssens, V., Van Soest, S. & Broeck, J.V. (2006) Regulation of *Schistocerca gregaria* neuroparsin transcript levels by juvenile hormone and 20-hydroxyecdysone. *Archives of Insect Biochemistry and Physiology*, 62 (3), 107–115.
- Claeys, I., Simonet, G., Van Loy, T., De Loof, A. & Vanden Broeck, J. (2003) cDNA cloning and transcript distribution of two novel members of the neuroparsin family in the desert locust, *Schistocerca gregaria*. *Insect Molecular Biology*, 12 (5), 473–481.
- Claeys, I., Simonet, G., Poels, J., Van Loy, T., Vercammen, L., De Loof, A. & Vanden Broeck, J. (2002) Insulin-related peptides and their conserved signal transduction pathway. *Peptides*, 23 (4), 807–816.
- Coast, G.M., Orchard, I., Phillips, J.E. & Schooley, D.A. (2002) Insect diuretic and antidiuretic hormones. In: Academic Press, 279–409.
- Colbourne, J.K., Pfrender, M.E., Gilbert, D., Thomas, W.K., Tucker, A., Oakley, Todd H., Tokishita, S., Aerts, A., Arnold, G.J. Basu, M.K., Bauer, D.J., Cáceres, C.E., Carmel, L., Casola, C., Choi, J. H., Detter, J.C. Dong, Q., Dusheyko, S., Eads, B.D., Fröhlich, T., Geiler-Samerotte, K.A., Gerlach, D., Hatcher, P., Jogdeo, S., Krijgsveld, J., Kriventseva, E.V., Kültz, D., Laforsch, C. Lindquist, E., Lopez, J., Manak, J.R., Muller, J., Pangilinan, J., Patwardhan, R.P., Pitluck, S., Pritham, E.J., Rechtsteiner, A., Rho, M., Rogozin, I.B., Sakarya, O., Salamov, A., Schaack, S., Shapiro, H., Shiga, Y., Skalitzyk, C., Smith, Z., Souvorov, A., Sung, W., Tang, Z., Tsuchiya, D., Tu, H., Vos, H., Wang, M., Wolf, Y.I., Yamagata, H., Yamada, T., Ye, Y., Shaw, J.R., Andrews, J., Crease, T.J., Tang, H., Lucas, S.M., Robertson, H.M., Bork, P., Koonin, E.V., Zdobnov, E.M., Grigoriev, I.V., Lynch, M., Boore, J.L. (2011) The ecoresponsive genome of *Daphnia pulex*. *Science (New York, N.Y.)*, 331 (6017), 555–561.
- Collett, M., Despland, E., Simpson, S.J. & Krakauer, D.C. (1998) Spatial scales of desert locust gregarization. *Proceedings of the National Academy of Sciences of the United States of America*, 95 (22), 13052–13055.
- Collip, J.B. (1923) The demonstration of an insulin-like substance in the tissue of clam (*Mya arenaria*). *Journal of Biology Chemistry*, 55, 39.
- Cooke, I.M. & Sullivan, R.E. (1982) Hormones and neurosecretion. In: H. Atwood & D. Sandem eds. *The Biology of Crustacea*. New York, Academic Press, 205–290.
- Devon, R.S., Porteous, D.J. & Brookes, A.J. (1995) Splinkerettes-improved vectorettes for greater efficiency in PCR walking. *Nucleic Acids Research*, 23 (9), 1644–1645.
- Dirksen, H., Neupert, S., Predel, R., Verleyen, P., Huybrechts, J., Strauss, J., Hauser, F., Stafflinger, E., Schneider, M., Pauwels, K., Schoofs, L. & Grimmelikhuijzen, C.J.P. (2011) Genomics, transcriptomics, and peptidomics of *Daphnia pulex*

- neuropeptides and protein hormones. *Journal of Proteome Research*, 10 (10), 4478–4504.
- Fernlund, P. & Josefsson, L. (1972) Crustacean Color-Change Hormone: Amino Acid Sequence and Chemical Synthesis. *Science*, 177 (4044), 173–175.
- Fournier, B., Herault, J.P. & Proux, J. (1987) Antidiuretic factor from the nervous corpora cardiaca of the migratory locust: Improvement of an existing in vitro bioassay. *General and Comparative Endocrinology*, 68 (1), 49–56.
- Girardie, J., Bourême, D., Couillaud, F., Tamarelle, M. & Girardie, A. (1987) Anti-juvenile effect of neuroparsin A, a neuroprotein isolated from locust corpora cardiaca. *Insect Biochemistry*, 17 (7), 977–983.
- Girardie, J., Girardie, A., Huet, J.C. & Pernollet, J.C. (1989) Amino acid sequence of locust neuroparsins. *FEBS Letters*, 245, 4–8.
- Girardie, J., Huet, J.C., Atay-Kadiri, Z., Ettaouil, S., Delbecque, J.P., Fournier, B., Pernollet, J.C. & Girardie, A. (1998) Isolation, sequence determination, physical and physiological characterization of the neuroparsins and ovary maturing parsins of *Schistocerca gregaria*. *Insect Biochemistry and Molecular Biology*, 28 (9), 641–650.
- Girardie, J., Richard, O., Huet, J.C., Nespoulous, C., Dorselaer, A. & Pernollet, J.C. (1991) Physical characterization and sequence identification of the ovary maturing parsin. A new neurohormone purified from the nervous corpora cardiaca of the African locust (*Locusta migratoria migratorioides*). *European Journal of Biochemistry*, 202 (3), 1121–1126.
- Gäde, G. (2009) Peptides of the adipokinetic hormone/red pigment-concentrating hormone family: a new take on biodiversity. *Annals of the New York Academy of Sciences*, 1163, 125–136.
- Gäde, G. (2004) Regulation of intermediary metabolism and water balance of insects by neuropeptides. *Annual Review of Entomology*, 49, 93–113.
- Gäde, G. (1997) The explosion of structural information on insect neuropeptides. In: F. L. Tallahassee, G. W. Kirby, R. E. Moore, W. Steglich, & C. Tamm eds. *Progress in the Chemistry of Organic Natural Products*. New York, Springer, 1–128.
- Gäde, G. (1992) The hormonal integration of insect flight metabolism. *Zoologische Jahrbücher. Abteilung für allgemeine Zoologie und Physiologie der Tiere*, 96 (2), 211–225.
- Gäde, G. & Auerswald, L. (2002) Beetles' choice — proline for energy output: control by AKHs. *Physiology*, 132, 117–129.
- Gäde, G. & Auerswald, L. (1998) Insect neuropeptides regulating substrate mobilisation. *South African Journal of Zoology*, 33 (2), 65–70.
- Gäde, G. & Auerswald, L. (2003) Mode of action of neuropeptides from the adipokinetic hormone family. *General and Comparative Endocrinology*, 132 (1), 10–20.
- Gäde, G., Auerswald, L., Simek, P., Marco, H.G. & Kodrík, D. (2003) Red pigment-concentrating hormone is not limited to crustaceans. *Biochemical and Biophysical Research Communications*, 309 (4), 967–973.
- Gäde, G. & Holwerda, D.A. (1976) Involvement of adenosine 3':5'-cyclic monophosphate in lipid mobilization in *Locusta migratoria*. *Insect Biochemistry*, 6 (5), 535–540.
- Gäde, G. & Marco, H.G. (2009) Peptides of the adipokinetic hormone/red pigment-concentrating hormone family with special emphasis on Caelifera: primary sequences and functional considerations contrasting grasshoppers and locusts. *General and Comparative Endocrinology*, 162 (1), 59–68.

- Gäde, G. & Marco, H.G. (2006) Structure, function and mode of action of select arthropod neuropeptides. In: *Studies in Natural Products Chemistry ed. Bioactive Natural Products*, 33, 169–139.
- Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95–98.
- Hergenbahn, H.G., Aspán, A. & Söderhäll, K. (1987) Purification and characterization of a high-Mr proteinase inhibitor of pro-phenol oxidase activation from crayfish plasma. *The Biochemical Journal*, 248 (1), 223–228.
- Hietter, H., Dorsselaer, A. & Luu, B. (1991) Characterization of three structurally-related 8–9 kDa monomeric peptides present in the Corpora cardiaca of *Locusta*: A revised structure for the neuroparsins. *Insect Biochemistry*, 21 (3), 259–264.
- Van der Horst, D.J. & Rodenburg, K.W. (2010) Locust flight activity as a model for hormonal regulation of lipid mobilization and transport. *Journal of Insect Physiology*, 56 (8), 844–853.
- Innis, M.A. & Gelfand, D.H. (1990) Optimization of PCRs. In: M. A. Innis, D. H. Gelfand, J. J. Sninsky, & T. J. White eds. *PCR protocols: a guide to methods and application*. New York, Academic Press, 3–12.
- Ishizaki, H. & Suzuki, A. (1994) The brain secretory peptides that control moulting and metamorphosis of the silkworm, *Bombyx mori*. *International Journal of Developmental Biology*, 38, 301–310.
- Janssen, T., Claeys, I., Simonet, G., De Loof, A., Girardie, J. & Vanden Broeck, J. (2001) cDNA cloning and transcript distribution of two different neuroparsin precursors in the desert locust, *Schistocerca gregaria*. *Insect Molecular Biology*, 10 (2), 183–189.
- Kaufmann, C., Merzendorfer, H. & Gäde, G. (2009) The adipokinetic hormone system in Culicinae (Diptera: Culicidae): molecular identification and characterization of two adipokinetic hormone (AKH) precursors from *Aedes aegypti* and *Culex pipiens* and two putative AKH receptor variants from *A. aegypti*. *Insect Biochemistry and Molecular Biology*, 39 (11), 770–781.
- Kavar, T., Pavlovčič, P., Sušnik, S., Meglič, V. & Virant-Doberlet, M. (2007) Genetic differentiation of geographically separated populations of the southern green stink bug *Nezara viridula* (Hemiptera: Pentatomidae). *Bulletin of Entomological Research*, 96 (2), 117–128.
- Kodrík, D. (2008) Adipokinetic hormone functions that are not associated with insect flight. *Physiological Entomology*, 33 (3), 171–180.
- Kodrík, D., Marco, H.G., Simek, P., Socha, R., Štys, P. & Gäde, G. (2010) The adipokinetic hormones of Heteroptera: a Comparative study. *Physiological Entomology*, 35 (2), 117–127.
- Lampert, W. (2006) *Daphnia*: model herbivore, predator and prey. *Polish Journal of Ecology*, 54, pp.607–620.
- Lea, A.O. (1972) Regulation of egg maturation in the mosquito by the neurosecretory system: the role of the corpus cardiacum. *General and Comparative Endocrinology*, 3, 602–608.
- Lea, A.O. (1967) The medial neurosecretory cells and egg maturation in mosquitoes. *Journal of Insect Physiology*, 13 (3), 419–429.
- Liang, Z., Sottrup-Jensen, L., Aspán, a, Hall, M. & Söderhäll, K. (1997) Pacifastin, a novel 155-kDa heterodimeric proteinase inhibitor containing a unique transferrin chain.

- Proceedings of the National Academy of Sciences of the United States of America*, 94 (13), 6682–6687.
- Lindemans, M., Liu, F., Janssen, T., Husson, S.J., Mertens, I., Gäde, G. & Schoofs, L. (2009) Adipokinetic hormone signaling through the gonadotropin-releasing hormone receptor modulates egg-laying in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America*, 106 (5), 1642–1647.
- De Loof, A., Claeys, I., Simonet, G., Verleyen, P., Vandersmissen, T., Sas, F. & Huybrechts, J. (2006) Molecular markers of phase transition in locusts. *Insect Science*, 13 (1), 3–12.
- Ma, Z., Yu, J. & Kang, L. (2006) LocustDB: a relational database for the transcriptome and biology of the migratory locust (*Locusta migratoria*). *BMC Genomics*, 7, 11.
- Marco, H.G. & Gäde, G. (2010) Biological activity of the predicted red pigment-concentrating hormone of *Daphnia pulex* in a crustacean and an insect. *General and Comparative Endocrinology*, 166 (1), 104–110.
- Marco, H.G., Hansen, I.A., Scheller, K. & Gäde, G. (2003) Molecular cloning and localization of a cDNA encoding a crustacean hyperglycemic hormone from the South African spiny lobster, *Jasus lalandii*. *Peptides*, 24 (6), 845–851.
- Matsumoto, S., Brown, M.R., Suzuki, A. & Lea, A.O. (1989) Isolation and characterization of ovarian ecdysteroidogenic hormones from the mosquito, *Aedes aegypti*. *Insect Biochemistry*, 19 (7), 651–656.
- Michitsch, J. & Steele, J.E. (2008) Carbohydrate and lipid metabolism in cockroach (*Periplaneta americana*) fat body are both activated by low and similar concentrations of Peram-AKH II. *Peptides*, 29 (2), 226–234.
- Moreau, R., Gourdoux, L. & Girardie, J. (1988) Neuroparsin: A new energetic neurohormone in the African locust. *Archives of Insect Biochemistry and Physiology*, 8 (2), 135–145.
- Nagasawa, H., Kataoka, H., Isogai, A., Tamura, S., Suzuki, A., Ishizaki, H., Mizoguchi, A. & Fujiwara, Y. (1984) Amino-terminal amino Acid sequence of the silkworm prothoracicotropic hormone: homology with insulin. *Science (New York, N.Y.)*, 226 (4680), 1344–1345.
- Nässel, D.R., Passier, P.C.C.M., Elekes, K., Dircksen, H., Vullings, H.G.B. & Cantera, R. (1995) Evidence that *Locusta* tachykinin I is involved in release of adipokinetic hormone from locust corpora cardiaca. *Regulatory Peptides*, 57 (3), 297–310.
- Park, Y., Kim, Y.J. & Adams, M.E. (2002) Identification of G protein-coupled receptors for *Drosophila* PRXamide peptides, CCAP, corazonin, and AKH supports a theory of ligand-receptor coevolution. *Proceedings of the National Academy of Sciences of the United States of America*, 99 (17), 11423–11428.
- Pener, M.P. & Yerushalmi, Y. (1998) The physiology of locust phase polymorphism: an update. *Journal of Insect Physiology*, 44, 365–377.
- Predel, R., Neupert, S., Huetteroth, W., Kahnt, J., Waidelich, D. & Roth, S. (2012) Peptidomics-based phylogeny and biogeography of *Mantophasmatodea* (*Hexapoda*). *Systematic Biology*, 61 (4), pp.609–29.
- Rao, K.R. (2001) Crustacean Pigmentary-Effector Hormones: Chemistry and Functions of RPCH, PDH, and Related Peptides. *Integrative and Comparative Biology*, 41 (3), 364–379.
- Roth, S., Fromm, B., Gäde, G. & Predel, R. (2009) A proteomic approach for studying insect phylogeny: CAPA peptides of ancient insect taxa (*Dictyoptera*, *Blattoptera*) as a test case. *BMC Evolutionary Biology*, 9, p.50.

- Richard, O., Tamarelle, M., Girardie, J. & Geoffre, S. (1994) Restricted occurrence of *Locusta migratoria* ovary maturing parsin in the brain-corpora cardiaca complex of various insect species. *Histochemistry*, 102 (3), 233–239.
- Sherff, C.M. & Mulloney, B. (1991) Red pigment concentrating hormone is a modulator of the crayfish swimmeret system. *The Journal of Experimental Biology*, 155, 21–35.
- Showler, A.T. & Potter, C.S. (1991) Synopsis of the 1986-1989 Desert Locust (Orthoptera: Acrididae) Plague and the Concept of Strategic Control. *American Entomologist*, 37 (2), 106–110.
- Spencer, I.M. & Candy, D.J. (1976) Hormonal control of diacyl glycerol mobilization from fatbody of the desert locust, *Schistocerca gregaria*. *Insect Biochemistry*, 6 (3), 289–296.
- Staubli, F., Jorgensen, T.J.D., Cazzamali, G., Williamson, M., Lenz, C., Sondergaard, L., Roepstorff, P. & Grimmlikhuijzen, C.J.P. (2002) Molecular identification of the insect adipokinetic hormone receptors. *Proceedings of the National Academy of Sciences of the United States of America*, 99 (6), 3446–3451.
- Stoffel, W. & Hofmann, K. (1993) TMbase - A database of membrane spanning proteins segments. *Biological Chemistry Hoppe-Seyler*, 376, 166.
- Stone, J.V., Mordue, W., Batley, K.E. & Morris, H.R. (1976) Structure of locust adipokinetic hormone, a neurohormone that regulates lipid utilisation during flight. *Nature*, 263 (5574), 207–211.
- Tamarelle, M. & Girardie, J. (1989) Immunohistochemical investigation of locust neuroparsin-like substances in several insects, in some other invertebrates, and vertebrates. *Histochemistry*, 91 (5), 431–435.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*.
- Vanden Broeck, J. (2001) Neuropeptides and their precursors in the fruitfly, *Drosophila melanogaster*. *Peptides*, 22 (2), 241–254.
- Vanhems, E., Delbos, M. & Girardie, J. (1990) Insulin and Neuroparsin Promote Neurite Outgrowth in Cultured Locust CNS. *European Journal of Neuroscience*, 2 (9), 776–782.
- Veenstra, J.A (2010) What the loss of the hormone neuroparsin in the melanogaster subgroup of *Drosophila* can tell us about its function. *Insect Biochemistry and Molecular Biology*, 40 (4), 354–361.
- Vullings, H.G.B., Ten Voorde, S.E.C.G., Passier, P.C.C.M., Diederer, J.H.B., Van der Horst, D.J. & Nässel, D.R. (1998) A possible role of Schisto FLRFamide in inhibition of adipokinetic hormone release from locust corpora cardiaca. *Journal of Neurocytology*, 27 (12), 901–913.
- Wicher, D., Agricola, H.J., Söhler, S., Gundel, M., Heinemann, S.H., Wollweber, L., Stengl, M. & Derst, C. (2006) Differential receptor activation by cockroach adipokinetic hormones produces differential effects on ion currents, neuronal activity, and locomotion. *Journal of Neurophysiology*, 95 (4), 2314–2325.
- Ziegler, R., Isoe, J., Moore, W., Riehle, M.A & Wells, M.A (2011) The putative AKH receptor of the tobacco hornworm, *Manduca sexta*, and its expression. *Journal of Insect Science*, 11 (40), 40.
- Zralá, J., Kodrík, D., Zahradníková, H., Zemek, R. & Socha, R. (2010) A novel function of red pigment-concentrating hormone in crustaceans: *Porcellio scaber* (Isopoda) as a model species. *General and Comparative Endocrinology*, 166 (2), 330–336.