Investigation of metabolic regulatory genes and hormones in pest snails

Kevin John Adamson

A thesis submitted for the degree of Doctor of Philosophy at The University of the Sunshine Coast in August 2015
School of Science and Engineering
Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the General Award Rules of The University of the Sunshine Coast, immediately made available for research and study in accordance with the Copyright Act 1968.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material.
Statement of Contributions to Jointly Authored Works Contained in the Thesis

1. Kevin J Adamson, Tianfang Wang, Min Zhao, Francesca Bell, Anna V Kuballa, Kenneth B Storey, Scott F Cummins 2015. Molecular insights into land snail neuropeptides through transcriptome and comparative gene analysis. BMC Genomics 16:308

KJA carried out the bioinformatics and experimental analysis, constructed figures, tables and drafted the manuscript. MZ helped to perform the comparative transcriptome analysis. FB maintained animals and prepared tissue for experimental analysis. TW carried out mass spectral proteome work. AVK, KBS and SFC conceived the idea and obtained funding for the experiments and drafted the manuscript. All authors read and approved the final manuscript.


KJA carried out the experimental analysis, constructed figures, tables and drafted the manuscript. BR prepared immunocytochemical tissues. TW carried out mass spectral proteome work. TK helped to perform the histological reconstruction of the animal CNS. AVK, KBS and SFC conceived the idea and obtained funding for the experiments and drafted the manuscript. All authors read and approved the final manuscript.


KJA carried out the experimental analysis, constructed figures, tables and drafted the manuscript. BR analysed immunocytochemical tissues. TW carried out mass spectral proteome work. AVK helped with qPCR and to draft the final manuscript. KBS and SFC conceived the idea and obtained funding for the experiments and drafted the manuscript. All authors read and approved the final manuscript.
Statement of Contributions by Others to the Thesis as a Whole

Min Zhao helped to perform the comparative transcriptome analysis in Chapter 2. Francesca Bell assisted with maintenance of animals and tissue preparation for experimental analysis in Chapter 2. Tianfang Wang carried out mass spectral peptide identification in Chapters 2 and 3. Bronwyn Rotgans prepared immunocytochemical tissues in Chapters 2 and 3. Thanopong Kruangkum helped to perform the histological reconstruction of the animal in Chapter 3 and assisted with in situ hybridisation. Anna Kuballa assisted with qPCR in Chapters 3 and 4, and with Kenneth Storey and Scott Cummins, the primary supervisor, conceived the idea and obtained funding for the experiments and drafted the manuscripts.

Statement of Parts of the Thesis Submitted to Qualify for the Award of Another Degree

None.

Published Works by the Author Incorporated into the Thesis

1. Kevin J Adamson, Tianfang Wang, Min Zhao, Francesca Bell, Anna V Kuballa, Kenneth B Storey, Scott F Cummins 2015. Molecular insights into land snail neuropeptides through transcriptome and comparative gene analysis. BMC Genomics 16:308
Incorporated as Chapter 2.

Incorporated as Chapter 3.

Incorporated as Chapter 4.
Additional Published Works by the Author Relevant to the Thesis but not Forming Part of it


This Thesis is dedicated to the hundreds of snails who donated their bodies to science.
Acknowledgements

This Ph.D would never have happened without the help of a great number of people. You are too numerous to name, but if you ever read this, you know who you are, and thanks. Included in the list are my immediate and extended family and friends, who have encouraged and supported me along the way. Without family and friends, nothing is worth doing. A large number of fellow Honours and Ph.D students and staff at USC have assisted me, with tech support, assistance in the lab, shoulders to cry on at times, but mostly just by being there and being friends who I’ll value forever.

I would not have got here without the help of my amazing supervisors, Scott Cummins and Anna Kuballa. I’m well aware of the difficulties of having a grumpy old man as a student, but you’ve survived it! Well done and thanks. Two other people at USC deserve special mention. Lesley Brooker and Abigail Elizur, it has been a pleasure knowing you both, and thanks for having faith in me. There is no doubt in my mind that I would never have started my Ph.D without your help and encouragement both before and during Honours.

Finally, thanks to USC for being a university that I enjoyed attending, to the Australian Research Council for helping fund the project, and to the Grain Research and Development Corporation for their generous financial support that removed some of the financial pressures that can so often hamper research.
Abstract

Aestivation is a state of hypometabolism used by a variety of animals to survive conditions that are too hot or dry to sustain life. Many physiological and molecular changes take place during this dormancy, but little is known of the global control necessary to initiate and maintain these changes. An understanding of this control in the land snail *Theba pisana* could help with the development of a pest control method for this invasive species, and could also enable a greater understanding of metabolic control in other species including humans, with possible medical implications.

Neuropeptides control or influence various complex metabolic events, so this study initially set out to identify the entire set of neuropeptides in *T. pisana*, some of which may be involved in the control of aestivation. *In silico* gene data mining of three tissue transcriptomes found a total of 24,920 central nervous system (CNS), 37,661 foot muscle and 40,766 hepatopancreas gene transcripts, which together encode for 5,236 functional protein domains. We predict that as many as 5,504 genes encode for proteins that may be destined for extracellular secretion. Of these, we identified 35 full-length neuropeptide genes encoding precursors that release molluscan-type bioactive neuropeptide products, and are abundantly expressed within *T. pisana* CNS. These included achatin, allototropin, conopressin, elevenin, FMRFamide, LFRFamide, LRFNVamide, myomodulins, neurokinin Y, PKYMDT, PXFVamide, sCAPamides and several insulin-like peptides. The presence of many of these neuropeptides was confirmed using liquid chromatography-mass spectrometry of neural ganglia.

Comparison of the CNS neuropeptidome in active and aestivating states using mass spectral-based analysis, initially revealed 19 differentially produced peptides; 2 were more abundant in active animals and 17 were more abundant in aestivated animals. Using a different extraction method we found a total of 22 peptides differentially produced; 4 upregulated in active animals and 18 upregulated in aestivation. Two peptides upregulated in aestivation, buccalin and sCAP, are known to regulate muscle contractions in a variety of molluscs and for that reason were chosen for further investigation to analyse temporal and spatial expression of their precursor genes and bioactive peptides. One peptide upregulated in active animals, an unknown aestivation-associated peptide (AAP12), was also chosen for further analysis.
We found no significant difference in CNS gene expression levels between active and aestivated animals for either buccalin or sCAP, suggesting that regulation may reside at the level of post-translational control of peptide generation. Spatial gene and peptide expression analysis of aestivated snail CNS confirmed that buccalin and sCAP have widespread distribution within neural regions that control several physiological roles. CNS gene expression levels for the AAP12 precursor gene were upregulated in active animals, while the peptide was upregulated in aestivated animals.

This study has provided the most comprehensive list of genes and peptides present within the CNS of a land snail. It is also the first detailed molecular analysis of the peptides associated with hypometabolism in a gastropod snail, providing new insight into the molecular basis of aestivation through CNS peptide control.

**Keywords**
Snail, *Theba pisana*, peptides, neuropeptides, aestivation, central nervous system, muscle, hepatopancreas
# TABLE OF CONTENTS

Declaration by author.........................................................................................................................2  
Statement of Contributions to Jointly Authored Works Contained in the Thesis.................................4  
Statement of Contributions by Others to the Thesis as a Whole .....................................................6  
Statement of Parts of the Thesis Submitted to Qualify for the Award of Another Degree ...............6  
Published Works by the Author Incorporated into the Thesis..........................................................6  
Additional Published Works by the Author Relevant to the Thesis but not Forming Part of it ...........8  
Acknowledgements............................................................................................................................12  
Abstract.............................................................................................................................................14  
LIST OF FIGURES...............................................................................................................................20  
LIST OF TABLES.................................................................................................................................22  
LIST OF ABBREVIATIONS..................................................................................................................24  
Chapter one: Introduction ..................................................................................................................26  
Hibernation, torpor, diapause and aestivation....................................................................................27  
Primary molecular factors that are modified during hypometabolism..............................................35  
  Protein stability..................................................................................................................................35  
  Anti-oxidants.................................................................................................................................35  
  Gene expression.............................................................................................................................36  
  Neuropeptides and metabolism........................................................................................................37  
  Implications of neuropeptide research in land snails......................................................................41  
  Conclusions......................................................................................................................................44  
  Aims of this thesis............................................................................................................................44  
Chapter 2: Molecular insights into land snail neuropeptides through transcriptome and  
comparative gene analysis ..................................................................................................................48  
  Background......................................................................................................................................49  
  Results.............................................................................................................................................50  
  De novo assembly and comparison of Theba CNS, foot and hepatopancreas...............................50  
  Functional annotation......................................................................................................................52  
  Annotation of mollusc proteins secreted from CNS in Theba pisana...........................................54  
  Prohormone convertases................................................................................................................59  
  Discussion.........................................................................................................................................60  
  Conclusion.......................................................................................................................................64  
  Materials and Methods....................................................................................................................64  
  Animals and tissue collection..........................................................................................................64  
  RNA isolation and transcriptome sequencing....................................................................................65  
  Gene ontology and identification....................................................................................................65
Protein comparison and Pfam domain annotation .................................66
Neuropeptide prediction and sequence analysis ...................................67
Tissue distribution of selected genes of interest .....................................67
Peptide isolation from CNS and nanoHPLC, mass spectrometry LC-ESI-QTOF peptide identification ..........................................................68

Chapter 3: Genes and associated peptides involved with aestivation in a land snail ........80
Abstract .................................................................................................80
Background ............................................................................................81
Materials and Methods ..........................................................................83
Animals and tissue collection for proteome analysis ...............................83
Histology .................................................................................................83
Protein isolation from CNS and nanoHPLC - ESI-Triple TOF peptide identification and label-free quantification ...........................................84
Gene analysis .........................................................................................85
Absolute qPCR .......................................................................................86
In situ hybridization ................................................................................87
Immunolocalisation .................................................................................87
Results and Discussion .........................................................................88
Histology of the Theba pisana CNS........................................................88
Identification of differentially expressed peptides in active and aestivated Theba pisana CNS .................................................................89
Temporal and spatial expression of Tpi-buccalin-2 in the CNS ...............97
Conclusions ............................................................................................101

Chapter 4: Differential peptide expression in the central nervous system of the land snail
Theba pisana, between active and aestivated ..........................................110
Abstract .................................................................................................110
Background ............................................................................................111
Materials and methods ..........................................................................113
Animals and tissue collection for proteome analysis ...............................113
Protein isolation from CNS and nanoHPLC, mass spectrometry LC-ESI-QTOF peptide identification ..........................................................113
Gene analysis .........................................................................................115
Absolute qPCR .......................................................................................115
Immunolocalisation .................................................................................116
Results ....................................................................................................117
Identification of differentially expressed peptides in active and aestivated Theba pisana CNS .................................................................117
Characterization of differentially expressed sCAP .................................119
LIST OF FIGURES

Figure 1.1. Hibernating animals. (A) chipmunk (B) bats (C) bear ........................................27
Figure 1.2. Aestivating animals (A) frog (B) lungfish (C) land snails .....................................29
Figure 1.3. Aestivating land snail Theba pisana showing epiphragm sealing the shell opening. .................................................................30
Figure 1.4. Typical neuropeptide precursor with signal peptide (yellow), proteolytical cleavage sites (red) and peptides (blue). .................................................................38
Figure 1.5. Schematic diagram of the Lymnaea CNS with identified neurons responsive to glutamate. ........................................................................................................40
Figure 1.6. Distribution of Theba pisana in Australia. Blue circles indicate locations where T.pisana have been reported .................................................................43
Figure 2.1. Summary of transcriptome and annotation of genes from Theba pisana CNS, hepatopancreas and foot muscle. ........................................................................51
Figure 2.2. Predicted functional analysis. ........................................................................53
Figure 2.3. Distribution and level of gene expression, based on FPKM ........................................54
Figure 2.4. Schematic representation showing organisation of Theba pisana and Lottia gigantea neuropeptides .................................................................56
Figure 2.5. Characterization of Theba pisana allatostatin C, CCAP, FFamide, sCAP and NKY. .................................................................57
Figure 2.6. Identification and characterization of Theba pisana insulins ....................................59
Figure 2.7. Identification and characterization of Theba pisana prohormone convertases (PCs). .................................................................................................60

Additional Figures and Tables ............................................................................................70
Figure S2.1. Length distribution of unigenes derived from a combined transcriptome assembly of Theba pisana CNS, foot muscle and hepatopancreas ........................................................................70
Figure S2.2. Graphs showing transcript abundance of top 50 transcripts in each Theba pisana transcriptome. Numbers for each correspond to transcripts listed in Table S2.2. ........................................................................71
Figure S2.3. Graph showing transcript abundance for Theba pisana neuropeptides identified. .................................................................................................72
Figure S2.4. LC-MS/MS spectra showing peptides matching FFamide, NKY, sCAP and insulin precursors. .................................................................................................77
Figure S2.5. Theba pisana insulin-like precursors annotated with features characteristic of insulin peptides .................................................................................................78
Figure S2.6. Theba pisana prohormone convertases, PC1 and PC2 ........................................79
Figure 3.3. Hierarchical clustering of peptides relative expression levels in the Theba pisana CNS of aestivated versus active snails based on MS/MS analysis ........................................91
Figure 3.4. Analysis of the Theba pisana buccalins. ..................................................................96
Figure 3.5. In situ hybridisation analysis of Theba pisana buccalin-2 in the aestivated CNS. .................................................................................................99
Figure 3.6. Immunolocalisation analysis of Theba pisana buccalin ....................................100

Additional Figures and Tables ............................................................................................103
Figure S3.1. Histological sections through the T. pisana CNS and stained haematoxylin and eosin. Serial sections are numbered 1-20 ................................................................103
Figure S3.2. Immunolocalisation analysis of aestivating T. pisana buccalin R102LDKFGFSGGI-amide. Serial sections are numbered 1-24 ...................................................................105
Figure S3.3. Immunolocalisation analysis of active T. pisana buccalin R102LDKFGFSGGI-amide. Serial sections are numbered 1-11 ..................................................................106
Fig. 4.1. Hierarchical clustering of peptide expression in the *Theba pisana* CNS of active versus aestivated snails based on MS/MS analysis. The 22 differentially expressed peptides can be divided into two clusters.

Fig. 4.2. Precursor protein sequences with neuropeptide-like characteristics that have no defined function.

Fig. 4.3. Analysis of the *Theba pisana* sCAPs.

Fig. 4.4. Immunolocalisation analysis of *Theba pisana* sCAP.

Fig. 4.5. Analysis of the *Theba pisana* AAP12.

Fig. 4.6. Immunolocalisation analysis of *Theba pisana* AAP12.

Fig. 4.7. Quantitative PCR expression analysis of *Tpi*-sCAP and *Tpi*-AAP12 in CNS of active and aestivated *Theba pisana*.

Additional Figures and Tables.

Figure S4.2. Immunolocalisation analysis of *T. pisana* AAP12 in active CNS.

Figure 5.1. Small RNA analysis A.
LIST OF TABLES

Table 2.1. Summary of molluscan neuropeptides, distribution and characteristics of identified in *Theba pisana*. ................................................................. 55
Table 3.1. List of peptides up- and down-regulated between active and aestivated *Theba pisana* CNS. ........................................................................................................ 92
Table 3.2. List of abundance of differentially expressed gene transcripts of peptides in CNS, hepatopancreas and muscle tissue in *Theba pisana*. ......................................................................... 95
Table S3.1. List of sequences used for analysis................................................................ 107
Table 4.1. List of peptides up- and down-regulated during aestivation.......................... 119
Table 4.2. List of abundance of differentially expressed gene transcripts of peptides in CNS, hepatopancreas and muscle tissue in *Theba pisana*. .................................................. 121
Table S4.1. List of all protein sequences used in this study.......................................... 132
Table S4.2. Mass spectral analysis of CNS showing MS/MS fragmentation data for sCAP and AAP-12 precursor peptides................................................................. 135
Table 5.1. List of most highly differentially expressed miRNAs. ................................. 147
LIST OF ABBREVIATIONS

aa    amino acid
AAP   aestivation associated peptide
ATP   adenosine triphosphate
BCN   bag cell neuron
bp    base pair
CC    cerebral commissure
cDNA  complementary DNA
CG    cerebral ganglia
CNS   central nervous system
COG   cluster of orthologous groups
CCAP  crustacean cardioactive peptide
DPPH  2,2-diphenyl-1-picryl-hydrazyl-hydrate
ET    electron transfer
ELH   egg laying hormone
FABP  fatty acid binding protein
FCAP  feeding circuit-activating peptide
FPKM  fragments per kilobase of exon region in a given gene per million mapped fragments
GDH   glutamate dehydrogenase
G6PDH  glucose-6-phosphate dehydrogenase
GDH   glutamate dehydrogenase
GAP   GnRH associated peptide
GO    gene ontology
HAT   hydrogen atom transfer
HPLC  high performance liquid chromatography
LC-MS liquid chromatography–mass spectrometry
LDA   love dart allohormone
miRNA microRNA
MR    metabolic rate
mRNA  messenger RNA
MS    mass spectrometry
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na⁺/K⁺</td>
<td>sodium/potassium</td>
</tr>
<tr>
<td>NADPH</td>
<td>nicotinamide adenine dinucleotide phosphate</td>
</tr>
<tr>
<td>NPY</td>
<td>neuropeptide Y</td>
</tr>
<tr>
<td>PaG</td>
<td>parietal ganglia</td>
</tr>
<tr>
<td>PC</td>
<td>prohormone convertase</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PDH</td>
<td>pyruvate dehydrogenase</td>
</tr>
<tr>
<td>PE, PeG</td>
<td>pedal ganglia</td>
</tr>
<tr>
<td>PFK</td>
<td>phosphofructokinase</td>
</tr>
<tr>
<td>PK</td>
<td>pyruvate kinase</td>
</tr>
<tr>
<td>PKA</td>
<td>protein kinase A</td>
</tr>
<tr>
<td>PKC</td>
<td>protein kinase C</td>
</tr>
<tr>
<td>PKG</td>
<td>protein kinase G</td>
</tr>
<tr>
<td>PiG</td>
<td>pleural ganglia</td>
</tr>
<tr>
<td>PP</td>
<td>protein phosphatase</td>
</tr>
<tr>
<td>qPCR</td>
<td>quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
</tr>
<tr>
<td>RPP</td>
<td>reversible protein phosphorylation</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>reverse transcriptase PCR</td>
</tr>
<tr>
<td>SCP, SCAP</td>
<td>small cardioactive peptide</td>
</tr>
<tr>
<td>Tb</td>
<td>body temperature</td>
</tr>
<tr>
<td>Tpi</td>
<td><em>Theba pisana</em></td>
</tr>
<tr>
<td>USC</td>
<td>University of the Sunshine Coast</td>
</tr>
<tr>
<td>VG</td>
<td>visceral ganglia</td>
</tr>
</tbody>
</table>
Chapter one: Introduction

Nature abounds with examples of adaptation of form and behaviour to overcome adversity. For instance, to escape predators, many insects are shaped and coloured to mimic their surroundings, while animals such as chameleons (Stuart-Fox, Whiting et al. 2006) and octopuses (Cheney, Grueter et al. 2008) are able to change their colour and pattern to either blend into the background or to mimic another species. Some animals rely on speed and manoeuvrability to escape predation (Howland 1974), while many animals and plants rely on morphological defences such as spines (Cooper and Ginnett 1998, Mikolajewski and Johansson 2004) and armour (Barshaw, Lavalli et al. 2003, Vamosi and Schluter 2004) or produce poison or other chemical deterrents (Dumbacher and Pruett-Jones 1996, Clark, Raxworthy et al. 2005). However, predation is not the only danger, with environmental conditions posing an even greater threat to survival and the range of living organisms is as diverse as the conditions in which they live. From the high temperature and acidity of thermal pools, to the extreme pressure of the deep ocean and to the cold of polar regions, organisms have adapted to conditions that seem uninhabitable.

One climatic challenge faced by many organisms is the extremes of temperature that can occur in the annual cycle of seasons. While some species of mammals, birds and insects overcome this by annual migration (Chapman, Reynolds et al. 2003, Hobson and Wassenaar 2008, Egevang, Stenhouse et al. 2010), many species are unable to travel long distances and must take a vastly different approach to extremes of heat and cold, and the resultant lack of supply of water and/or food. Many mammals enter a state of suspended animation through slowed metabolism (hypometabolism). This can be for a short period of time (daily torpor) or for an extended period (hibernation), when conditions get too cold to ensure an adequate food supply and survival (Wilz and Heldmaier 2000). Other animals, both vertebrate and invertebrate, enter a similar hypometabolic state called aestivation, when conditions become too hot or dry (Storey and Storey 2012). Similar to hibernation, aestivation helps the animal to conserve energy and moisture supplies, allowing the body to withstand temperatures well outside the normal habitable range, thus ensuring survival.

The numerous physiological and cellular level changes that take place during hypometabolism are well studied, however little is known of the global control that must occur to coordinate these changes.
Hibernation, torpor, diapause and aestivation

There are several different states of hypometabolism that animals enter to ensure survival in adverse conditions. These include hibernation, torpor, diapause and aestivation, with a large reduction in metabolic rate (MR) being a common factor.

For endothermic mammals, maintaining normal body temperature in cold conditions requires large amounts of metabolic heat production, and so a large food intake. Hibernators can store fat as a primary fuel source or, in the case of chipmunks (Figure 1.1A) and hamsters, can store food, that can be eaten during arousal periods (Carey, Andrews et al. 2003). Maintaining body temperature is a greater problem for small mammals where the surface area to volume ratio increases (Geiser 2004). As a result of this, those animals that enter a state of hypometabolism are mostly small mammals such as ground squirrels (Hampton, Melvin et al. 2011), rodents (Bartholomew and Cade 1957), bats (Figure 1.1B) (Beer and Richards 1956, Smalley and Dryer 1963), several marsupials (Geiser 1994) and a few species of birds (Geiser and Ruf 1995, Brigham, Körtner et al. 2000). However, few larger animals with the exception of some bears (Figure 1.1C) enter a hibernation-like state (Folk, Larson et al. 1976, Tøien, Blake et al. 2011).

Figure 1.1. Hibernating animals. (A) chipmunk (B) bats (C) bear.


Torpor is characterised by a short term (less than a day) reduction in body temperature ($T_b$) and metabolism, while hibernation is torpor for several weeks or possibly even months in some cases (Storey and Storey 2004), with $T_b$ and MR maintained at even lower levels. In both cases thermoregulation continues, however, the thermoregulation temperature is changed to a level usually just a few degrees above ambient temperature (Ortmann and
Heldmaier 2000). Mean comparison rates for $T_b$ were 17°C for daily torpor (with large variation between species) and 6°C for hibernation. MR reduced to 30% for torpor and 5% for hibernation (as low as 1% for small species) (Geiser 1998). Body temperatures of marsupials, for example, have been shown to fall from 35°C to between 11 - 28°C, depending on species, during torpor, with MR falling to 10-60% of basal MR (BMR). During hibernation, however, $T_b$ falls to between 1-5°C and MR to 2-6% of BMR (Geiser 1994). Metabolic rate during the daily torpor period of a hamster was measured to be about 25% of the euthermic (optimal body temperature) rate, while $T_b$ dropped from 34.2°C to 17.9°C (Heldmaier, Ortmann et al. 2004). Heart rate and respiration are also lowered during both torpor and hibernation, with heart rate in ground squirrels reducing during hibernation from a normal rate of 350-400 beats/min to 6-25 beats/min (Frerichs, Kennedy et al. 1994), and respiration reducing from 100-200 breaths/min to 4-6 breaths/min with long apneic periods (periods with no breathing) in some species (McArthur and Milsom 1991). Typically, animals in hibernation arise for short periods during hibernation (Geiser 1988), in some cases leading to debate on distinctions between torpor and hibernation and the origins of both (Geiser 1998). For example, bouts of deep hibernation in marmots are interrupted by short periods (1-2 days) of activity at normal $T_b$, averaging 11.7 days over a 6-7 month period. The active periods account for 72% of energy use during the overall hibernation time (Heldmaier, Ortmann et al. 2004).

Many mammals that hibernate (and some that don’t) can also enter a state of daily torpor, where $T_b$ drops and metabolism is reduced. However, with daily torpor it is thought that the drop in $T_b$ is responsible for the lowered MR, whereas in hibernation, inhibition of MR occurs (Geiser 2004). In these instances daily torpor can typically last for several hours while hibernation can last for 1-3 weeks.

Diapause is an extreme case of metabolic depression, and there is evidence to suggest that there may be total but reversible cessation of metabolism in some species (Clegg 1997). Diapause is most commonly seen in invertebrates, including flies of the Sarcophaga species (Denlinger 1972), Cecropia silkworms (Schneiderman and Williams 1953), Danaus monarch butterflies (Herman and Tatar 2001), Melanoplus grasshoppers (Tatar, Gray et al. 1997) and the crustaceans Artemia brine shrimp (Qiu, Tsoi et al. 2007) and Daphnia (Stross and Hill 1968). Diapause can occur in embryos, larvae, pupae and adult stages of the life cycle (Tatar and Yin 2001) and development slows or may cease altogether, depending on seasonally changing environmental conditions (Danks 2002). In some species (such as the marine planktonic copepods Labidocera aestiva) diapause is triggered primarily by day length, with
temperature also having some influence (Marcus 1982), while in others (such as the egg parasitoid *Trichogramma cordubensis*) it appears to be triggered solely by temperature (Ventura Garcia, Wajnberg et al. 2002). Not all species enter diapause at the same time, with some species avoiding the cold of winter by entering diapause, while others remain active in winter and are in diapause from spring to autumn (Danks 2002).

In contrast to hibernation, aestivation is dormancy during the summer or dry season. Aestivation is normally entered during times of restricted water or food, usually during periods of extreme summer temperatures (Storey and Storey 2012). Both vertebrates and invertebrates use this state of hypometabolism. Aestivating vertebrates include amphibians such as frogs (*Figure 1.2A*) (Hudson, Lonhienne et al. 2008), toads (Armentrout and Rose 1971) and salamanders (Barry and Shaffer 1994), and reptiles such as lizards (Christian, Webb et al. 2003), crocodiles (Firth, Christian et al. 2010) and snakes (Winne, Willson et al. 2006). A limited number of fish, such as the air-breathing lungfish (*Figure 1.2B*), are also known to aestivate in dry conditions (Frick, Bystriansky et al. 2008). Aestivating invertebrates primarily include pulmonate land snails (*Figure 1.2C*) (Storey 2002), while some earthworms are also known to aestivate (Bayley, Overgaard et al. 2010).

**Figure 1.2.** Aestivating animals (A) frog (B) lungfish (C) land snails.

Aestivation is not solely associated with ectothermic animals, as a diverse range of endotherms are known to enter short term torpor in hot or dry conditions to aid survival. These include echidnas, *Tachyglossus aculeatus* (Brice, Grigg et al. 2002), several marsupials (Geiser and Baudinette 1987), bats (Turbill, Law et al. 2003, Stawski and Geiser 2010) and rodents (Bartholomew and Cade 1957, Wilz and Heldmaier 2000). There are also examples of birds exhibiting short bouts of summer dormancy, as in the case of the Rufous
hummingbird *Selasphorus rufus* (Carpenter and Hixon 1988). These animals can survive harsh conditions by decreasing their MR from 80% (normal resting rate) to less than 1%, although a reduction of 5-40% is more common (Guppy and Withers 1999).

One of the challenges that aestivating animals encounter is the conservation of water within the body, and this can be partially achieved by the animal sealing itself away from the environment. For example, aestivating frogs stay in burrows and utilise layers of shed epidermal cells to form a cocoon to reduce water loss through evaporation (Withers 1998, Tracy, Reynolds et al. 2007). Similarly, African lungfish form a cocoon of dried mucus in the mud (Chew, Chan et al. 2004). Land snails typically elevate themselves to avoid higher temperatures at ground level or seek shelter in places such as crevices or under logs (Storey 2002). In addition, they form an “epiphragm” consisting of a special mucus to seal the opening in their shell (Figure 1.3) (Jokinen 1978, McNair, Kier et al. 1981). The epiphragm in the terrestrial snail *Otala lactea* has a region known as the “Kalkfleck” that allows diffusion of gas, while the remaining area of epiphragm is quite impermeable. It is estimated that the epiphragms in this snail provide up to 20% of the protection against evaporative water loss (Barnhart 1983).

Further protection from desiccation is offered by reduced respiration, as evaporation from the lung causes small quantities of water to be lost during exhalation (Barnhart 1986). Oxygen consumption in the aquatic amphibian salamander *Siren intermedia* was found to decrease faster in the second half of aestivation, and attain a level just 20% of normal (Gehlbach, Gordon et al. 1973). Interestingly, atrophy of the external gills is observed until their tenth week of aestivation. Australian arid zone frogs of the species *Neobatrachus* and *Cyclorana* (both non-aquatic) can reduce their respiration rate to between 14 and 38% of the standard rate after 30 days of aestivation (Withers 1993). In African lungfish, an initial increase in breathing frequency was observed in the early stages of aestivation, followed by a gradual decrease (Delaney, Lahiri et al. 1974).
Land snails have a closable vent, the pneumostome, through which air enters the lungs. This is open continuously in active snails, but only intermittently in aestivating snails. Aestivating snails may take as few as 2-3 breaths/hour (Storey 2002), and patterns of oxygen uptake and carbon dioxide release are not continuous (Barnhart and McMahon 1987). The respiration rate in some snails reduces to approximately 16% of the normal resting rate in the first 3 or 4 days of aestivation, and then stabilises (Horne 1973, Rees and Hand 1990).

During aestivation, urea synthesis continues, however this is not excreted as would be the case in active animals. A build-up of urea, possibly as a result of protein catabolism, ammonia detoxification and the absence of urination during dormancy (Horne 1973) assists with water retention by increasing the osmolality of body fluids to high concentrations (Storey 2002). Urea concentrations many times higher than that of active animals, have been found in aestivating frogs (Withers and Guppy 1996), lungfish (Chew, Chan et al. 2004), earthworms (Bayley, Overgaard et al. 2010) and land snails (Rees and Hand 1993). While water retention using the methods previously described is important, there is also a high tolerance to tissue dehydration in aestivating species (Hillman 1980).

A further result of metabolic depression is the reduction in heart rate observed in aestivating animals. It is known that decreased heart rate is a common feature of aestivation (Storey 2002), however there are a very limited number of animals on which the reduced rate has been recorded. It is known that heart rate in the first 4 weeks of salamander S. intermedia aestivation decreases dramatically, and then continues to progressively decline over 16 weeks to a final rate slightly under half the normal rate (Gehlbach, Gordon et al. 1973). In the African Lungfish, heart rate drops from 22-30 beats/min to 11-16 beats/min in 60 days (Delaney, Lahiri et al. 1974).

Conservation of water is not the only challenge to dormancy. There is also the requirement to withstand long periods without food, and so conservation of fuel reserves is a major issue (Storey 2002). In vertebrates, during aestivation, energy metabolism is mainly achieved by aerobic oxidation of lipids, with some protein catabolism and a small amount of carbohydrates (Jones 2005). Protein use, which is low in the early stages of aestivation, rises as water content decreases, to increase urea synthesis (Storey 2002). In contrast to vertebrates, two species of aestivating snails (Oreohelix strigosa and Oreohelix subrudis) are known to catabolise primarily polysaccharides for the first 2-4 months, then once these are exhausted, to catabolise protein, while a low rate of lipid catabolism occurs throughout aestivation (Rees and Hand 1993).
Metabolic rate depression enables a reduction in energy use, so a greater metabolic depression ensures longer lasting fuel reserves. Decrease in MR can be partly attributed to ceasing digestion and movement, as well as from lowered heart rate and reduced (and irregular) breathing, resulting in altered pH and oxygen consumption. Reduced rates of energy turnover in tissues are responsible for further reductions in energy use (Storey 2002). The mechanisms responsible for controlling MR depression must be quickly reversible, as many animals are known to become active within 5 minutes of conditions improving (Storey 2002). The mechanisms of MR depression include suppressing fuel catabolism, ion channel arrest and reducing protein synthesis (Churchill and Storey 1989, Rees and Hand 1991).

One of the biggest energy consumers is adenosine triphosphate (ATP) turnover, which is used in numerous cellular functions, including maintaining membrane potential. To achieve long-term aestivation a low net ATP turnover and a changed ratio of ATP-producing to ATP-using functions, is essential (Ramnanan and Storey 2006). This can be achieved by a reorganisation of energy use (Storey and Storey 2004) like, for example, reducing protein turnover (synthesis and breakdown), normally the biggest consumer of ATP, by 93-94% during anoxia (a hypometabolic state similar to aestivation). Also, reductions are known to occur in the activity of an ATP membrane ion pump, which maintains gradients of transmembrane sodium and potassium. In the aestivating land snail O. lactea, activity of the membrane enzymes Na⁺/K⁺-ATPase, one of the biggest energy consumers in the cell (Ramnanan and Storey 2006), is reduced by 30%. In turtle liver cells experiencing anoxia, while ATP use by Na⁺ pumping is significantly reduced, the ratio of total ATP used by Na⁺ pumping increases from 28% to 75%, clearly demonstrating not only the importance of maintaining transmembrane gradients, but also the reorganisation of energy use (Hochachka, Buck et al. 1996, Ramnanan and Storey 2006). During aestivation, it is not the amount of the enzyme Na⁺/K⁺-ATPase present in the cell that changes, but rather the enzyme’s activity level. In snails, Na⁺/K⁺-ATPase activity has been found to be related to the presence or absence of phosphate groups, where high phosphate forms of Na⁺/K⁺-ATPase are of low activity, while low phosphate forms (achieved by dephosphorylation) are highly active (Ramnanan and Storey 2006).

Reversible protein phosphorylation (RPP) appears to be a major mechanism controlling metabolic suppression, altering the states of many proteins and cellular enzymes (Storey and Storey 2010). As described for Na⁺/K⁺-ATPase, the removal or addition of a phosphate group can change the activity in proteins, including enzymes, like an on/off control. Other possible actions include changes in affinity for substrates, sensitivity to activators and
inhibitors, and changes in binding to other proteins. In both vertebrates (e.g. toads) and invertebrates (e.g. land snails), phosphorylation has been shown to control the glycolysis regulating enzymes pyruvate kinase (PK) and phosphofructokinase (PFK) during aestivation (Storey and Storey 2010). The less active, low phosphate forms increase in abundance during aestivation (Whitwam and Storey 1990, Whitwam and Storey 1991). Interestingly, PK in the foot muscle of the snail *O. lactea* went from active to aestivating levels between 12 and 48 hours after withdrawal of food and water. However, after 22 days of aestivation, PK levels returned to active levels in just 10 minutes after arousal (Storey 2002). Glycogen phosphorylase and pyruvate dehydrogenase (PDH) are both enzymes involved in carbohydrate catabolism, and are down-regulated by RPP in aestivating snails (Brooks and Storey 1990, Brooks and Storey 1992).

Another enzyme, glucose-6-phosphate dehydrogenase (G6PDH), is also under the control of RPP with the low phosphate form found in active snails and the high phosphate form found in aestivating snails (Ramnanan and Storey 2006). G6PDH controls the entry of carbohydrates into the pentose phosphate pathway, and so is involved in anabolism of biomolecules such as nucleotides, and nicotinamide adenine dinucleotide phosphate (NADPH) which is involved in antioxidant defence (Storey and Storey 2010). Similarly to PK, active PDH levels decreased from 98% to 60% in the first 25-30 hours of aestivation, but returned to normal within 1 hour of arousal (Brooks and Storey 1992). Another important enzyme also regulated by RPP is glutamate dehydrogenase (GDH). GDH is involved in several metabolic processes and is important in both carbohydrate and nitrogen metabolism, and is thought to be critical in alterations to cellular biochemistry during aestivation. Major biochemical changes during aestivation include the increase of osmolytic concentration in tissue, to retain water. This is achieved by the production of urea, mainly from ammonium released by oxidative deamination of glutamate by GDH. GDH can be as much as three times more active in aestivating snail foot muscle than in muscle from control snails (Bell, Dawson et al. 2012). The phosphorylation of several enzymes has been shown to control metabolic pathways in hypometabolism, although the mechanism for phosphorylation of the enzymes is not clear. The glycolysis regulating enzymes PK and PFK are phosphorylated by the cAMP-dependant protein kinase A (PKA) (Storey 2000). The phospholipid-dependant protein kinase C (PKC), involved in cellular functions including proliferation and differentiation, is down-regulated during aestivation, indicating suppression of metabolic processes that are PKC controlled. Similarly, GDH appears to be regulated by the protein kinase G (PKG) (Bell, Dawson et al. 2012). The action of these protein kinases is reversed by protein phosphatases.
(PP), and the activity of PP types 1, 2A, 2B and 2C are reduced during aestivation (Cowan, MacDonald et al. 2000). It is interesting to note that several of these enzymes are active not only in aestivating animals, but have also been found to regulate cellular functions in mammals during hibernation (Storey 2000, Bell and Storey 2010).

Protein turnover, including both synthesis and breakdown is energetically expensive, and suppression of protein synthesis appears to be a common factor in most types of hypometabolism (Storey and Storey 2004). Reductions of 50-80% in protein synthesis have been measured in frogs and snails (Fuery, Withers et al. 1998, Pakay, Withers et al. 2002, Ramnanan, Allan et al. 2009). With protein synthesis slowed, there is also a reduction in proteolysis (breakdown of protein), which not only reduces net turnover of protein but also extends the lifespan of existing proteins during the period of hypometabolism (Storey and Storey 2012). The ribosomal factors that transport initiating methionine and mRNA into the ribosome include the initiation factors 2 (eIF2) and 4 (eIF4) and the eukaryotic elongation factor 2 (eEF2), all of which are regulated by RPP. Phosphorylation of eIF2 prevents that factor from delivering and attaching the initiating methionine to the ribosome; this phosphorylation is increased by 15- to 20-fold in O. lactea tissues after 14 days of aestivation (Storey and Storey 2012). An increase in phosphorylated eIF2 has also been found in aestivating frog liver (Pakay, Hobbs et al. 2003), however, that study found no difference between levels in active and aestivating land snail Helix aspersa. Phosphorylation of eIF4 increases translation (Proud 2007), so reduced phosphorylation observed in the foot muscle of O. lactea (Ramnanan, Allan et al. 2009) is consistent with translational inhibition. Peptide elongation is also inhibited by a 5- to 10-fold increase in eEF2 phosphorylation (Ramnanan, Allan et al. 2009). For the most part, protein synthesis is inhibited during hypometabolism, although an increase in the synthesis of certain proteins has been found in the liver of the lungfish Protopterus annectens during the first 6 days of aestivation, and this is thought to be necessary for reconstruction of protein structures for the maintenance phase of aestivation (Loong, Hiong et al. 2012).

It is clear that RPP is involved in the regulation of multiple cellular processes during aestivation, and although the mechanism of phosphorylation is known, the molecular master control elements that trigger RPP are as yet unknown.
Primary molecular factors that are modified during hypometabolism

Protein stability

Proteins, in hibernators in particular, must be able to function over a wider range of temperatures than in non-hibernating animals. One example of this is the fatty acid binding protein (FABP), which transports long chain fatty acids in cells, and is essential in lipid oxidation, the main metabolic fuel supply during hibernation. FABPs must function at temperatures ranging from ~37°C to <5°C, and FABP from rat (a non-hibernator) and ground squirrel livers were compared for temperature tolerance. Decreased fatty acid binding capacity was found at 5°C in rat FABP, while in squirrel FABP, at 5°C, binding capacity was at least equal to that at 37°C (Storey 2000). While the mechanism that enables this is not known, it is thought that possibly just a few amino acid substitutions could be responsible.

Temperature is not the only condition during hypometabolism that can impact on protein stability. Proteins can be denatured by urea, and urea concentrations are high in some animals, notably in marine elasmobranchs (e.g. sharks, rays, and skates). These animals produce methylamine products, mostly trimethylamine N-oxide (TMAO) to counteract the effects of urea (Yancey and Somero 1979, Bennion and Daggett 2004). Aestivating anurans, including spadefoot toads and Australian desert frogs, also accumulate high levels of urea without protein degradation, but as they do not appear to counteract these with methylamine, it appears that urea insensitivity has developed (Withers and Guppy 1996, Storey 2000).

Anti-oxidants

Enhancement of anti-oxidant defences is essential during hypometabolism to help protect existing macromolecules from damage by reactive oxygen species (ROS). This protection from damage is important during hypometabolism, as the synthesis of macromolecules to replace those damaged is restricted (Pałyga 2011). Anti-oxidants are of particular importance during arousal, where there is a rapid rise in ROS production due to the sudden large increase in oxygen consumption. Anti-oxidants that have been described include non-enzymatic antioxidants such as ascorbic and lipoic acid, polyphenols and carotenoids, derived from dietary sources (Uttara, Singh et al. 2009), and enzymes including superoxide dismutase, catalase, glutathione S-transferase and the low-molecular weight compounds
glutathione, α-tocopherol and ascorbic acid. These act to either remove or change ROS into metabolites that are less toxic (Nowakowska, Świderska-Kołacz et al. 2009). There are also numerous peptides known to have antioxidant properties. These have been found in a diverse range of organisms, including in the muscle of the prawn *Penaeus japonicus* (Suetsuna 2000), in fermented milk products (Hernández-Ledesma, Miralles et al. 2005) and chicken eggs (Rao, Sun et al. 2012), tuna and sharks (Je, Qian et al. 2007, Wang, Li et al. 2012), frogs (Liu, Hong et al. 2010) and in a number of plants (Suetsuna and Chen 2002, Chen, Yang et al. 2012).

Increased levels of anti-oxidant defence have been found in squirrels during hibernation (Drew, Osborne et al. 1999, Drew, Tøien et al. 2002), and in anoxia-tolerant turtles, freeze-tolerant snakes and frogs and aestivating snails (Storey 1996). Reduced oxygen consumption and blood oxygen (PO$_2$) content during aestivation should lower production of ROS, yet levels of antioxidant enzymes were observed to increase in aestivating *O. lactea* snails (Hermes-Lima, Storey et al. 1998). It is thought that this is necessary to allow for high levels of ROS upon initiation of arousal. However, a different strategy may be used by the land snail *H. pomatia*, where antioxidant enzymes were maintained at relatively high levels in both active and aestivating animals (Nowakowska, Świderska-Kołacz et al. 2009), a strategy that is also seen in some species of freshwater turtle (Storey 1996). It appears that where animals are frequently subjected to oxidative stress, constant high antioxidant levels are the preferred method to counteract ROS damage, whereas when the stress is infrequent, as in most cases of aestivation, the preferred method seems to be increasing antioxidant levels during low oxygen stress in preparation for high ROS during arousal (Storey 1996).

**Gene expression**

Only a limited amount of research has been undertaken on differential gene expression, both in hibernation and aestivation. Of those studies, cDNA library screening and cDNA microarray analyses have found that the majority of mRNA transcripts evaluated are down-regulated or remain unchanged, while only a few transcripts appear to be up-regulated (Storey 2000). Up-regulated genes found in hibernating mammals include genes encoding a myosin light chain 1 (*MLC1*) and the mitochondrial NADH-ubiquinone oxidoreductase (*Nad2*), both found in heart muscle of the squirrel *Spermophilus lateralis* (Fahlman, Storey et al. 2000); genes encoding pancreatic lipase (PL) and pyruvate dehydrogenase kinase isozyme
4 (PDK-4) in the heart of *S. tridecemlineatus* (Andrews, Squire et al. 1998); and genes encoding heart-specific fatty-acid proteins (*h-fabp* and *a-fabp*) from the bat *Myotis lucifugis* (Eddy and Storey 2004).

Similar results are observed in aestivating animals. In the liver of the female spadefoot toad, *Scaphiopus couchii*, a riboflavin-binding protein involved with egg production was found to be up-regulated (Storey, Dent et al. 1999), while in the liver of the lungfish, *P. annectens*, several genes involved with urea synthesis, including carbamoyl phosphate synthetase, argininosuccinate synthetise and glutamine synthetise, are up-regulated (Loong, Hiong et al. 2012, Loong, Chng et al. 2012). In a study of heat shock proteins in the snail *Sphincterochila*, heat shock protein 90 was found to be up-regulated during aestivation, while other heat shock proteins were only up-regulated during arousal (Arad, Mizrahi et al. 2010). Currently, the specific function of most up-regulated genes is not fully understood, however, their differential expression indicates that gene expression levels are involved in the control of hypometabolism.

Most recently, transcriptome sequencing and characterisation of specific transcripts has been useful in the identification of genes involved in sea cucumber (*Apostichopus japonica*) aestivation (Du, Bao et al. 2012). A comparison of transcripts from active and aestivating animals revealed 446 differentially expressed genes. Of these, 253 were down-regulated during aestivation and 193 were up-regulated. Of these genes, the functions of 34% are unknown (Du, Bao et al. 2012). A transcriptome analysis of skeletal muscle in the green-striped burrowing frogs (*Cyclorana alboguttata*), which show little muscle dystrophy in spite of long periods of disuse, found a number of genes differentially expressed during aestivation. These genes are involved in cytoskeletal remodelling, avoidance of oxidative stress, energy metabolism, the cell stress response, and apoptotic signalling (Reilly, Schlipalius et al. 2013).

**Neuropeptides and metabolism**

*Background*

A diverse array of cellular modifications is necessary to regulate and co-ordinate changes at the organismal level during hypometabolism. These modifications need to be global, and tightly controlled throughout the organism. Such wide ranging control is typically achieved by chemical signalling using peptides, most of which are neuropeptides (Kiss 2011). Peptides occur in all animals and in plants (Hökfelt, Broberger et al. 2000), and have been
widely studied since the early 1950’s, with the characterisation and identification of the neuropeptides oxytocin and vasopressin (du Vigneaud, Ressler et al. 1953, Popenoe and du Vigneaud 1954).

Most neuropeptides derive from larger inactive precursors containing a signal sequence and there is proteolytic cleavage into a number of small bioactive neuropeptides, sometimes all the same, (Figure 1.4) sometimes with different biological actions (Newcomb and Scheller 1987, Hökfelt, Broberger et al. 2000).

**Figure 1.4.** Typical neuropeptide precursor with signal peptide (yellow), proteolytical cleavage sites (red) and peptides (blue).

In the fruit fly *Drosophila*, a single gene encodes a precursor from which 13 copies of the neuropeptide FMRFamide are produced (Schneider and Taghert 1988), while in the nematode *Ascaris suum* the FMRFamide precursor encodes up to 6 FMRFamide-like neuropeptides (Edison, Messinger et al. 1997). In invertebrates, insulin is expressed in neurons and the insulin precursor is also a good example of a neurohormone derived from a larger inactive precursor. Insulin consists of A and B chains, linked by a connecting (C) peptide. This C peptide is cleaved from the precursor and the A and B chains connect by cysteine bridges to form the mature insulin peptide (Huybrechts, Bonhomme et al. 2010). In many cases, following cleavage of the precursor, the final peptides are subjected to post-translational modification such as glycosylation, C-terminal amidation, acetylation, phosphorylation or sulfation (Hökfelt, Broberger et al. 2000).
Molluscs and neuropeptides

Comprising about 7% of living animals, molluscs are the second largest phylum on earth and occupy habitats ranging from high alpine regions to deep sea vents, with a diverse range of life styles including predatory, scavenging, herbivorous, detritivorous and filter-feeding (Benkendorff 2010). It is estimated that there are possibly 200,000 species of molluscs, with only 80,000-100,000 described (Strong, Gargominy et al. 2008). Gastropod molluscs (commonly referred to as snails), are the most diverse class of mollusc and include marine, freshwater and terrestrial species (Benkendorff 2010) with possibly 30,000-35,000 species of land snails worldwide (Solem 1984).

The sea slug Aplysia has a central nervous system (CNS) similar to many molluscs, with nine connected ganglia distributed around the body (Strumwasser 1971). The main study of neuropeptides in molluscs has been on Aplysia and has been largely focused on peptides involved with reproduction, including the egg laying hormone (ELH) (Chiu and Strumwasser 1981, Painter 1992). In A. californica, ELH is produced in the bag cell neurons (BCNs) and interestingly, studies indicate that ELH synthesis is stimulated by mRNA translation up-regulation rather than by activating gene transcription and that total protein synthesis is inhibited (Lee and Wayne 1998, Wayne, Lee et al. 2004). Other neuropeptides that have been identified in Aplysia include the FMRF-amide, which is possibly responsible for physiological control of gills (Weiss, Goldberg et al. 1984), and has also been found in the heart (Harris, Lesser et al. 1995), the neurohormones SCP A and SCP B, (small cardioactive peptides) which increase heart rate and the amplitude of the beat (Harris, Lesser et al. 1995), and the insulin prohormone, which is processed into one of the largest known insulins (Floyd, Li et al. 1999).

The freshwater snail L. stagnalis has also been used as an experimental model for discovery of molluscan neuropeptides. The CNS (representative diagram from the freshwater snail L. stagnalis, Figure 1.5) is made up of cerebral, buccal, pedal, pleural, parietal and visceral ganglia (Nesic, Magoski et al. 1996, Rigon, de Castilhos et al. 2010). Those well studied neuropeptides in L. stagnalis include an ovulation neuropeptide (caudodorsal cell hormone) similar to Aplysia ELH (Ebberink, Van Loenhout et al. 1985) and five genes coding for the neurohormones APGWamide, neuropeptide Y, conopressin, molluscan insulin-related peptide and pedal peptide, that are involved in muscle contraction and modulation in males (Smit, Jiménez et al. 1992, de Lange, van Golen et al. 1997). Oesophagus contractions in Lymnaea are inhibited by the neurohormone GAPRFVamide (Li, Van Minnen et al. 1996).
Less is known about neuropeptides in other molluscs. However, APGWamide which in pulmonate gastropods appears to be involved in male reproductive organ control, is also found in the deep sea scallop *Placopecten megellanicus*, the mussel *Mytilus edulis* and the oyster *Crassostrea virginica*, although the function in these bivalves, is not yet known (Smith, Nason et al. 1997). Neuropeptides showing sequence homology to pedal peptides in
Aplysia, and the neurons in which they are synthesised, have been identified in the marine nudibranch gastropod Tritonia (Lloyd, Phares et al. 1996), the precursor for the APGWamide neuropeptides has been characterised in the cerebral ganglia of the marine mussel M. edulis (Favrel and Mathieu 1996) and an excitatory neuropeptide, TEP-1, believed to be involved in digestive and reproductive system contractions in the rock-shell Thais clavigera (Morishita, Minakata et al. 2006). In the limpet mollusc Lottia gigantea genome, 59 genes have been annotated from its genome that are predicted to produce neurohormones, with another 8 predicted to produce insulin-related peptides (Veenstra 2010). A number of other molluscan neuropeptides have also been reported (Li, Moroz et al. 1998, Chase 2002, Morishita, Furukawa et al. 2010, Pirger, Lubics et al. 2010, York, Cummins et al. 2012).

There has only been limited high-throughput study on neuropeptides in land snails. Mass spectrometry has been used to determine changes in peptide profiles of the brain and haemolymph of the snail H. pomatia during hibernation (Pirger, Lubics et al. 2010). The output of selected peptides was found to be increased or decreased in response to low temperatures. In the brain, 19 neuropeptides were predominantly present in the active state, with 10 in the haemolymph. Other peptide/polypeptides (11 in the brain and 13 in the haemolymph) were present only in hibernation, while several peptides showed no difference between states (Pirger, Lubics et al. 2010).

**Implications of neuropeptide research in land snails**

The ability to manipulate land snail metabolism, reproduction and growth may be possible through a better understanding of their neuropeptides. This may include enabling the more efficient culture of edible land snails or the control of pest snail species. Several species of non-marine gastropods are of importance as a human food source. There is evidence to suggest that this has been the case in the Mediterranean region for over 10,000 years, where H. aspersa, H. melanostoma, Leucochroa candissima, Helicella setifensis and Otala. sp. were the main species eaten, and are all still found in the region (Lubell 2004). Cooked snail meat, or ‘escargot’ is popular in many countries in Europe, especially France, and is usually H. pomatia or H. lucorum (Abdalmawjood and Bülte 2006). In Italy, H. aspersa and H. pomatia are the main species produced for food (Begg 2003). There are several species of the Giant African snail (Achatina spp and Archachatina spp) eaten in Africa (Barwa 2011, Ogunjimi
and Ajala 2012) and the South American golden apple snail *Pomacea canaliculata*, was introduced into Asia as a dietary protein source for the poor, and subsequently became a major crop pest (Teo 2004).

Many molluscs have been, and are still being used in traditional medicine worldwide. These include marine and land gastropods, cephalopods and chitons in South Africa (Herbert, Hamer et al. 2003), the green lipped mussel in New Zealand (Whitehouse, Macrides et al. 1997), a variety of freshwater gastropods and bivalves in India (Prabhakar and Roy 2009) and a small bivalve in the Torres Strait (Rowland 1994). The use of land snails in medicine started as far back as Hippocrates, with *Helix* spp. being used as treatment for a wide range of ailments, from pain relief to curing hernias and pneumonia (Bonnemain 2005). Mucus secretions from *Cryptomphalus aspersa* (*H. aspersa*) contain antioxidants and stimulate the proliferation of fibroblasts, assisting wound healing (Brieva, Philips et al. 2008).

Neuropeptide research could help control invasive snail populations. There are a number of invasive snail species that have become major pests worldwide. These include aquatic snail pests in reservoirs in Israel (Ben-Ami and Heller 2001), the New Zealand mud snail *Potamopyrgus antipodarum* in Europe, Australia, Japan and North America (Levri, Dermott et al. 2008), the golden apple snail, *P. canaliculata*, in rice fields throughout Asia (Yusa, Sugiura et al. 2006), the giant African snail *A. fulica* throughout the tropics and subtropics (Thiengo, Faraco et al. 2007) and the aquatic snail *Biomphalaria glabrata*, carrier of the *Schistosoma mansoni* parasite (Lewis, Patterson et al. 2000, Kloos, Passos et al. 2004). The common garden snail *H. aspersa* Müller is common in large areas of Western Europe in man-made habitats (Madec, Desbuquois et al. 2000), and is considered a pest in areas it has been introduced into, including North America (Selander and Kaufman 1975), Australia (Rudman 1999) and New Zealand (Barker and Watts 2002).

The small white Italian snail, *Theba pisana* is one of several snails native to the Mediterranean but now established as pests in many countries including Australia, South Africa and North America. In grain crops, pastures and vineyards in southern Australia (Figure 1.6), *Theba* not only eats seedling crops and pastures, but also clogs harvesting equipment and contaminates grain harvests (Baker 1991, Baker 1998, Baker 2008). This snail is a pulmonate (air breathing) gastropod mollusc in the family Helicidae. In Southern Australia the snails are active during autumn, winter and spring, and aestivate in early summer as temperatures increase (Baker and Vogelzang 1988).
Figure 1.6. Distribution of *Theba pisana* in Australia. Blue circles indicate locations where *T.pisana* have been reported.

*Aplysia* has provided an excellent model of neuroscience discoveries for the past 4 decades; the ability to regenerate damaged neurons and the ease with which neurons can be cultured *in vitro* (Moffett 1995), along with their simple neuronal organisation and giant cells makes them very suitable for neurobiological studies (Price 1977). Some of these attributes are shared in land snails. Studies conducted on the CNS of land snails have included the characterisation and localisation of nitric oxide synthase, which produces nitric oxide (NO), a neurotransmitter, in the CNS of the land snail *Megalobulimus abbreviates* (Rigon, de Castilhos et al. 2010) and in *H. pomatia* (Huang, Kerschbaum et al. 1997). The distribution of annexins, involved in regulation of several functions including cell growth, blood coagulation and protein kinase C activity has been studied in *H. pomatia* (Kerschbaum, Donato et al. 1997), while glutamate has been found to act as a neurotransmitter on some neurons in *L. stagnalis* (Nesic, Magoski et al. 1996). Octopamine, which was first discovered in the octopus, and has a variety of functions in molluscs and insects, has been shown to work as a transmitter between buccal feeding neurons in the *L. stagnalis* CNS (Vehovszky, Hiripi et al. 2000). In the African giant snail, *A. fulica* Férussac, five specific pairs of neurons that control buccal muscle movement involved with feeding have been identified, along with possible excitatory transmitters acetylcholine, glutamate and aspartate (Yoshida and Kobayashi 1991).
Further knowledge of neuropeptides present in molluscs and the functions of these neuropeptides will compliment these neuroscience studies and give a more complete picture of the neural control of a variety of physiological and cellular functions not only within molluscs, but possibly in animals in general. This could be of particular interest in relation to control of metabolic processes.

Conclusions

The ability to enter a state of hypometabolism ensures survival of a large range of animals in times of environmental stress. There are several different forms of hypometabolism, however, the basic mechanisms of each are similar. A lowering of metabolic rate and heart rate, steps to conserve moisture and energy use and a change in protein turnover at a cellular level are common to animals entering torpor, hibernation, aestivation and diapause. In some cases the mechanisms for these changes, such as reversible protein phosphorylation, to increase or decrease specific enzyme activity, and so regulate many cellular functions, is now well studied and understood. Changes in anti-oxidant defences also occur but are less consistent between species and are less well understood, although the requirements for the changes are clear. It is known that differential gene expression and epigenetic silencing of genes occurs between normal and hypometabolic states, and also that small RNAs are in some cases involved in gene regulation during hypometabolism.

With such a diverse range of biological functions altered during hypometabolism, it is clear that some global control is necessary to not only regulate but also coordinate these changes, and little is known about this control. Typically, many, if not all of these changes will be controlled by neurohormones, and this is an area where little study has been done.

Aims of this thesis

The intent of this research project was to increase our knowledge and understanding of the control of hypometabolism, specifically aestivation in the pest snail *T. pisana*. The focus in this study was to identify neuropeptides present in *Theba*, determine which were differentially regulated between active and aestivated states, and to target those neuropeptides likely to be involved in the global control and maintenance of aestivation. To achieve these goals, the project has the following aims:
Aim 1) To build a comprehensive transcriptome resource for *T. pisana*, including annotation (Chapter 2). The scope of molecular metabolic studies in molluscs is currently limited due to a scarcity of completed genomes available. Up until now only two molluscs, neither of which are known to aestivate, have had their genomes well annotated. As a result, it is necessary to obtain an extensive database of all transcribed genes in *T. pisana* and this will be procured from CNS, hepatopancreas and muscle tissue, to enable comparison and identification of genes specific to neural tissue.

Aim 2) To identify neuropeptide genes expressed in *T. pisana* (Chapter 2). From the transcriptome (Aim 1) genes encoding neuropeptides can be identified, by comparison of sequences to known neuropeptides, and by specific algorithms that specify neuropeptides-like motifs (such as signal sequences). Bioinformatic analysis will produce a comprehensive list of known and novel genes encoding neuropeptides in the CNS.

Aim 3) To identify neuropeptides expressed in *T. pisana* CNS (Chapter 2). A list of neuropeptides present in the CNS of *T. pisana* will be prepared using liquid chromatography-mass spectrometry of neural ganglia.

Aim 4) To determine which neuropeptide genes and their products are up- or down-regulated during aestivation (Chapters 3 and 4). The presence of neuropeptides in active and aestivating snails will be compared using liquid chromatography-mass spectrometry of neural ganglia. This will provide a list of candidate peptides that are differentially expressed in the CNS during activity and aestivation. Neuropeptide gene expression will be investigated using quantitative polymerase chain reaction (qPCR), using primers designed from identified target neuropeptides gene sequences.

Aim 5) To determine temporal and spatial location of neuropeptide genes, and their products (Chapters 3 and 4). To predict the function of candidate genes and neuropeptides in aestivation (identified in Aims 1-4), an indication of when and where they are expressed is necessary. To achieve this, whole-mount in situ hybridisation and immunolocalisation will be used on CNS tissue from *T. pisana* at various stages of activity and aestivation.

The extensive transcriptome database produced in this study will be of benefit in future molecular studies of not only *Theba*, but many molluscs. Identification and
characterisation of *T. pisana* neuropeptides, particularly those involved in the control of aestivation, may eventually allow production of a peptide mimetic capable of controlling these and a number of other snail pests worldwide.
Chapter 2: Molecular insights into land snail neuropeptides through transcriptome and comparative gene analysis

Published in BMC Genomics 2015, 16:308 Abstract. Available online: http://www.biomedcentral.com/1471-2164/16/308/

Background: Snails belong to the molluscan class Gastropoda, which inhabit land, freshwater and marine environments. Several land snail species, including *Theba pisana*, are crop pests of major concern, causing extensive damage to agriculture and horticulture. A deeper understanding of their molecular biology is necessary in order to develop methods to manipulate land snail populations.

Results: The present study used *in silico* gene data mining of *T. pisana* tissue transcriptomes to predict 24,920 central nervous system (CNS) proteins, 37,661 foot muscle proteins and 40,766 hepatopancreas proteins, which together have 5,236 unique protein functional domains. Neuropeptides, metabolic enzymes and epiphragmin genes dominated expression within the CNS, hepatopancreas and muscle, respectively. Further investigation of the CNS transcriptome demonstrated that it might contain as many as 5,504 genes that encode for proteins destined for extracellular secretion. Neuropeptides form an important class of cell-cell messengers that control or influence various complex metabolic events. A total of 35 full-length neuropeptide genes were abundantly expressed within *T. pisana* CNS, encoding precursors that release molluscan-type bioactive neuropeptide products. These included achatin, allotropin, conopressin, elevenin, FMRFamide, LFRFamide, LRFNVamide, myomodulins, neurokinin Y, PKYMDT, PXFVamide, sCAPamides and several insulin-like peptides. Liquid chromatography-mass spectrometry of neural ganglia confirmed the presence of many of these neuropeptides.

Conclusions: Our results provide the most comprehensive picture of the molecular genes and proteins associated with land snail functioning, including the repertoire of neuropeptides that likely play significant roles in neuroendocrine signalling. This information has the potential to expedite the study of molluscan metabolism and potentially stimulate advances in the biological control of land snail pest species.

Key words: Snail, *Theba pisana*, neuropeptides, central nervous system, muscle, hepatopancreas
Background

Molluscs are the second largest animal phylum, comprising about 7% of living animals and occupying habitats ranging from high alpine regions to deep sea vents, with a diverse range of lifestyles including predatory, scavenging, herbivorous, detritivorous and filter-feeding (Benkendorff 2010). The most abundant class of mollusc are the gastropods, which include the land snails that have evolved independently to life on land. Such a transition has required adaptations towards water saving and breathing dry air.

The central nervous system (CNS), hepatopancreas (or digestive gland) and foot muscle are all key organs controlling or having some influence on a snail’s metabolic rate, whereby a diverse array of cellular modifications is necessary to regulate and co-ordinate changes at the organismal level. Many of these modifications need to be global, and tightly controlled. Such wide ranging control is typically achieved by chemical signalling using neurohormones, many of which are peptides (Kiss 2011). Neuropeptides, which occur in all animals with a nervous system (Hökfelt, Broberger et al. 2000), have been widely studied since the early 1950’s when oxytocin and vasopressin neuropeptides were first identified and characterised (du Vigneaud, Ressler et al. 1953, Popenoe and du Vigneaud 1954). Many neuropeptides derive from larger inactive precursors which are proteolytically cleaved into a number of smaller bioactive neuropeptides, generally all the same, but occasionally with different biological actions (Newcomb and Scheller 1987, Hökfelt, Broberger et al. 2000).

In molluscs, most neuropeptide research has been on the aquatic gastropods, Aplysia, Lymnaea and Lottia. In the marine gastropod, Aplysia, several neuropeptides have been well described that are involved with reproduction and learning (Chiu and Strumwasser 1981, Painter 1992, Veenstra 2010). In A. californica, the best known neuropeptide is FMRF-amide, which appears to provide physiological control of gills (Weiss, Goldberg et al. 1984), and has also been found in heart tissue (Harris, Lesser et al. 1995), along with the small cardioactive peptides (SCP A and SCP B) which increase heart rate and the amplitude of the beat (Harris, Lesser et al. 1995). The freshwater snail L. stagnalis has also been a model for mollusc neuropeptide research where studies have discovered five genes coding for the neuropeptides APGWamide, neuropeptide Y, conopressin, molluscan insulin-related peptide, and pedal peptide that are involved in muscle contraction and modulation in males (Smit, Jiménez et al. 1992, de Lange, van Golen et al. 1997). In silico genome and transcriptome database mining have proven effective for high-throughput annotation of the presence and expression of neuropeptides in Lottia, as well as in bivalve oysters (Veenstra 2010, Stewart, Favrel et al.
Less is known about the neuropeptide repertoire of terrestrial pulmonate gastropods, although recently, mass spectrometry has been used to identify and determine changes in neuropeptide profiles in the brain and haemolymph of the snail Helix pomatia during activity versus hibernation (Pirger, Lubics et al. 2010). In that study, 19 neuropeptides were identified as being more highly produced within the brain of active snails. Meanwhile, during hibernation, 11 neuropeptides were exclusively present (Pirger, Lubics et al. 2010).

In the current study, we investigated the CNS, hepatopancreas and foot muscle transcriptomes of the land snail Theba pisana through gene and peptide analysis. We found numerous neuropeptide precursors that show similarity with other known molluscan neuropeptide precursors, and also demonstrate the existence of numerous other genes that encode peptides destined for secretion. This represents the most extensive analysis of neuropeptide genes and their products in a terrestrial snail.

Results

De novo assembly and comparison of Theba CNS, foot and hepatopancreas

Transcript libraries derived from T. pisana CNS, foot and hepatopancreas tissues were sequenced using Illumina technologies and assembled. All sequence data was deposited in the NCBI Genbank under SRA file SRP056280. A summary of the number of high quality raw reads, contigs and unigenes for each tissue is shown in Figure 2.1. The CNS, foot and hepatopancreas transcriptomes encoded 220,602, 201,746, and 186,132 unigenes, respectively. A unigene is typically interpreted as representing a single genomic locus; hence, these groups represent the first comprehensive non-redundant putative transcript database for T. pisana. Collectively, contigs and unigenes could be assembled into a total of 250,848 consensus sequences ranging in size from 200 to 26,246 bp; size distribution for the pooled tissue transcriptome is shown in Figure S2.1. Approximately 156,386 representative sequences are shared within the three transcriptomes (Fig. 2.1). Although CNS has more unique unigenes compared to foot or hepatopancreas, there was no substantial difference based on protein functional domain annotation (Fig. 2.1). In total, there were 24,920 CNS proteins, 37,661 foot proteins and 40,766 hepatopancreas proteins annotated, with 5236 unique Pfam domains. Combining the unigene and Pfam comparison results, the data revealed that the three tissues have a moderate number of tissue-specific transcripts that encode proteins that play similar functions in cellular processes.
Figure 2.1. Summary of transcriptome and annotation of genes from *Theba pisana* CNS, hepatopancreas and foot muscle. Figure shows the CNS, including regions of cerebral ganglia (CG), cerebral commissure (CC), mesocerebrum (meso), procerebrum (pro), metacerebrum (meta) and pedal ganglia (Pe). Venn diagrams show comparisons of representative sequences and protein domain annotation between each tissue transcriptome.
**Functional annotation**

*T. pisana* sequences were annotated against protein databases (Nr, Nt, SwissProt, KEGG, GO and COG) using BLASTX (E-value < 0.00001). From the 250,848 consensus sequences, 69,799 (27.8%) had at least one hit. The sequence names and annotation information of all sequences are provided in Table S1 (online: http://www.biomedcentral.com/1471-2164/16/308/). The majority of transcripts had either a significant match with those from the pacific oyster (*Crassostrea gigas*) or did not match to any known genes; this result is most likely due to insufficient sequences being available in public databases from phylogenetically closely related species. The annotation rate in our study is comparable to those that have been reported in previous *de novo* transcriptome sequencing studies for molluscs (Pairrett and Serb 2013, Artigaud, Thorne et al. 2014).

Gene ontology was performed to classify functions to *T. pisana* genes (Fig. 2.2A). Of these, 77,715 transcripts were assigned to functional categories of ‘Biological Process’ (48.5%), 27,187 to ‘Molecular Function’ (34.5%) and 55,346 to ‘Cellular Component’ (17%). Functional annotation of all transcripts combined against the cluster of orthologous groups (COG) database is shown in Figure 2.2B. These were assigned to four primary COG classes: Information storage and processing (8156 transcripts), Cellular processes and signalling (7756 transcripts), and Metabolism (9932 transcripts) as well as poorly characterized genes (7,911 transcripts). The COG functional classification demonstrates that the most abundant classification is “general function prediction only”, followed by “translation, ribosomal structure and biogenesis” and “transcription”.

Transcript abundance was determined based on fragments per kilobase of exon per million fragments mapped) [FPKM; Fig. S2.2 and Table S1 (online: http://www.biomedcentral.com/1471-2164/16/308/)]. In the CNS, of the 50 most abundant transcripts (besides ribosomal), 22% encoded for neuropeptides such as insulin, neuropeptide Y, myomodulin and achatin, whereas unannotated transcripts comprised 40% of transcripts. In the hepatopancreas, many of the top 50 abundant transcripts (minus ribosomal) encoded for catabolic enzymes including cathepsin peptidase, serine peptidase, chitotriosidase-1, myosinase, lysozyme, while 20% were unannotated including the most abundant transcript (Unigene64357_All). In the foot muscle, the epiphragmin-encoding transcripts were most prominent. Common transcripts that were of high abundance in all three transcriptomes included those encoding a polyubiquitin protein, heat shock protein 70 and an elongation factor 1 alpha.
**Figure 2.2.** Predicted functional analysis. (A) Analysis of gene ontology to classify functions to *T. pisana* genes (B) Graph showing the assignment of the *T. pisana* unigenes to categories of the eukaryotic cluster of orthologous groups of proteins (COG). The main COG categories (functional classes) are represented with different colours.
Annotation of mollusc proteins secreted from CNS in Theba pisana

Investigation of the T. pisana CNS transcriptome, including identification of those precursors containing signal peptides and no transmembrane domains, revealed that *Theba* might contain as many as 5504 precursor proteins that are destined to secrete peptides (Table S2 online: http://www.biomedcentral.com/1471-2164/16/308/). Of those, 4649 putative proteins are expressed in all three transcriptomes (Fig. 2.3A). Another 213 were expressed in both CNS and foot, while there appeared to be no overlap in expression between CNS and hepatopancreas. The remaining 642 proteins were expressed exclusively in CNS. On the basis of a fold change higher than 5, we defined a list of 849 putative secreted proteins in CNS. As shown in Figure 2.3B, the expression FPKM for the majority of the putative secreted peptides were less than 3. Of those where expression was >100 and 10-100, 81% and 22% are designated as encoding hypothetical proteins, respectively (Table S2 online: http://www.biomedcentral.com/1471-2164/16/308/).

![Figure 2.3](image_url)

**Figure 2.3.** Distribution and level of gene expression, based on FPKM. (A) Relative distribution of CNS predicted secreted proteins. (B) Relative expression level of secreted protein transcripts in the CNS.
Of the putative secreted proteins in the CNS, full-length precursors were identified in *T. pisana* for 35 known molluscan neuropeptides (Table 2.1), including isoforms and insulin that in molluscs can be secreted from neural tissue. In Table 2.1, a “white” cell is defined as indicating only a trace amount or no transcript (<10). In general, for these neuropeptide transcripts, abundance was most prominent within neural tissue, particularly for insulin2, NPY, myomodulin1 and achatin (Table 2.1 and Fig. S2.3).

**Table 2.1.** Summary of molluscan neuropeptides, distribution and characteristics of identified in *Theba pisana.*

<table>
<thead>
<tr>
<th>Name</th>
<th>full length</th>
<th>signal</th>
<th>repetitive</th>
<th>cysteine-rich</th>
<th>MS</th>
<th>CNS</th>
<th>Hepatopancreas</th>
<th>Foot</th>
<th>full length</th>
<th>signal</th>
<th>repetitive</th>
<th>cysteine-rich</th>
<th>MS</th>
<th>CNS</th>
<th>Hepatopancreas</th>
<th>Foot</th>
</tr>
</thead>
<tbody>
<tr>
<td>achatin</td>
<td>126</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>155</td>
<td>22</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>allatostatinC</td>
<td>80</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>193</td>
<td>25</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>allatotrocin</td>
<td>130</td>
<td>25</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>543</td>
<td>19</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCAP-1</td>
<td>259</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>107</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCAP-2</td>
<td>246</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>343</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cerebrin</td>
<td>84</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>323</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>elevenin</td>
<td>148</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FFamide-1</td>
<td>91</td>
<td>19</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>142</td>
<td>22</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FFamide-2</td>
<td>110</td>
<td>41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>137</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMRFamide-1</td>
<td>369</td>
<td>25</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>85</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMRFamide-2</td>
<td>352</td>
<td>25</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>112</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGG-1</td>
<td>153</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>172</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGG-2</td>
<td>133</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>695</td>
<td>20</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>insulin1a</td>
<td>200</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>136</td>
<td>24</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>insulin1b</td>
<td>193</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>136</td>
<td>24</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>insulin1c</td>
<td>186</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>insulin1d</td>
<td>179</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>insulin2</td>
<td>131</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>insulin3</td>
<td>174</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>insulin4</td>
<td>185</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scale (FPKM):</th>
<th>&lt;10</th>
<th>50</th>
<th>100</th>
<th>500</th>
<th>1000</th>
<th>&gt;1000</th>
</tr>
</thead>
</table>

Full-length precursors for some of these neuropeptides were also found to be of intermediate abundance in foot muscle while absent from the hepatopancreas. Peptide extraction followed by LC-MS analysis was used to confirm the usefulness of transcriptome-derived protein libraries and complement *in silico* neuropeptide predictions (Table 2.1 and Fig. S2.4). Figure 2.4 shows the organisation of *T. pisana* neuropeptide precursors compared
with previously identified homolog precursors of a marine gastropod, the limpet *Lottia gigantea*, with general high conservation in the spatial organisation of bioactive peptide sequences, as well as precursor size, cleavage sites and position of cysteine residues. Several of these neuropeptides were targeted for more in-depth tissue expression analysis.

<table>
<thead>
<tr>
<th>Theba pisana</th>
<th>Lottia gigantea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achatin</td>
<td>Achatin</td>
</tr>
<tr>
<td>Atlatostatin-C</td>
<td>Atlatostatin</td>
</tr>
<tr>
<td>Atlatophin</td>
<td>Atlatophin</td>
</tr>
<tr>
<td>CCAP-1</td>
<td>CCAP</td>
</tr>
<tr>
<td>CCAP-2</td>
<td></td>
</tr>
<tr>
<td>Cerebrin</td>
<td>Cerebrin</td>
</tr>
<tr>
<td>Elevecin</td>
<td>Elevecin</td>
</tr>
<tr>
<td>FFamide</td>
<td>FFamide</td>
</tr>
<tr>
<td>FMRF-a-1</td>
<td>FMRF-a</td>
</tr>
<tr>
<td>FMRF-a2</td>
<td></td>
</tr>
<tr>
<td>GOGQ-1</td>
<td>GOGQ-1</td>
</tr>
<tr>
<td>GOGQ-2</td>
<td></td>
</tr>
<tr>
<td>LRF-1</td>
<td>LRF-1</td>
</tr>
<tr>
<td>LRF-1a</td>
<td>LRF-1a</td>
</tr>
<tr>
<td>Lupin</td>
<td>Lupin</td>
</tr>
<tr>
<td>Narmelonin-1</td>
<td>Narmelonin-1</td>
</tr>
<tr>
<td>Narmelonin-2</td>
<td></td>
</tr>
<tr>
<td>NPY</td>
<td>NPY</td>
</tr>
<tr>
<td>NPY-1</td>
<td>NPY-1</td>
</tr>
<tr>
<td>PKY/MDT</td>
<td>PKY/MDT</td>
</tr>
<tr>
<td>Phewhin</td>
<td>Phewhin</td>
</tr>
<tr>
<td>PXF-a</td>
<td>PXF-a</td>
</tr>
<tr>
<td>scCAP-1</td>
<td></td>
</tr>
<tr>
<td>scCAP-2</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2.4.** Schematic representation showing organisation of *Theba pisana* and *Lottia gigantea* neuropeptides. Yellow, signal peptide; blue, bioactive peptides; C, cysteine. Scale bars at bottom show length in amino acids (aa).
The Tpi-Allatostatin C precursor comprised 80 amino acids, including signal peptide and cleavage sites to release the 15 amino acid active allatostatin C located in the C-terminal region of the precursor (Fig. 2.5). Most similarity with Allatostatin C from other species was found within the bioactive peptide, primarily at the N-terminal SH and spatial conservation of two cysteine residues that likely form disulfide bonds to help stabilise the peptide. Tpi-allatostatin C was only identified in the CNS transcriptome, at a low expression level (Table 2.1) but RT-PCR analysis showed that it was expressed not only in cerebral and pleural ganglia but also in T. pisana digestive tract and skin tissue (Fig. 2.5). A representative RT-PCR of the 18s rRNA gene is also shown, confirming the integrity of all tissue RNA.

Figure 2.5. Characterization of Theba pisana allatostatin C, CCAP, FFamide, sCAP and NKY. Each protein is represented by; (1) primary amino acid sequence with signal sequence (yellow), cleavage sites (red), cysteine residues (pink) and bioactive region (blue); (2) multiple amino acid sequence alignment with other gastropod and bivalve molluscs; and (3) RT-PCR expression of the neuropeptide gene within various tissues.
Two Crustacean cardioactive peptide (CCAP) precursor transcripts were identified in *T. pisana*, both encoding 3 CCAP moieties and showing very high amino acid identity with *Lottia* and *Aplysia* homologs (Fig. 2.5). Tpi-CCAP1 and 2 transcripts were identified in CNS and foot muscle transcriptomes, with CNS CCAP-2 being more abundant (Table 2.1). RT-PCR demonstrated that Tpi-CCAP1 is additionally expressed within sensory tentacles. The Tpi-FFamide transcript encoded 2 different amidated peptides; an LLFamide peptide and an LFFamide peptide, which show conservation with known homologs from other species within the C-terminal regions (Fig. 2.5). Mass spectral analysis of the CNS of active *T. pisana* identified peptides corresponding to both LLFamide and LFFamide. Both Tpi-FFa transcripts were identified in the CNS transcriptome, with FF-1 very highly expressed (Table 2.1). Tpi-FFa-1 was additionally expressed within the tentacles, digestive tract, skin and foot muscle tissues, despite not being identified from the foot muscle transcriptome.

Two Tpi-NKYamide transcripts were identified that encode neuropeptide Y (NKY) precursors of 142 and 137 residues, each containing cleavage sites to release two bioactive peptides, a 24-residue peptide and the 38-residue NKY peptide (Fig. 2.5). Little amino acid conservation exists for the 24-residue peptide with other species. For the 38-residue NKY precursor, conservation with other species exists primarily within the N- and C-terminal regions, with very limited amino acid identity in the middle region. Mass spectral analysis of the CNS identified peptides corresponding to both NKY precursor peptides. Both Tpi-NKY1 transcripts were observed in the CNS and foot muscle transcriptomes, while Tpi-NKY2 was detected in very low amounts in the CNS transcriptome only. RT-PCR found that Tpi-NKY1 was also present in the tentacles and Tpi-NKY2 was additionally found in the tentacles, digestive tract, skin and foot muscle tissue. Two Tpi-sCAP transcripts were identified that both encode precursor proteins of 136 amino acids, including spatial conservation of putative cleavage sites and cysteine residues (Fig. 2.5). High conservation between species was observed in the N-terminal region of the bioactive peptide. Mass spectral analysis of the CNS identified peptides corresponding to both sCAP precursor peptides. Both Tpi-sCAP transcripts were identified in high abundance within the CNS, and less in the foot muscle, while RT-PCR showed expression of Tpi-sCAP1 in the tentacles, digestive tract and skin.

We identified from the three *T. pisana* transcriptomes, 7 transcripts with strong homology to insulins. Multiple sequence comparison and phylogeny demonstrated 4 groups of insulin-like peptides with strong bootstrap support; Tpi-insulin1a, Tpi-insulin2, Tpi-insulin3 and Tpi-insulin4 (Table 2.1, Fig. 2.6 and Fig. S2.5). Tpi-insulin precursors vary in size from the largest at 200 amino acids (Tpi-insulin1a) to the smallest at 131 amino acids.
(Tpi-insulin2). Consistent with other species, the organisation of T. pisana insulins consists of a “B peptide” with 3 cysteines, followed by a “C peptide” and then a C-terminal “A peptide” containing 5 cysteines. Only the Tpi-insulin2 A peptide contains a glycine residue site for amidation. Tpi-insulin2 and Tpi-insulin4 were highly expressed in the CNS, while only Tpi-insulin1c could be identified in another transcriptome (foot muscle). No further RT-PCR analysis was performed to investigate insulin expression in other snail tissues.

Figure 2.6. Identification and characterization of Theba pisana insulins. Phylogeny of 7 insulins identified and representative schematic showing the organisation of the signal peptides (yellow), bioactive peptides (blue), cleavage sites at paired basic and tetrabasic amino acids (vertical lines) and cysteine (C) residues. Scale bar represents amino acid substitutions.

Prohormone convertases

We identified transcripts in each transcriptome database encoding enzymes that cleave precursor proteins, including the prohormone convertase 1 (PC1) and PC2 (Fig. 2.7A, B and Fig. S2.6). Tpi-PC1 and Tpi-PC2 encode precursors of 648 and 652 amino acids, respectively, and contain regions typical of these processing enzymes, including signal, pro
and catalytic regions that are required for cleavage of dibasic amino acids. Another type of convertase, the *Tpi-furin* was also identified that encodes a partial-length furin (554 amino acids) with a large 3’ untranslated region. Multiple sequence alignment and phylogeny support their enzyme grouping (Fig. 2.7C).

**Figure 2.7.** Identification and characterization of *Theba pisana* prohormone convertases (PCs). *Theba pisana* PCs cluster within the PC1, PC2 and furin-type convertases. Amino acid alignment between PC1 and PC2 shows strong homology, especially within the catalytic region, shown schematically; shown are signal (pre; yellow), pro domain (blue) and catalytic region (red).

**Discussion**

In this study, we extend on our understanding of gastropod biology by investigating transcriptomes of the land snail, *Theba pisana*. Deep transcript sequencing followed by *in silico* predictions and mass spectral analysis enabled the identification of numerous
molluscan genes in this snail, including previously unknown neuropeptides. This greatly enhances our ability to utilise this species as an experimental model in neuroscience and metabolism, with the potential to enable development of molecular tools that could manipulate the physiology and behaviour of this agricultural and horticultural pest species.

Deep sequencing was performed to enable the assembly of three transcriptomes representing the CNS, hepatopancreas and foot muscle. Since gene expression levels may vary considerably depending on animal behavioural and physiological state, tissue samples for RNA transcriptome sequencing were combined from stages representing active and aestivated, reproductive and juvenile snails. With over 300 million raw reads providing approximately 200,000 unigenes for each tissue, this ensured a comprehensive analysis of tissue genes. De novo assembly was necessary since no annotated reference is currently available for Theba or any other helicid snail. BLAST analysis showed that the majority of genes matched most closely with oyster, Crassostrea gigas, which represents the most expansive list of molluscan genes within the NCBI database due to the release of its genome and associated transcriptomes (Zhang, Fang et al. 2012). As a bivalve, C. gigas belongs to a different taxonomic group, thus our submission of the Theba transcripts derived from this study ensures that helicid snail gene sequences are now far better represented in the NCBI databases. Abundant unannotated transcripts suggest that there may be numerous helicid-specific sequences, or this species is at least considerably different to those currently annotated.

Gastropods have long been used in experimental neuroscience projects, providing a rich source for neurological discoveries. The vast majority of these studies have used marine and freshwater gastropods, including Aplysia and Lymnaea. In Lymnaea stagnalis, two separate studies analysed neural transcriptomes, first using normalized EST screening to identify 7,712 distinct transcripts (Feng, Zhang et al. 2009), which was then expanded via Illumina sequencing to provide 116,355 contigs (Sadamoto, Takahashi et al. 2012). A large-scale neural EST screen of Aplysia CNS has revealed 175,000 transcripts, including 19,814 unique (at the time) neuronal gene products (Moroz, Edwards et al. 2006). In our study, Illumina sequencing enabled a more comprehensive identification of T. pisana CNS transcripts, providing 220,602 unigenes. Several of the Theba neuropeptides identified have been shown to be significantly up- or down-regulated in the brain or hemolymph of hibernating Helix pomatia when compared to active individuals (Pirger, Lubics et al. 2010). For instance, levels of peptides derived from the FMRFa, myomodulin, and probably LRNFVa precursors are much more abundant when H. pomatia are in the hibernation state,
suggesting they may be critical for initiating and/or maintaining metabolic arrest. The CNS, hepatopancreas and foot muscle are crucial tissues regulating how snails initiate and maintain deep hypometabolic states, and therefore our comprehensive transcriptomes and associated transcriptome-derived protein databases should provide a useful platform to use for comprehensive gene-associated hypometabolic studies in these animals (Ito, Okada et al. 2012, Lama, Bell et al. 2013).

It has been speculated that in invertebrates, insulins, which are found within neurons, may have a role in neurotransmission (Boyd, Clarke et al. 1985). Similar to vertebrate homologs, the invertebrate insulin precursor contains A and B chains, linked by a connecting (C) peptide. This C peptide is cleaved from the precursor and the A and B chains connect by cysteine bridges to form the mature insulin peptide (Huybrechts, Bonhomme et al. 2010). In many cases, following cleavage of the precursor, the mature peptides are subjected to post-translational modification such as glycosylation, C-terminal amidation, acetylation, phosphorylation or sulfation (Hökfelt, Broberger et al. 2000). Insulin precursors are known to undergo a series of processing events that yield the functional insulin, consisting of A and B peptides bound by intermolecular disulfide bonds. We found that T. pisana have 7 insulins, with each sharing similar attributes common to all other insulins. The largest reported molluscan insulin has been Aplysia insulin; its involvement in regulating glucose metabolism was determined through the demonstration that Aplysia insulin can reduce hemolymph glucose levels (Floyd, Li et al. 1999). Of all the T. pisana insulins, the Tpi-insulin2 transcript appears to be the most highly expressed insulin in the CNS. Future studies are planned to determine the molecular role for this insulin.

This study shows that a prohormone convertase is present within a land snail. The ability for a cell to secrete peptides relies on specific molecular secretory machinery, including processing enzymes that contribute to the synthesis of a mature bioactive peptide. In molluscs, PCs have been reported for Lymnaea, Aplysia and Haliotis (Nagle, Garcia et al. 1995, Smit, Spijker et al. 1996, Cummins, York et al. 2012). The Lymnaea PC2 gene is exclusively expressed in the neuroendocrine system and may be present as two alternatively-spliced isoforms, similar to the Haliotis genes which are differentially expressed in neural tissues (Cummins, York et al. 2012). We found only one PC2 gene (as well as one PC1) in T. pisana, indicating that one of each is sufficient in this species to perform precursor cleavages, including cleavage of allatostatin C, FFamide and sCAP.

For gastropods, the hepatopancreas and foot muscle are major sites for lipid metabolism and storage, and therefore are critical during periods of starvation and
hypometabolism (Böer, Graeve et al. 2006, Lama, Bell et al. 2013). The hepatopancreas is also a major source of digestive enzymes and is involved in nutrient absorption, food storage and excretion (Barker 2002). For that reason, it is not surprising that metabolic enzyme genes dominate transcripts annotated within the hepatopancreas. The foot muscle, on the other hand, contains numerous epiphragmin gene transcripts that encode for the epiphragmin protein, a major constituent of the epiphragm mucus that seals off the aperture during hibernation and aestivation (Li and Graham 2007). The dried epiphragm is necessary to enable moisture retention while still allowing for gaseous exchange (Barnhart 1983). In the snail Cernuella virgata, a major source of epiphragmin appears to be the mucous glands (Li and Graham 2007), while our study also implicates the foot muscle as a rich source of this protein. The foot contains its own type of mucus-secreting cells that were likely incorporated into our foot tissue transcriptome. During prolonged hypometabolism, muscle disuse atrophy (possibly related to reactive oxygen species) may be a major issue and we expect that future use of this transcriptome may enable us to establish molecular pathways by which these snails circumvent damage during hibernation and aestivation. Enzymes that have been implicated as key proteins in metabolic depression in other gastropod species, such as pyruvate kinase, phosphofructokinase and glutamate dehydrogenase (GDH) (Brooks and Storey 1992, Ramnanan and Storey 2006) are present in T. pisana transcriptomes and also represent obvious targets for future metabolic studies in this species. Furthermore, glutamate dehydrogenase (GDH) activity can be as much as three times higher in aestivating snails than in active animals; GDH is crucial to amino acid metabolism and in aestivating snails helps to funnel nitrogen into urea biosynthesis which leads to the elevated urea in body fluids that contributes to desiccation resistance (Bell and Storey 2010, Bell, Dawson et al. 2012).

The identification of genes involved in the maintenance of land snails, as reported here, has numerous implications. Prior to this study, there was a lack of gene data for land snail species, despite their importance in ecological functioning. Theba pisana is one of several snails native to the Mediterranean region that are now established as pests in grain crops, pastures and vineyards in southern Australia, where they not only eat seedling crops and pastures, but also clog harvesting equipment and contaminate grain harvests (Baker 1991, Baker 1998, Baker 2008). The current study provides a foundation for further studies into the manipulation of pest snail behavioural/physiological states that could lead to novel treatments to control their populations. For example, as demonstrated in insect pests (Nachman, Pietrantonio et al. 2009, Nachman and Pietrantonio 2010), neuropeptide mimetic analogs hold potential for blocking or overstimulating receptors. There are also a number of other
invasive snail species that have become major pests worldwide, including aquatic snail pests in reservoirs in Israel (Ben-Ami and Heller 2001), the New Zealand mud snail *Potamopyrgus antipodarum* in Europe, Australia, Japan and North America (Levri, Dermott et al. 2008), the golden apple snail, *P. canaliculata*, in rice fields throughout Asia (Yusa, Sugiura et al. 2006), and the giant African snail *A. fulica* throughout the tropics and subtropics (Thiengo, Faraco et al. 2007). In addition, the common garden snail *H. aspersa* Müller now inhabits large areas of Western Europe in man-made habitats (Madec, Desbuquois et al. 2000) and is an introduced pest of many areas of the world, including North America (Selander and Kaufman 1975), Australia (Rudman 1999) and New Zealand (Barker and Watts 2002).

**Conclusion**

This study has identified the genes expressed in three tissues of *Theba pisana* and identified numerous CNS neuropeptide gene products, confirming their expression *in vivo* by RT-PCR and mass spectrometry. Investigations can now progress into defining the function of gene products.

**Materials and Methods**

**Animals and tissue collection**

*Theba pisana* were collected from agricultural sites surrounding Warooka, located on the Yorke Peninsula, South Australia in early spring (September). Snails were transported to the University of the Sunshine Coast (USC) and housed within purpose-built enclosures. Snails were provided with water and food (cucumber and carrot) ad *libitum* and maintained at room temperature.

Active snails were conditioned by feeding and supplied with water daily over 7 days. Snails to be used for aestivation experiments were placed into glass jars without food or water and kept in an incubator on a cycle of 12 h at 30°C with light, 12 h at 20°C dark, to emulate South Australian summer conditions. The positions of the snails were marked on the jars after 14 days. Any snails that had not moved following a further 21 days were deemed to be in aestivation. To obtain waking snails, aestivated snails were sprayed with water and harvested as soon as they became active (10-60 min). CNS, hepatopancreas (digestive gland) and foot muscle tissue was dissected out of *T. pisana* and immediately frozen. In each case,
tissue from active, waking and aestivating snails was combined to maximise transcript representation from each metabolic state.

**RNA isolation and transcriptome sequencing**

RNA was extracted from tissue using TRIzol Reagent (Invitrogen), as per the manufacturer’s protocol. Following extraction, RNA was assessed for quality by visualisation on a 1.2% denaturing formaldehyde agarose gel, quantified using a Nanodrop spectrophotometer (Thermo scientific). For complimentary DNA (cDNA) synthesis, RNA samples were subjected to oligo-dT selection for mRNA purification and fragmented into small fragments. Fragmented RNA samples were subsequently repaired before adapter ligation. Suitable fragments were selected and reversed-transcribed into double-stranded cDNAs. The cDNA libraries were constructed by PCR amplification using random hexamer primed cDNAs. Finally, the samples were used for library construction and sequenced using an Illumina HiSeq 2000 sequencing (BGI, Hong Kong).

**Gene ontology and identification**

*De novo* assemblies for each tissue type were performed by Trinity software (Grabherr, Haas et al. 2011) using trimmed reads from Illumina sequencing. The assembler was run with the parameter sets as following: seqType, fq; minimum kmer coverage = 4; minimum contig length of 200 bp. Sequences without Ns and which could not be extended on either end were defined as unigenes. When there were several samples from the same species, TGICL (Pertea, Huang et al. 2003) was used to assemble all the unigenes from different samples to form a single set of non-redundant unigenes. After clustering, the unigenes were divided into two classes, clusters with the prefix CL, and singletons with the prefix Unigene.

Transcripts were annotated with the databases of NR, NT, Swiss-Prot, KEGG, COG, and GO, using BLAST and BLAST2GO software. Relative abundances of all transcripts among different tissues were estimated by SOAP software version 2.21. Searches of molluscan neurohormone precursors were also conducted by performing tBLASTn search of all tissue transcripts, which were reported by previous “omics” analysis studies (Gard, Lenz et al. 2009, Ma, Bors et al. 2009, Christie, Durkin et al. 2010, Christie, McCoole et al. 2011, Veenstra, Rombauts et al. 2012). BLAST searches were performed in CLC Main Workbench.
(Version 6.0). All hits were then analyzed manually with their orthologous peptides from various species and then their structures were characterized.

Analysis of protein identity/similarity between different protein receptors was performed by protein alignment using clustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/). The percent identity was calculated as the number of identical amino acid residues, as indicated by "*" symbol in the Clustal output, divided by the total number of amino acid residues of the longest sequence (x100). All hits (lowest E-value) were run through the SignalP 3.0 website for signal peptide prediction with the Neural Networks algorithm (Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark; http://www.cbs.dtu.dk/services/SignalP/ ref). Proteolytic cleavage sites as well as post-translational modifications were predicted based on homology to other known peptides and the Neuropred tool (neuroproteomics.scs.illinois.edu/neuropred.html).

**Protein comparison and Pfam domain annotation**

To generate the most complete possible set of *T. pisana* peptides, we predicted protein-coding regions using OrfPredictor (Min, Butler et al. 2005) by default parameters on each tissue-specific assembly. We only retained the predicted longest ORFs and translated those into amino acid sequences over 30 amino acids. These sequences represented the full transcriptome-derived proteome. These were grouped into identical amino acid sequences using BLASTCLUST (BLAST score-based single-linkage clustering, [ftp://ftp.ncbi.nih.gov/blast/documents/blastclust.html]). We required the minimal length coverage threshold to be 0.7, which means that the minimum alignment length should cover at least 70% of full length of the shortest member in a group of sequences. For each resulting sequence clustering group, the longest amino acid sequence was chosen as the representative. The first optimal protein sequence in the sequence group was selected if multiple peptides had identical amino acid lengths. To provide an overview for the biological function, we annotated all the predicted proteins using the Pfam database (version 27.0). HMMSEARCH (Nachman, Pietrantonio et al. 2009) was adopted to associate proteins with Pfam domains. We used a threshold of 0.01 for the e-value to identify reliable hits.
**Neuropeptide prediction and sequence analysis**

Neuropeptides are generally secreted out of the cell, which is facilitated by signal peptides in the premature protein form. To systematically identify putative neuropeptides in the CNS of *T. pisana*, we initially utilized four bioinformatics tools on all putative CNS proteins to predict the presence of a signal peptide (SignalP 3.0) (Bendtsen, Nielsen et al. 2004) and PrediSi (Hiller, Grote et al. 2004)) and any transmembrane domain (TMHMM 2.0) (Krogh, Larsson et al. 2001) and (HMMTOP 2.1) (Tusnady and Simon 2001). For all these tools, we used default settings and parsed the results using in-house perl script. Then, resulting proteins were used as input to the NeuroPred tool to predict cleavage products. Schematic diagrams of protein domain structures were prepared using Domain Graph (DOG, version 2.0) software (Ren, Wen et al. 2009). Protein sequences from *T. pisana* were aligned against a database prepared from known sequences from NCBI (January, 2014) using the MEGA 5.1 (Tamura, Peterson et al. 2011) platform with the clustalW protocol utilising the Gonnet protein weight matrix. Neighbour-joining trees were generated based off these alignments. Unrooted trees were generated with 1000 bootstrap trials and presented with a cut-off bootstrapping value of 50.

**Tissue distribution of selected genes of interest**

Tissue was dissected out of 4 animals and each tissue type was pooled. Tissue types collected were tentacles, digestive tract, hepatopancreas, skin from head, mantle, foot muscle, cerebral ganglia and pleural ganglia. Total RNA was extracted from the 8 tissue types using TRIzol Reagent (Invitrogen, Catalogue # 15596-018) as per the manufacturer’s protocol. RNA extracts were assessed using agarose gel electrophoresis, and quantified using a Nanodrop 2000 spectrophotometer (Thermo scientific). First-strand cDNA was synthesised using QuantiTect Reverse Transcription kit (Qiagen) as per the manufacturer’s protocol. To normalise cDNA for each sample, equal quantities of extracted RNA were used as templates. Amplification of cDNA was carried out using the *Taq* F1 DNA Polymerase kit (Fisher Biotec) in a total volume of 25 µl. The PCR reagent mix was prepared as recommended by the manufacturer, containing 1x PCR reaction buffer, 2 mM MgCl₂, 0.2 mM each dNTPs, 0.2 µM forward primer and reverse primer (sequences available upon request), 1 unit of *Taq* F1 DNA polymerase, 17.9 µl water and 1 µl cDNA template. A negative control, substituting water for template, was included in all PCR experiments. PCR was performed in a Kyratec Model
SC200 thermal cycler. Cycling parameters were: initial denaturation at 95°C for 5 min, followed by 35 cycles of denaturation at 94°C for 20 s, annealing at 55°C for 30 s, and extension at 72°C for 1 min. Final extension was at 72°C for 10 min, then held at 10°C. Primers used were designed from sequences for *Theba pisana* 18s rRNA gene and 6 peptide precursor genes (available upon request). A positive control reaction using cDNA from mixed tissues was also performed for each PCR reaction.

**Peptide isolation from CNS and nanoHPLC, mass spectrometry LC-ESI-QTOF peptide identification**

CNS were dissected out and combined from active and aestivated (3 weeks without movement) mature *T. pisana*, then immediately frozen in liquid nitrogen prior to storage at -80°C until use. Frozen samples of CNS were ground to a powder under liquid nitrogen in a mortar, then quickly weighed and homogenized in extraction buffer (90% methanol, 9% glacial acetic acid in deionized water) in a 1:5 w:v ratio. Crude extracts were then sonicated with three pulses, 30 s each, and centrifuged for 20 min (16,000 xg, 4°C). Supernatant was collected and lyophilised.

The CNS extracts were analyzed by LC-MS/MS on a Shimadzu Prominance Nano HPLC (Japan) coupled to a Triple-ToF 5600 mass spectrometer (ABSCIEX, Canada) equipped with a nano electrospray ion source. Aliquots (6 µl) of each extract were injected onto a 50 mm x 300 µm C18 trap column (Agilent Technologies, Australia) at 30µL/min. The samples were de-salted on the trap column for 5 minutes using solvent A [0.1% formic acid (aq)] at 30µL/min. The trap column was then placed in-line with the analytical nano HPLC column, a 150 mm x 75 µm 300SBC18, 3.5 µm (Agilent Technologies) for mass spectrometry analysis. Peptide elution used a linear gradient of 1-40% solvent B [90:10 acetonitrile: 0.1% formic acid (aq)] over 35 min at 300 nl/minute flow rate, followed by a steeper gradient from 40% to 80% solvent B over 5 min. Solvent B was then held at 80% for 5 min to wash the column and then returned to 1% solvent B for equilibration prior to the next sample injection. The ionspray voltage was set to 2400V, declustering potential (DP) 100V, curtain gas flow 25, nebuliser gas 1 (GS1) 12 and interface heater at 150°C. The mass spectrometer acquired 500ms full scan TOF-MS data followed by 20 by 50 ms full scan product ion data in an Information Dependent Acquisition (IDA) mode. Full scan TOFMS data was acquired over the mass range 350-1800 and for product ion ms/ms 100-1800. Ions observed in the TOF-MS scan exceeding a threshold of 100 counts and a charge state of +2 to
+5 were set to trigger the acquisition of product ion, ms/ms spectra of the resultant 20 most intense ions. The data was acquired and processed using Analyst TF 1.5.1 software (ABSCIEX, Canada).

Proteins were identified by database searching using PEAKS v7.0 (BSI, Canada) against the protein database built from the CNS transcriptome. Search parameters were as follows: no enzyme was used; variable modifications included methionine oxidation, conversion of glutamine to pyroglutamic acid, deamidation of asparagine and amidation. Precursor mass error tolerance was set to 20 ppm and a fragment ion mass error tolerance was set to 0.05 Da. Maximum expectation value for accepting individual peptide ion scores [-10*Log(\(p\))] was set to \(\leq 0.05\), where \(p\) is the probability that the observed match is a random event. Proteins and their supporting peptides were obtained and analysed.
Additional Figures and Tables

**Figure S2.1.** Length distribution of unigenes derived from a combined transcriptome assembly of *Theba pisana* CNS, foot muscle and hepatopancreas.

**Figure S2.2.** Graphs showing relative transcript abundance of top 50 transcripts in each *Theba pisana* transcriptome.

**Figure S2.3.** Graph showing relative transcript abundance for *Theba pisana* neuropeptides identified.

**Figure S2.4.** LC-MS/MS spectra showing peptides matching FFamide, NKY, sCAP and insulin precursors.

**Figure S2.5.** *Theba pisana* insulin-like precursors annotated with features characteristic of insulin peptides. Yellow, signal peptide; blue, bioactive insulin; pink, cysteine residues; green, putative amidated glycine; red, cleavage sites.

**Figure S2.6.** *Theba pisana* prohormone convertases, PC1 and PC2. Underline, signal peptide; red, catalytic region; boxed, conserved catalytic residues.

Figure S2.1. Length distribution of unigenes derived from a combined transcriptome assembly of *Theba pisana* CNS, foot muscle and hepatopancreas.
Figure S2.2. Graphs showing transcript abundance of top 50 transcripts in each *Theba pisana* transcriptome. Numbers for each correspond to transcripts listed in Table S2.2.
Figure S2.3. Graph showing transcript abundance for *Theba pisana* neuropeptides identified.
Tpi-Ffamide-1

1. MKESSLFFVVI SMVALLVTR SHVTQDNIK DTVQQPQPLF GRGLMPNRN HLFFGKRGAI NSEVSLKELK SACAMFLY

81. EDVDSSTVDDEE V

Tpi-NKY-1

1. MVNTINTLLK LALLAFVPLC YSRPPALLGE DDMGYFLDQFSPRLYVFVE VPDLALLAL RAAHQQSQFP NALLDKRASS

81. ASEILGSALP KLDDDTVDVG MKR KMPFWQPQPL GYLPASVRAH NSFTGSTQGE NQGSSSNVFR Y
Tpi-sCAP-1

MEMSLQBVSI SMSLLVILVC SAAKADNLAF PPR... SCYLAPPRM FSQAK SETHAEPFGNC CGVGLKSELV IGHDGKEELR

PVTYNECC QGLRELADQK POGTTYSMCV PDLPPSLESN GPSSEVLKKL KTLMKK
Tpi-insulin2

1 MSNLISYMLS RACLFTLILT IVPHGVTVP SAQYCDIHNS VPFRGTYGV NLSNVELVC ESFKRPTSSR YIVKNDIINDLGSIAL

Figure S2.4. LC-MS/MS spectra showing peptides matching FFamide, NKY, sCAP and insulin precursors.
Tpi-insulin1a

MVRNVEYYTLTVALAINLAHOVQQGQRRTCSLVARPHPNGYGGERLAAQHSGNFFLLRRTYPHLFPMSKRSPVNESHISQSLDSLPWDGDESTRYPSKSFPPPTLEDNSYLLSLLDTSDELATDSDEQA DTSDEQADTSDEQADTSDEQADTSERPTASAIFINLEKRKRKRKLVCECCYAPCSFRIARYC

Tpi-insulin1b

MVRNVEYYTLTVALAINLAHOVQQGQRRTCSLVARPHPNGYGGERLAAQHSGNFFLLRRTYPHLFPMSKRSPVNESHISQSLDSLPWDGDESTRYPSKSFPPPTLEDNSYLLSLLDTSDELATDSDEQA DTSDEQADTSDEQADTSDEQADTSERPTASAIFINLEKRKRKRKLVCECCYAPCSFRIARYC

Tpi-insulin1c

MVRNVEYYTLTVALAINLAHOVQQGQRRTCSLVARPHPNGYGGERLAAQHSGNFFLLRRTYPHLFPMSKRSPVNESHISQSLDSLPWDGDESTRYPSKSFPPPTLEDNSYLLSLLDTSDELATDSDEQA DTSDEQADTSDEQADTSDEQADTSERPTASAIFINLEKRKRKRKLVCECCYAPCSFRIARYC

Tpi-insulin1d

MVRNVEYYTLTVALAINLAHOVQQGQRRTCSLVARPHPNGYGGERLAAQHSGNFFLLRRTYPHLFPMSKRSPVNESHISQSLDSLPWDGDESTRYPSKSFPPPTLEDNSYLLSLLDTSDELATDSDEQA DTSDEQADTSDEQADTSDEQADTSERPTASAIFINLEKRKRKRKLVCECCYAPCSFRIARYC

Tpi-insulin2

MSNLISYMLSRACLFTLILTIVPHTGPVTPSAQKYCDHNSVPHPRGGCQVNLNSMVELVESFSRTPTS SRYIVKAQTKILNDNLGIALNKDDASYLAKRATTGTTCEQYHACTFEELAQYCPLRGS

Tpi-insulin3

MSGNADHCAALTCLITTFLAVNQQQGCRSCDLSRPHPNQGGSMLAQVHENVGFMLRQAYPHF FPLKLSPVLGDEDHGGLFFPLLKNVLVSGDGYRGPAVGDTPFSLRRNSNAPPRKTFGLPGDI PQDQSENPIQRVLNKARNRSRSLVCECCYGCSTRRILASYC

Tpi-insulin4

MASLVKVCGLVAICVLVDVISQGMDTENFRFAITTTNFSRSLSATELLNAWTHECHRRCNYOLTW HVEIAHFDPYRIQQGRRRSSIEKPRNLTNILTTSDKVTONDNSQPAPKDPQFMPRNSAMSFL KDHSYKRKKDVISRECO问KABSWEEFAEQCGSHRRAERDSVCTYD

Figure S2.5. *Theba pisana* insulin-like precursors annotated with features characteristic of insulin peptides. Yellow, signal peptide; blue, bioactive insulin; pink, cysteine residues; green, putative amidated glycine; red, cleavage sites.
Figure S2.6. *Theba pisana* prohormone convertases, PC1 and PC2. Underline, signal peptide; red, catalytic region; boxed, conserved catalytic residues.
Chapter 3: Genes and associated peptides involved with aestivation in a land snail

Published in General and Comparative Endocrinology 2015, S0016-6480(15)30013-7

Abstract

Some animals can undergo a remarkable transition from active normal life to a dormant state called aestivation; entry into this hypometabolic state ensures that life continues even during long periods of environmental hardship. In this study, we aimed to identify those central nervous system (CNS) peptides that may regulate metabolic suppression leading to aestivation in land snails. Mass spectral-based neuropeptidome analysis of the CNS comparing active and aestivating states, revealed 19 differentially produced peptides; 2 were upregulated in active animals and 17 were upregulated in aestivated animals. Of those, the buccalin neuropeptide was further investigated since there is existing evidence in molluscs that buccalin modulates physiology by muscle contraction. The \textit{T. pisana} CNS contains two buccalin transcripts that encode precursor proteins that are capable of releasing numerous buccalin peptides. Of these, \textit{Tpi-buccalin-2} is most highly expressed within our CNS transcriptome derived from multiple metabolic states. No significant difference was observed at the level of gene expression levels for \textit{Tpi-buccalin-2} between active and aestivated animals, suggesting that regulation may reside at the level of post-translational control of peptide abundance. Spatial gene and peptide expression analysis of aestivated snail CNS demonstrated that buccalin-2 has widespread distribution within regions that control several physiological roles. In conclusion, we provide the first detailed molecular analysis of the peptides and associated genes that are related to hypometabolism in a gastropod snail known to undergo extended periods of aestivation.

\textbf{Key words:} Snail, \textit{Theba pisana}, peptides, neuropeptides, aestivation, central nervous system
Background

Aestivation refers to a state of metabolic rate suppression (a dormant or resting state) during the summer or dry season, a state that is normally entered during times of restricted water or food, often during periods of extreme high summer temperatures (Storey and Storey 2012). Both vertebrates and invertebrates use hypometabolism as a survival strategy; aestivating vertebrates include amphibians such as frogs (Hudson, Lonhienne et al. 2008), toads (Armentrout and Rose 1971) and salamanders (Barry and Shaffer 1994), and reptiles such as lizards (Christian, Webb et al. 2003), crocodiles (Firth, Christian et al. 2010) and snakes (Winne, Willson et al. 2006). Aestivation among invertebrates has been well studied in pulmonate land snails (Storey 2002) and also in some earthworms and insects (Bayley, Overgaard et al. 2010, Benoit 2010) but other states that are highly comparable to aestivation (e.g. diapause in insects, dauer/diapause in nematodes) are also well studied (Hahn and Denlinger 2011, Padilla and Ladage 2012).

One of the challenges that aestivating animals encounter is to conserve water within the body, and this can be partially achieved by the animal sealing itself away from the environment. Snails typically elevate themselves to avoid higher temperatures at ground level or seek shelter in places such as crevices or under logs (Storey 2002) while also sealing the operculum with an mucous epiphragm that greatly reduces water loss from the body. Regulated metabolic rate depression is also used to achieve a strong reduction in energy use that ensures longer lasting endogenous fuel reserves. Typically, aestivating animals show a reduction in metabolic rate to 5-40% of normal resting rate in the active state (Guppy and Withers 1999). The mechanisms controlling metabolic rate depression must also be quickly reversible, since some species are known to become active within as little as 5 minutes after environmental conditions improve. Metabolic rate depression involves strongly reducing ATP turnover by coordinated reductions in both fuel catabolism and the rates of many ATP-expensive metabolic functions (e.g. ion channels, gene expression, protein synthesis, cell division). For example, in the aestivating land snail Otala lactea, activity of the membrane enzyme Na⁺/K⁺-ATPase, one of the biggest energy consumers in the cell, is reduced by 30% (Storey and Storey 2004, Ramnanan and Storey 2006). Indeed, ion channel activity (Na⁺/K⁺-ATPase, Ca²⁺-calmodulin ATPase) and both initiation and elongation factors of protein synthesis (another major energy expenditure in cells) are strongly suppressed during aestivation via reversible protein phosphorylation (Ramnanan and Storey 2008, Ramnanan, Allan et al. 2009, Storey and Storey 2012). In addition, aestivation includes the specific
implementation of cell protection measures via the selective up-regulation of some genes (e.g. chaperone proteins, antioxidant defenses) (Storey and Storey 2012).

However, much remains to be learned about the gene expression and protein adaptation responses that support aestivation and/or regulate transitions to and from the aestivating state. To date, there have been very few gene/protein screening studies of aestivating species (Storey and Storey 2010, Storey and Storey 2012). Recent transcriptome sequencing projects have identified differential gene expression during aestivation in sea cucumbers (Apostichopus japonica) and green-striped burrowing frogs (Cyclorana alboguttata) (Du, Bao et al. 2012, Reilly, Schlipalius et al. 2013). In sea cucumbers, a comparison of transcripts from active and aestivating animals revealed 446 differentially expressed genes and, of these, 253 were down-regulated during aestivation and 193 were up-regulated; the functions of 34% of these genes are unknown (Du, Bao et al. 2012). In frogs, a comparison between active and 4-month estivating animals found differential up-regulation of genes associated with energy metabolism, antioxidant defence, cytoskeletal remodelling, and anti-apoptotic signalling during aestivation.

Peptides have been implicated in helicid snail hypometabolism. For instance, mass spectrometry has been used to determine changes in peptide profiles of the brain and haemolymph of the snail Helix pomatia during hibernation (Pirger, Lubics et al. 2010). The output of selected peptides was increased or decreased in response to low temperatures. In the brain, 19 peptides/polypeptides were predominantly present in the active state, with 10 in the haemolymph. Other peptide/polypeptides (11 in the brain and 13 in the haemolymph) were present only in hibernation, while several peptides showed no difference between states. With such a diverse range of biological functions that need to be altered during hypometabolism, it is clear that some global control is necessary to not only regulate but also coordinate these changes, and little is known about this control.

The small white Italian snail, Theba pisana, is one of several snails native to the Mediterranean that are now established as pests in grain crops, pastures and vineyards in southern Australia (Baker 1991, Baker 1998, Baker 2008). In Australia the snails are active during autumn, winter and spring, and aestivate in early summer as temperatures increase (Baker and Vogelzang 1988). The central nervous system (CNS), hepatopancreas and foot muscle transcriptomes of T. pisana have been analysed, providing a summary of peptides (Adamson, Wang et al. 2015). Of those, a total of 22 neuropeptides were found at significantly different levels in the CNS at activity compared to aestivation, including the small cardioactive peptide A (Adamson, Wang et al. 2015).
In this study, we have performed a comprehensive molecular investigation of the responses of the CNS peptidome during *T. pisana* aestivation to confirm and help clarify precisely which extracellular neuropeptides (and their genes) are associated with regulating entry into and/or arousal from aestivation. We analyzed the CNS peptidome of *T. pisana* during active and aestivating states, finding 19 differentially expressed peptides. One of those peptides upregulated the buccalin, was further investigated at the gene and peptide level through gene and protein expression analyses.

**Materials and Methods**

**Animals and tissue collection for proteome analysis**

Mature *Theba pisana* were collected from agricultural sites surrounding Warooka, located on the Yorke Peninsula, South Australia in early spring (September). Snails were transported to the University of the Sunshine Coast (USC) and housed within purpose-built enclosures. Snails were provided with water and food (cucumber and carrot) *ad libitum* and maintained at room temperature.

Active snails were conditioned by feeding and supplied with water daily over 7 days. Snails to be used for aestivation experiments were placed into glass jars without food or water and kept in an incubator on a cycle of 12 h at 30°C with light, 12 h at 20°C dark, to emulate South Australian summer conditions. The positions of the snails were marked on the jars after 14 days. Any snails that had not moved following a further 21 days were deemed to be in aestivation.

**Histology**

Histological analysis was assessed to determine the cell structure of the circumesophageal nerve ring central nervous system (CNS) minus the buccal ganglia. Tissue samples that had been fixed in 4% paraformaldehyde were further dehydrated in ethanol before being embedded in a paraffin wax. The samples were then sectioned in serial transverse cross sections using a rotary microtome and stained with Harris’s hematoxylin and eosin stains. The slides were permanently mounted using DePex (BDH Chemicals). The
sections were viewed and photographed with a light microscope (BX51; Olympus) equipped with a camera system (UC50; Olympus).

**Protein isolation from CNS and nanoHPLC - ESI-Triple TOF peptide identification and label-free quantification**

Both active moving and aestivated *T. pisana*, were killed by immersion in liquid nitrogen prior to CNS removal. CNS were then separately pooled before immediate re-freezing in liquid nitrogen prior to storage at -80°C until use. Frozen samples of CNS were ground to a powder under liquid nitrogen in a mortar, then quickly weighed while frozen and homogenized in extraction buffer (90% methanol, 9% glacial acetic acid in deionized water) in a 1:5 w:v ratio. Crude extracts were then sonicated with three pulses, 30 s each, and centrifuged for 20 min (16,000 x g, 4°C). Supernatant was collected and lyophilised.

The CNS extracts were analyzed by LC-MS/MS on a Shimadzu Prominance Nano HPLC (Japan) coupled to a Triple-ToF 5600 mass spectrometer (ABSCIEX, Canada) equipped with a nano electrospray ion source. Aliquots (6 µL) of each extract were injected onto a 50 mm x 300 µm C18 trap column (Agilent Technologies, Australia) at 30 µL/min. The samples were de-salted on the trap column for 5 minutes using solvent A [0.1% formic acid (aq)] at 30 µL/min. The trap column was then placed in-line with the analytical nano HPLC column, a 150 mm x 75 µm 300SBC18, 3.5 µm (Agilent Technologies) for mass spectrometry analysis. Peptide elution used a linear gradient of 1-40% solvent B [90:10 acetonitrile:0.1% formic acid (aq)] over 35 min at 300 nL/minute flow rate, followed by a steeper gradient from 40% to 80% solvent B over 5 min. Solvent B was then held at 80% for 5 min to wash the column and then returned to 1% solvent B for equilibration prior to the next sample injection. The ionspray voltage was set to 2400V, declustering potential (DP) 100V, curtain gas flow 25, nebuliser gas 1 (GS1) 12 and interface heater at 150°C. The mass spectrometer acquired 500 ms full scan TOF-MS data followed by 20 by 50 ms full scan product ion data in an Information Dependent Acquisition (IDA) mode. Full scan TOFMS data was acquired over the mass range 350-1800 and for product ion ms/ms 100-1800. Ions observed in the TOF-MS scan exceeding a threshold of 100 counts and a charge state of +2 to +5 were set to trigger the acquisition of product ion, ms/ms spectra of the resultant 20 most intense ions. The data was acquired and processed using Analyst TF 1.5.1 software (ABSCIEX, Canada).
Proteins were identified by database searching using PEAKS v7.0 (BSI, Canada) against the protein database built from the CNS transcriptome (Adamson, Wang et al. 2015). Search parameters were as follows: no enzyme was used; variable modifications included methionine oxidation, conversion of glutamine to pyroglutamic acid, deamidation of asparagine and amidation. Precursor mass error tolerance was set to 20 ppm and a fragment ion mass error tolerance was set to 0.05 Da. Maximum expectation value for accepting individual peptide ion scores [-10*Log(p)] was set to ≤0.01, where p is the probability that the observed match is a random event. Proteins and their supporting peptides were obtained and analysed. The quantitative analysis of proteins was carried out using the label-free quantification module (PEAKS Q) of PEAKS v7.0, and relative concentrations of proteins at two stages were compared. Biological triplicates [i.e., (A1, B1), (A2, B2) and (A3, B3)] of each stage were used in tandem repeats, and the average values were calculated as the final results. Extracted peptides were quantified based on absorbance at 280 nm using a NanoDrop spectrophotometer for later normalization. For each run, about 1.5 µg of the mix was then analyzed via liquid chromatography combined with mass spectrometry on a Triple-TOF. In addition to the protein identification method mentioned above, peptide feature fold change and protein unique peptide were set to 2, peptide feature significance filter was set to 0.01 and other parameters were adjusted according to the ‘volcano plot’ generated by PEAKS.

**Gene analysis**

To characterize which of the significantly up- and down-regulated CNS peptides identified from MS/MS corresponded to protein precursors, a BLAST analysis was performed against an in-house derived *Theba pisana* protein database (http://thebadb.bioinfo-minzhao.org/). Schematic diagrams of protein domain structures were prepared using Domain Graph (DOG, version 2.0) software (Ren, Wen et al. 2009). Buccalin-like sequences were derived by BLASTp analysis of the NCBI database, then protein sequences were aligned using the MEGA 5.1 (Tamura, Peterson et al. 2011) platform with the clustalW protocol utilising the Gonnet protein weight matrix. Neighbour-joining trees were generated based on these alignments. Gene expression levels were determined from RNA-seq data (NCBI Genbank under SRA file SRP056280) using the FPKM method, which represents the fragments per kilobase of transcript per million fragments mapped.
**Absolute qPCR**

RNA was extracted from tissue using TRIzol Reagent (Invitrogen), as per the manufacturer’s protocol. Each sample contained CNS from 3 *T. pisana*. A total of 15 aestivating and 14 active snail samples were prepared. Following extraction, RNA was assessed for quality by visualisation on a 1.2% denaturing formaldehyde agarose gel, and quantified using a Nanodrop spectrophotometer (Thermo Scientific). Approximately 500 ng of total RNA were reverse-transcribed using QuantiTect Reverse Transcription kit (Qiagen) as per the manufacturer’s protocol. Products were visualised by agarose gel electrophoresis, then cDNA was stored at -20°C until use.

Absolute qPCR was performed using a SensiFast HRM kit (Bioline) as per the manufacturer’s protocol, on a Rotor-gene 6000 cycler (Corbett research) using Rotor-gene 1.7.87 software. A total volume of 10 µl per reaction containing 1x SensiFast, and 400 nM of each primer, with 1 µl of template cDNA. All reactions, including a no template control (NTC) for each primer pair, were done in duplicate.

Cycling parameters were: an initial hold at 95°C for 2 min followed by 40 cycles of denaturation at 95°C for 5 s, annealing at 60°C for 10 s and extension at 72°C for 20 s acquiring to cycle A (Green). Melt curve analysis was performed rising by 1°C per cycle from 70°C to 95°C and was held for 90 s on the first step and 5 s on subsequent steps. Quantitative PCR primers for the gene of interest (*Tpi-buccalin-2*) were designed from transcriptome sequences using Primer3 software. Primers were sourced from Genworks and checked via PCR. A standard curve of known concentration was prepared for the gene of interest (*Tpi-buccalin*). Complementary DNA was amplified using the primers forward ACAGCCGAGCCCATATACAC, reverse GCGATGCTCCAGAATAAAGC and the PCR products were used as template for the standard curves. Following PCR the amplified products were visualised to check for correct size and then purified using the QIAquick kit (Qiagen). Purified products were quantified three times and mean concentrations calculated. From this and the known length of the products, the copy number was calculated using the calculator of URI Genomics & Sequencing Centre ([http://www.uri.edu/research/gsc/resources.cdna.html](http://www.uri.edu/research/gsc/resources.cdna.html)). A DNA concentration of 2 x 10⁹ was then prepared and serially diluted six times into 1:10 dilutions. A standard curve for the gene using the serially diluted PCR product was prepared (in triplicate) using the reaction conditions described above. No template control (NTC) was carried out in duplicate. Target gene expression levels of the CNS tissue from active and aestivating animals was quantified.
using the qPCR described above, samples were run in duplicate using cDNA synthesised from CNS RNA.

*In situ hybridization*

The *T. pisana* buccalin-2 open reading frame was cloned into a pUC plasmid (Genscript). Sense and antisense digoxigenin (DIG)-labelled riboprobes were prepared using a DIG RNA labelling mix kit (Roche) as per protocol (Cummins, Tollenaere et al. 2011), using SP6 and T7 polymerase (SP6 5’-TAATACGACTCACTATAGGG -3’, T7 5’-ATTTAGGTGACACTATAG3’). *T. pisana* CNS derived from active and aestivated snails was fixed in 4% paraformaldehyde overnight at 4°C and stored in 70% ethanol, before dehydration in ascending concentrations of ethyl alcohol for 30 min each, cleared in xylene three times, infiltrated, and embedded in paraffin. Serial transverse sections of the tissues were cut at 5 μm thickness using a microtome. Sections were then deparaffinized in xylene and rehydrated in a descending concentration of ethanol. Sections were then washed and rehydrated into phosphate buffered saline (PBS) with 0.1% Tween 20. Sections were pre-hybridised for 3 hours in prehybridisation solution [50% formamide, 5x sodium saline citrate, 5 mM EDTA, 1% Denhardt’s solution (Sigma), 100 µg/ml heparin, 100 µg/ml tRNA, 0.1% Tween20] at 55°C. Hybridisation was performed using the same solution but adding 200 ng/ml DIG-labelled riboprobe overnight at 42°C. Washing, detection and mounting for viewing was performed as described by Cummins et al. (Cummins, Tollenaere et al. 2011). Sections were viewed under a confocal laser-scanning microscope (Nikon).

*Immunolocalisation*

Rabbit polyclonal antibodies were generated to the Tpi-buccalin-2 precursor by Genscript using keyhole limpet hemocyanin-coupled peptides-RLDKGFSGGI-amide. CNS was isolated from active and aestivated *T. pisana* then fixed in 4% paraformaldehyde overnight at 4°C, before dehydration in ascending concentrations of ethyl alcohol for 30 min each, cleared in xylene three times, infiltrated, and embedded in paraffin. Serial transverse sections of the tissues were cut at 5 μm thickness using a microtome. Sections were then deparaffinized in xylene and rehydrated in a descending concentration of ethanol. Subsequently, sections were incubated in 0.1% glycine in 0.1 M PBS for 30 min, and washed three times with PBS with
0.1 M PBS containing 0.4% Triton X-100 (PBST). Non-specific binding was blocked in 4% normal goat serum in PBST for 2 h, followed by incubation in the primary antibody (anti-peptide) at an optimal dilution of 1:1000 in blocking solution, at 4°C overnight. Sections were then washed 3 times with PBST, and incubated for 2 h with Alexa Fluor 488 conjugated goat anti-rabbit IgG (Santa Cruz, USA) at room temperature. After washing with PBST, nuclei were stained with DAPI (Santa Cruz, USA) for 10 min. Finally, sections were washed with PBST and mounted with VECTA shield fluorescent mounting medium (Molecular Probes) before viewing under a confocal laser-scanning microscope (Nikon). In negative controls, tissues were processed by the same protocol, but preimmune mouse serum was used instead of primary antibody.

**Results and Discussion**

Previous investigation of the *T. pisana* CNS transcriptome revealed that this snail may contain as many as 5504 precursor proteins that are destined to secrete peptides. Of these, full-length precursors were identified in *T. pisana* for 35 known molluscan neuropeptides (Adamson, Wang et al. 2015). In this study, we were most interested in identifying which of those neuropeptides may be associated with the aestivation process.

**Histology of the Theba pisana CNS**

Prior to obtaining the aestivation-associated CNS peptides, it was important to explore the anatomical and histological make-up of the *T. pisana* CNS. To achieve this, the CNS was removed for light microscopic (Figure 3.1A) and hematoxylin/eosin histological examination (Figure 3.1B and Figure S3.1). The cerebral ganglia in *Theba* are located above the esophagus, and consist of a left and right ganglion, connected by a cerebral commissure. These ganglion are divided into 3 regions; the procerebrum, the mesocerebrum and the metacerebrum. Major nerve cords connect the cerebral ganglia to the subesophageal ganglia, including pairs of pedal, pleural and parietal ganglia, and a single visceral ganglion. This general organisation is consistent with that found in other land snails, including *Cornu* and *Achatina*, and the pulmonate freshwater snail *Lymnaea*, with a ring of ganglia circling the esophagus and inclusive of multiple ganglia (Chase and Toloćzkó 1993, Chase 2000). While the area of the body for which each ganglia control in molluscs are generally accepted, it
appears that there are often multiple ganglia involved in functions that regulate various metabolic events, such as breathing (Syed, Bulloch et al. 1990) and heart rate (Arshavsky, Deliagina et al. 1990). A summary schematic of the *T. pisana* CNS is shown in **Figure 3.1C**.

**Figure 3.1.** Investigation of the *T. pisana* CNS. (A) Whole-mount image of CNS. (B) Representative histological section of CNS with haematoxylin and eosin stain. (C) Schematic representation of CNS. CG, cerebral ganglia; CC, cerebral commissure; DBa, dorsal body area; LPaG, left parietal ganglia; RPaG, right parietal ganglia; PeG, pedal ganglia; LPlG, left pleural ganglia; RPlG, right pleural ganglia; PC, procerebrum; St, statocyst; TNv, tentacle nerve; MtC, metacerebrum; PCN, procerebrum.

*Identification of differentially expressed peptides in active and aestivated Theba pisana CNS*

Peptide extraction of the CNS followed by LC-MS analysis was used to identify those peptides present in the *T. pisana* CNS, then we proceeded to identify those differentially expressed between active and aestivating snails. A general workflow for this analysis is shown in **Figure 3.2**.
**Figure 2.** Workflow for the identification of aestivation-associated peptides. *T. pisana* CNS peptides from active and aestivated snails were extracted and purified by nanoHPLC-MS/MS for determination of differential expression of peptides. The protein database was provided from a CNS-derived transcriptome obtained from pooled RNA from active and aestivated *T. pisana* (Adamson, Wang et al. 2015).

To ensure that aestivated snails did not begin arousal, so that CNS samples accurately reflected the natural deep aestivation state, snails were rapidly killed by immersion in liquid nitrogen prior to CNS removal. Active snails were also rapidly killed with liquid nitrogen, an approach that had not been implemented in our prior analysis of aestivation peptides in the *T. pisana* CNS (Adamson, Wang et al. 2015). We identified 19 peptides that were differentially expressed, including 8 that contain a C-terminal amide (**Figure 3.3, Table 3.1 and File S3.1**).
Those CNS peptides that have not previously been characterised were termed Aestivation-associated peptides (AAPs).

![Hierarchical clustering of peptides relative expression levels in the Theba pisana CNS of aestivated versus active snails based on MS/MS analysis. The 19 differentially expressed peptides can be divided into two clusters. Names and further information about these peptides can be found in Table 3.1.](image)

Of the 19 peptides differentially expressed, only two peptides were up-regulated in active snails. One of these (AAP14) corresponds to a precursor protein that has not been previously identified in any species, while the other is Leu-Phe-Arg-Phe (LFRF). The remaining 17 peptides were up-regulated in aestivating snails, including the molluscan neuropeptides for *T. pisana* small cardioactive peptide (sCAP), feeding circuit-activating peptide (FCAP), enterin, pleurin, sensorin-A, a LASGLV-like peptide and buccalin.

*T. pisana* sCAP, FCAP, enterin, pleurin and sensorin-A have been investigated in some aquatic molluscs. The sCAP was first identified in the marine slug *Aplysia* where it appears to be highly concentrated in the buccal ganglia, implying an important role in feeding. Besides the buccal ganglia, it is also present in fibres and neuronal bodies throughout the CNS, suggesting its involvement in more widespread functions (Lloyd, Mahon et al. 1985). *T. pisana* contains 3 sCAP precursor isoforms, each containing an sCAPA peptide (Adamson, Wang et al. 2015). In our study, we found that Tpi-FCAP was exclusive to the CNS of
aestivating snails with none detected in active snails (Figure 3.3 and Table 3.1). FCAP was also initially identified in *Aplysia*, where it was implicated in the regulation of feeding. Similar to sCAP, this peptide is distributed throughout various areas of the CNS, suggesting multiple functions (Sweedler, Li et al. 2002). The peptide enterin, also recognised as a feeding-associated neuropeptide in *Aplysia*, is found throughout its CNS, indicating involvement in functions other than feeding (Furukawa, Nakamaru et al. 2001).

**Table 3.1.** List of peptides up- and down-regulated between active and aestivated *Theba pisana* CNS.

<table>
<thead>
<tr>
<th>Peptide name</th>
<th>Predicted avg mass</th>
<th>m/z</th>
<th>z (charge)</th>
<th>Aestivated</th>
<th>Active</th>
<th>Ratio</th>
<th>Sequence</th>
<th>Precursor length AA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tpi-FCAP</td>
<td>1312.41</td>
<td>656.82</td>
<td>2</td>
<td>2.60E+04</td>
<td>0</td>
<td>1.00:0</td>
<td>GLDSLGGSVHGW</td>
<td>173 [not FL]</td>
</tr>
<tr>
<td>Tpi-LFRF</td>
<td>1388.50</td>
<td>694.837</td>
<td>2</td>
<td>3.06E+04</td>
<td>2.57E+03</td>
<td>1.00:0.08</td>
<td>GLDSLGGYQVHG</td>
<td>155 [not FL]</td>
</tr>
<tr>
<td>Tpi-buccalin-1</td>
<td>1465.42</td>
<td>733.27</td>
<td>2</td>
<td>0</td>
<td>3.68E+03</td>
<td>00:1.00:00</td>
<td>SDSAGNPMDNEEE</td>
<td>265</td>
</tr>
<tr>
<td>Tpi-buccalin-2</td>
<td>1154.29</td>
<td>577.80</td>
<td>2</td>
<td>2.69E+04</td>
<td>9.98E+02</td>
<td>1.00:0.04</td>
<td>RLDSYFGAGGI</td>
<td>265</td>
</tr>
<tr>
<td>Tpi-enterin</td>
<td>1165.36</td>
<td>583.33</td>
<td>2</td>
<td>6.29E+04</td>
<td>1.01E+04</td>
<td>1.00:1.6</td>
<td>RVDKFGFAGGI</td>
<td>277</td>
</tr>
<tr>
<td>Tpi-sCAP</td>
<td>1195.39</td>
<td>598.33</td>
<td>2</td>
<td>2.68E+04</td>
<td>3.75E+03</td>
<td>1.00:1.14</td>
<td>RLDKGFSGGI</td>
<td>277</td>
</tr>
<tr>
<td>Tpi-pleurin</td>
<td>1040.25</td>
<td>488.70</td>
<td>2</td>
<td>6.35E+03</td>
<td>6.49E+02</td>
<td>1.00:1.10</td>
<td>GPNFQHSPV-amide(-.98)</td>
<td>279 [not FL]</td>
</tr>
<tr>
<td>Tpi-sCAP</td>
<td>1474.64</td>
<td>492.26</td>
<td>3</td>
<td>8.78E+03</td>
<td>1.57E+02</td>
<td>1.00:0.03</td>
<td>GVFQGAHGSPV-amide(-.98)</td>
<td>136</td>
</tr>
<tr>
<td>Tpi-pleurin</td>
<td>1497.78</td>
<td>797.91</td>
<td>2</td>
<td>4.01E+03</td>
<td>1.43E+03</td>
<td>1.00:0.36</td>
<td>TFFYGGNGIHYPR-amide(-.98)</td>
<td>179</td>
</tr>
<tr>
<td>Tpi-sensorin-A</td>
<td>1113.38</td>
<td>567.30</td>
<td>2</td>
<td>7.62E+03</td>
<td>0</td>
<td>1.00:0.0</td>
<td>AKYRVGYMF-amide(-.98)</td>
<td>110</td>
</tr>
<tr>
<td>Tpi-LASGLV_like</td>
<td>977.46</td>
<td>489.24</td>
<td>2</td>
<td>6.31E+03</td>
<td>0</td>
<td>1.00:0.1</td>
<td>RPFDDELGSG</td>
<td>368 [not FL]</td>
</tr>
<tr>
<td>Tpi-AAP1</td>
<td>980.09</td>
<td>490.27</td>
<td>2</td>
<td>4.49E+03</td>
<td>0</td>
<td>1.00:0.0</td>
<td>GSQSSFVRI-amide(-.98)</td>
<td>408</td>
</tr>
<tr>
<td>Tpi-AAP2</td>
<td>2269.58</td>
<td>757.07</td>
<td>3</td>
<td>5.21E+03</td>
<td>0</td>
<td>1.00:0.0</td>
<td>SLEAALRAPPSYSEALIEAPA</td>
<td>108</td>
</tr>
<tr>
<td>Tpi-AAP6</td>
<td>1699.88</td>
<td>567.29</td>
<td>3</td>
<td>3.52E+03</td>
<td>1.53E+03</td>
<td>1.00:1.14</td>
<td>EEAASVKKETIHTEK</td>
<td>689 [not FL]</td>
</tr>
<tr>
<td>Tpi-AAP4</td>
<td>1106.31</td>
<td>369.20</td>
<td>3</td>
<td>4.41E+03</td>
<td>3.04E+01</td>
<td>1.00:0.1</td>
<td>RMHNFVRF-amide(-.98)</td>
<td>234</td>
</tr>
<tr>
<td>Tpi-AAP5</td>
<td>1117.32</td>
<td>373.20</td>
<td>3</td>
<td>7.40E+03</td>
<td>3.27E+02</td>
<td>1.00:0.04</td>
<td>AHHIGLGTALK</td>
<td>384 [not FL]</td>
</tr>
<tr>
<td>Tpi-AAP3</td>
<td>1936.02</td>
<td>645.97</td>
<td>3</td>
<td>5.57E+03</td>
<td>3.76E+02</td>
<td>1.00:0.07</td>
<td>FFYGLTPRSSNPGAD</td>
<td>486</td>
</tr>
<tr>
<td>Tpi-AAP8</td>
<td>1494.68</td>
<td>747.85</td>
<td>2</td>
<td>1.43E+05</td>
<td>2.99E+04</td>
<td>1.00:0.21</td>
<td>IMDSLSSATVTVR</td>
<td>103 [not FL]</td>
</tr>
<tr>
<td>Tpi-AAP7</td>
<td>1496.6</td>
<td>748.37</td>
<td>2</td>
<td>4.42E+03</td>
<td>1.11E+03</td>
<td>1.00:0.25</td>
<td>RFDSSHGSSFGSGL</td>
<td>431 [not FL]</td>
</tr>
<tr>
<td>Tpi-AAP9</td>
<td>1172.41</td>
<td>396.89</td>
<td>3</td>
<td>6.31E+02</td>
<td>0</td>
<td>1.00:0.0</td>
<td>GVM(+15.99)GKSAGQLPK</td>
<td>205</td>
</tr>
<tr>
<td>Tpi-AAP14</td>
<td>1863.95</td>
<td>621.98</td>
<td>3</td>
<td>4.40E+02</td>
<td>1.90E+03</td>
<td>1.00:4.31</td>
<td>AKAVSGDAVESGSKTEDVN</td>
<td>211 [not FL]</td>
</tr>
</tbody>
</table>

m/z, mass to charge; FL, full-length
The peptide pleurin has previously been identified in *Lottia* and in *Aplysia*, where the precursor is found in the right pleural ganglion (Moroz, Edwards et al. 2006, Veenstra 2010). The sensorin-A peptide has also been found in the cerebral and buccal ganglia of *Aplysia* (Brunet, Shapiro et al. 1991), and in one neuron of each pedal ganglion of *Lymnaea* (Steffensen, Syed et al. 1995). An extensive *in silico* analysis of *Pinctata fucata* and *Crassostrea gigas* failed to identify pleurin and sensorin homologs, suggesting they have been lost in bivalve molluscs (Stewart, Favrel et al. 2014). Furthermore, we identified a peptide cleaved from the LFRF precursor that appears to be down-regulated in the *T. pisana* CNS during aestivation (Table 3.1); this peptide precursor contains the sequence SDSAQNPMMDNEEE and includes 6 LFRF-like peptides. This peptide is not present in other mollusc LFRF precursors and has not been functionally characterised. A LASGLV-like precursor with 13 predicted cleavage peptides, including 6 copies of RPFDELGSG which is up-regulated in aestivation is highly conserved in *Aplysia californica*, within a predicted buccalin-like peptide precursor, however there is little overall similarity to buccalin. The majority of the remaining peptides identified as differentially expressed, besides buccalin (Figure 3.3 and Table 3.1), match to precursor sequences that resemble either uncharacterised peptides or unknown proteins.

We identified 5 different buccalin peptides that are cleaved from two *T. pisana* buccalin precursor proteins and are up-regulated during snail aestivation (Figure 3.3 and Table 3.1), corresponding to Tpi-buccalin-1 and Tpi-buccalin-2. Our initial *in silico* screen for neuropeptides within the *T. pisana* CNS had not revealed these buccalin precursors (Adamson, Wang et al. 2015), possibly due to divergence with other species buccalin. The neuropeptide buccalin was first described in the buccal ganglia of *Aplysia* where it was shown to decrease the number of muscle contractions in the accessory radula closer, used during biting (Cropper, Miller et al. 1988). Since then, the closely related neuropeptides, buccalin B and buccalin C, were noted to be more effective (2-3x) at depressing radula contractions (Weiss, Cropper et al. 1988, Miller, Beushausen et al. 1993, Vilim, Cropper et al. 1994). Similar to Tpi-buccalins, the *Aplysia* buccalins contain multiple identical peptide repeats; for example, the buccalin C peptide is cleaved from a precursor containing a total of 19 buccalin-related peptides (Miller, Beushausen et al. 1993). There is evidence that buccalin is not only involved in *Aplysia* feeding, but also in functions controlled by the central ganglia, including cerebral and pleural ganglia (Raymond, Schulman et al. 1989, Rosen, Susswein et al. 1989, Miller, Alevizos et al. 1992). In other molluscs, little functional knowledge has been obtained for buccalin, although we do know that in *L. stagnalis*, it is primarily found in the
pedal and buccal ganglia, ganglia that regulate feeding and locomotion. Smaller numbers of buccalin neurons were found in the cerebral, right parietal and visceral ganglia (Santama, Wheeler et al. 1994).

Besides *Aplysia* and *Lymnaea*, buccalin precursor genes have been found in a variety of other molluscs through *in silico* genome or transcriptome mining, including those from *Lottia gigantea* (Veenstra 2010), *C. gigas* (Stewart, Favrel et al. 2014), and *Biomphalaria glabrata* (Lockyer, Spinks et al. 2007). Such *in silico* analyses can only predict that a bioactive peptide is produced, yet alternative splicing events may produce overlapping sets of peptides (Buck, Bigelow et al. 1987, Weiss, Bayley et al. 1989). In the pelagic sea slug *Clione limacine*, buccalin was detected in all central ganglia except the pleural ganglia using an *Aplysia* buccalin A antibody (Norekian and Satterlie 1997).

Since buccalin has an obvious role in the regulation of muscle contraction in the aquatic molluscs, as well as noted widespread distribution throughout the CNS, we speculated that the buccalin peptide could regulate metabolic events in aestivating snails. For that reason, Tpi-buccalin was further investigated via transcriptome screening and comparative sequence analysis, followed by spatial expression analysis. Of the buccalins identified, the *Tpi-buccalin-2* transcript and associated up-regulated peptide was targeted since its relative expression profile in the CNS far exceeded that of other buccalin isoforms (Adamson, Wang et al. 2015). In fact, *Tpi-buccalin-2* appears to be one of the most highly expressed transcripts within the CNS (combined active and aestivated adult animals; FPKM=2785), with expression relatively low in hepatopancreas (FPKM=1.59) and muscle (FPKM=8.3). *Tpi-buccalin-1* had relatively low levels of expression in all tissues (FPKM <10) (Table 3.2).
Table 3.2. List of abundance of differentially expressed gene transcripts of peptides in CNS, hepatopancreas and muscle tissue in *Theba pisana*.

<table>
<thead>
<tr>
<th>Transcripts</th>
<th>FPKM Rank</th>
<th>FPKM Rank</th>
<th>FPKM Rank</th>
<th>FPKM Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tpi-FCAP</td>
<td>2</td>
<td>GLDSLGGSQVHGW, GLDSLGGSQVHGW</td>
<td>24</td>
<td>5282</td>
</tr>
<tr>
<td>Tpi-LFRF</td>
<td>1</td>
<td>SDAGNPMDSNEE</td>
<td>454</td>
<td>268</td>
</tr>
<tr>
<td>Tpi-buccalin-1</td>
<td>1</td>
<td>RLDKFGFAGGI</td>
<td>2</td>
<td>74800</td>
</tr>
<tr>
<td>Tpi-buccalin-4</td>
<td>5</td>
<td>RVDKFGF, RVDKFGFAGGI, RLDKFGFAGGI, SRDNGLSGLGSL, ELDPYGFSAR(-98)</td>
<td>2785</td>
<td>14</td>
</tr>
<tr>
<td>Tpi-enterin</td>
<td>1</td>
<td>GPMFGHFSV(-98)</td>
<td>679</td>
<td>174</td>
</tr>
<tr>
<td>Tpi-αCAP</td>
<td>1</td>
<td>SGYIAFPRM(-98)</td>
<td>732</td>
<td>157</td>
</tr>
<tr>
<td>Tpi-pleurin</td>
<td>3</td>
<td>GVFTQGAHGSYPRV(-98), GVFTQGAHGSYPRV(-98), TFYTGGNSHYPRV(-98)</td>
<td>727</td>
<td>159</td>
</tr>
<tr>
<td>Tpi-sensorin-A</td>
<td>1</td>
<td>AKRVQVRM(-98)</td>
<td>55</td>
<td>2244</td>
</tr>
<tr>
<td>Tpi-LASGLV like</td>
<td>1</td>
<td>RPFDELSG</td>
<td>108</td>
<td>1134</td>
</tr>
<tr>
<td>Tpi-AAP1</td>
<td>1</td>
<td>GSQSSFKVRI(-98)</td>
<td>54</td>
<td>2264</td>
</tr>
<tr>
<td>Tpi-AAP2</td>
<td>1</td>
<td>SLEAARKPSISYEALAPA</td>
<td>470</td>
<td>252</td>
</tr>
<tr>
<td>Tpi-AAP6</td>
<td>1</td>
<td>IEAASVKEHHTEX</td>
<td>0.13</td>
<td>196410</td>
</tr>
<tr>
<td>Tpi-AAP4</td>
<td>1</td>
<td>PNNFVPRVRI(-98)</td>
<td>67</td>
<td>1823</td>
</tr>
<tr>
<td>Tpi-AAP5</td>
<td>1</td>
<td>AMHGLTALK</td>
<td>2.7</td>
<td>55715</td>
</tr>
<tr>
<td>Tpi-AAP3</td>
<td>1</td>
<td>FHYGTKPEFSGMPGAD</td>
<td>6</td>
<td>24113</td>
</tr>
<tr>
<td>Tpi-AAP6</td>
<td>1</td>
<td>IMDSLSSADTVVR</td>
<td>2</td>
<td>75851</td>
</tr>
<tr>
<td>Tpi-AAP7</td>
<td>1</td>
<td>RDFSISGHSFGL</td>
<td>54</td>
<td>2279</td>
</tr>
<tr>
<td>Tpi-AAP9</td>
<td>2</td>
<td>GVM(+15.99)GKSGAGQLPK, GVMKGSAGQLPK</td>
<td>569</td>
<td>211</td>
</tr>
<tr>
<td>Tpi-AAP14</td>
<td>1</td>
<td>AKAVSGDAVESGSKTEDVN</td>
<td>18</td>
<td>7181</td>
</tr>
</tbody>
</table>

Schematics are presented in Figure 3.4A showing the organisation of Tpi-buccalin precursors and their comparison with known buccalin precursors from other species. Some features of interest include; Tpi-buccalin-like repeats vary from 10-11, depending on the precursor. From both buccalin precursors there are a total of 19 unique buccalin-like peptides and of these only 3 have more than 1 copy per precursor (Table 3.1). Other molluscan buccalin precursors contain a similar general organisation, although *L. gigantea* and *A. californica* show a greater number of buccalin-like peptides (Figure 3.4A).

A multiple sequence alignment of the four Tpi-buccalin precursors shows conservation between Tpi-buccalin-1 and Tpi-buccalin-2 within the predicted active peptides (Figure 3.4B). Based on our MS/MS data, only 1 of the buccalins showed post-translational amidation, ELDPYGFSARImide, however peptide amidation has been shown in most cases to be essential for bioactivity (Merkler 1994, Clarke, Sari et al. 2008). Phylogenetic tree analysis demonstrates that Tpi-buccalin-1 clusters most closely to the identified buccalin from *Cornu aspersum* (Figure 3.4C). Helicid buccalins appear to cluster more closely with bivalve buccalin, rather than the other gastropods represented, *L. gigantea* and *A. californica*.
Figure 3.4. Analysis of the *Theba pisana* buccalins. (A) Schematic representation of known buccalin precursor organisation in molluscs compared with a buccalin-like precursor of *Stegodyphus mimosarum*. (B) Multiple sequence alignment of Tpi-buccalin precursors. Blue shading shows amino acids that are highly conserved. Sequence log above alignment shows amino acid conservation, represented by letter size. MS/MS peptides identified are shown as red overline for buccalin-1 and green underline for buccalin-2. (C) Phylogenetic tree of mollusc buccalin precursors. *S. mimosarum* was used as an outgroup. Bootstrap values are shown and scale bar represents amino acid differences. All sequences used for analysis can be found in File S3.1. (D) Quantitative PCR expression analysis of Tpi-buccalin-2 in CNS of active and aestivated *T. pisana*. Absolute qPCR was performed on CNS samples from active and aestivated snails to determine relative gene expression of the Tpi-buccalin-2 precursor. While there was a trend for this gene to be up-regulated in CNS of aestivated snails, this was not statistically significant [Student t-test P(T<=t) two-tail 0.064894; Figure 3.4D]. This lack of differential gene expression indicates that the bioactive form of this neuropeptide is up-regulated in aestivation through some means other than an increase in gene expression, possibly via some...
form of post-translational modification of the peptide (Mann and Jensen 2003). Alternatively, buccalins may simply accumulate and be stored within neuron secretory vesicles during aestivation but not utilized until the snail returns to normal activity. For example, in gastropods there are neuropeptides known that are produced and stored within cell secretory vesicles, as demonstrated in Aplysia bag cell neurons (Fisher, Sossin et al. 1988).

The neuropeptide egg-laying hormone (ELH) is produced during periods of inactivity, in preparation for reproduction-associated egg laying. Upon stimulation of the bag cells via another neuropeptide, ELH is secreted into the hemolymph. Also, several bioactive peptides appear at greater immunoreactive intensity within the cerebral ganglia of hibernating C. apersa, likely a result of accumulation during inactivity (Bernocchi, Vignola et al. 1998). Similarly, the bucculin precursor is likely transcribed and post-translationally processed into its multiple peptides and stored within secretory vesicles, awaiting an unknown trigger for release. This could be beneficial considering the rapid speed with which Theba reanimates to full activity (~ 5 min, personal observation) and also to help conserve energy that would otherwise be required for protein synthesis. A role in modulation of acetylcholine release would ultimately impact on muscle excitability.

Other possibilities could include increased stability (ie. decreased turn-over) of the mRNA under aestivating conditions or a mechanism that confers preferential translation of the transcript versus the majority of transcripts that are suppressed either by (a) reduced gene transcription or (b) mRNA storage into stress granules during hypometabolism, or (c) a sequence that allows preferential translation by ribosomes during hypometabolism. For example, end-dependent translation can be strongly suppressed in hypometabolic or stressed states but mRNAs that have internal ribosome entry sites are translated when translation of most mRNAs is repressed (Johannes and Sarnow 1998, Hellen and Sarnow 2001) – e.g. most chaperones have this as well as various transcription factors that mediate the “stress response”.

Temporal and spatial expression of Tpi-buccalin-2 in the CNS

In situ hybridisation was employed to investigate spatial gene expression in the CNS tissue using a Tpi-buccalin-2-specific DIG-labeled riboprobe. Since we had identified no significant difference in expression of mRNA levels between active and aestivated snails using qPCR, we analysed gene expression at only one metabolic state, aestivation. Negative
controls using a sense riboprobe indicated little or no staining throughout the CNS (Figure 3.5A). This result implied an absence of non-specific binding in the target tissue. Antisense riboprobes showed staining in several areas of the CNS, including the cerebral ganglia (Figure 3.5B, C), and the cerebral-pleural connective nerve (Figure 3.5B, D). Two areas of expression appeared specifically within nerve fibres (Figure 3.5B, E, F). A high level of gene expression was observed in the region of the visceral ganglia (Figure 3.5B, G).

The Tpi-buccalin-2 peptide R\textsubscript{102}LDKFGFSGGI-amide was identified as being up-regulated during aestivation. This buccalin is not present in the other isoforms, although it may have similarity at the both the primary and structural level to some other buccalin peptides. A polyclonal antibody was generated to this peptide for spatial immunolocalization within the CNS of aestivated *T. pisana*. We found that this neuropeptide was widely distributed throughout the CNS of aestivated and active snails (Figure 3.6). Expression can be observed within regions of the cerebral ganglia, pleural ganglia, parietal ganglia, pedal ganglia and visceral ganglia (Figure 3.6A-L). Within the cerebral ganglia, immunopositive staining is clearly observed within the metacerebrum and cerebral-pleural connective fibre (Figure 3.6A,B,D,E,F,K). Individual neurons were observed within the regions of pleural-parietal-pedal (Figure 3.6A,C,G,I) and in close association (Figure 3.6J,L). Two individual immunoreactive neurons can be seen in the visceral ganglia region of the CNS (Figure 3.6G,H). No significant differences in spatial peptide expression were observed between active and aestivated CNS using this method (Figure S3.2, S3.3). Isotopic hybridization methods are generally considered to be more sensitive for detecting expression, thus would be more amenable to quantitative spatial analysis.
Figure 3.5. In situ hybridisation analysis of *Theba pisana* buccalin-2 in the aestivated CNS. (A) Control using sense DIG-labelled *buccalin*-2 riboprobe. (B) Section showing positive staining using antisense DIG-labelled *buccalin*-2 riboprobe. Regions boxed as (C-G) are shown magnified. CG, cerebral ganglia; DBa, dorsal body area; meso, mesocerebrum; meta, metacerebrum; pro, procerebrum; PeG, pedal ganglia; PlG, pleural ganglia; VG, visceral ganglia.
Figure 3.6. Immunolocalisation analysis of *Theba pisana* buccalin R_{102}LDKFGSGGI-amide. (A-L) Localisation of peptide within sections of the aestivated CNS. PaG, parietal ganglia; PeG, pedal ganglia; PlG, pleural ganglia Further immunolocalisation is shown in Figure S3.2-S3.3.

High levels of Tpi-buccalin in the cerebral ganglia, and the subesophageal ganglia are consistent with the widespread buccalin distribution observed in other molluscs (Raymond, Schulman et al. 1989, Rosen, Susswein et al. 1989, Miller, Alevizos et al. 1992). Based on the spatial expression of Tpi-buccalin-2 transcript and peptide within regions of the CNS, we may speculate upon certain physiological functions it may regulate. Within the cerebral ganglia of snails, the procerebrum region controls olfactory function, the mesocerebrum controls reproduction and the metacerebrum regulates motor actions (Chase and Tolloczko 1993). Within the pleural-parietal-visceral-pedal, the pairs of pedal control the foot, the pleural controls the mantle, parietal controls the pallial cavity, and a single visceral ganglia
regulates the visceral organs (Chase and Tolloczko 1993). Buccalin expression within these various ganglia implicates it in activities controlled by those ganglia. In the land snails studied, the visceral ganglia is known to regulate the heart (Chase 2000), thus, the presence of buccalin within this region is consistent with a functional role in controlling heart rate, which is known to slow dramatically during aestivation (Storey and Storey 1990). Similarly, relatively high concentrations of several bioactive peptides in the mesocerebrum of hibernating C. aspersum, could be regulating mating and avoidance behaviour upon reactivation (Bernocchi, Vignola et al. 1998). Therefore, the widespread distribution of Tpi-buccalin probably reflects the diversity of functions of buccalin, including the regulation of muscle contractions in the heart (visceral ganglia) and in breathing (pleural ganglia). Further work would be required to determine exactly how many buccalin-expressing neurons are present and their location within the various ganglia. In addition, it would be of interest to precisely assess buccalin concentration within individual neurons during aestivation and normal activity.

**Conclusions**

In summary, we have performed a comprehensive molecular investigation of CNS peptides associated with aestivation by comparative mass spectral analysis, followed by spatial gene and peptide expression of the Tpi-buccalin. We speculate that increased levels of Tpi-buccalin, and other peptides identified during aestivation, relate to accumulation within CNS neurons. When released, this peptide may modify the animal’s physiological response (e.g. muscle contraction strength or rate) that occurs in the aestivating to active transition (or is needed in a sustained manner in the active state). This work sets a foundation to clarify precisely what are the genes and extracellular peptides that trigger hypometabolism and maintains aestivation in land snails, and possibly other molluscs.

**Abbreviations**

CNS: central nervous system; FPKM: Fragments per kilobase of exon region in a given gene per million mapped fragments; GDH: glutamate dehydrogenase; GO: gene ontology; MS: mass spectrometry; FCAP: feeding circuit-activating peptide; sCAP: small cardioactive peptide.
Competing interests
The authors declare that they have no competing interests.

Author’s contributions
KJA carried out the experimental analysis, constructed figures, tables and drafted the manuscript. BAR prepared immunocytochemical tissues. TW carried out mass spectral proteome work. TK helped to perform the histological reconstruction of the animal CNS. AVK, KBS and SFC conceived the idea and obtained funding for the experiments and drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements
This work was supported by grants from the Australian Research Council (KBS, SFC) and the Grains Research Development Corporation (KA). We thank Dr Alun Jones (Institute for Molecular Bioscience, the University of Queensland) for advice and assistance with the tandem mass spectrometry. This research was undertaken with the assistance of resources from the National Computational Infrastructure (NCI), which is supported by the Australian Government.
Additional Figures and Tables

Figure S3.1. Histological sections through the *T. pisana* CNS and stained haematoxylin and eosin. Serial sections are numbered 1-20.
Figure S3.2. Immunolocalisation analysis of aestivating *T. pisana* buccalin R102LDKFGSGGI-amide. Serial sections are numbered 1-24.
Figure S3.3. Immunolocalisation analysis of active *T. pisana* buccalin R102LDKFGSGGI-amide. Serial sections are numbered 1-11.
<table>
<thead>
<tr>
<th>Table S3.1. List of sequences used for analysis.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tpi-AAP1</strong> CL16802.Contig2_All</td>
</tr>
<tr>
<td>MTSSKHAVLSTGLLLASLASSATESLIGNPSNPAKIANRADRSDPDRSKRMTDFAPLQGVDPVFKQGSIGFSETFPNVAADVET GKYLDLEYDHDLKAERAPPRRRNNFYIRGDRGDTNDSQMYSNDDETDKRASQMFRIGKLSSGFIVRIGKEISIESSNDNNVKRASSFVRI GRPSPHVRIRGPNPSKRPRDFLASEGPIKSRRSDGSHPFVAKRPDELNGITIGVKGPRDELGIGIVPGKPRDFE LENNKRRAFVFRIGKSTGTEDEINATKRASSFVIRGKSLKDSTDNFNSQKRASSFVRIGKSSLSQDNSALVSPADSSKLSESFWDETL EDNZPKNYKTVRGSAPFRIGKPSSAPFVRIGKDNFLDFAEGNRNTGRSQSSSVFRIGK</td>
</tr>
<tr>
<td><strong>Tpi-AAP2</strong> Unigene50087_All</td>
</tr>
<tr>
<td>MPSIAVCFPLVFAILHAVAPAESEGSSLAGASPRHTGLLRLKRSLLEAALRAPPSIYSEALIEAPAKRDGDGYIWMPAIGHYMPMP TDDAPEQGGGRDSGMNLIYRG</td>
</tr>
<tr>
<td><strong>Tpi-LASGLV_like</strong> Unigene5024_All</td>
</tr>
<tr>
<td>RNLFVKQGDHRTPGVLSSLKIISVSLLVISSIPSSPDCQSRNMRNTWETYMTKLALLSAMFCLGVSLGVSASEASRTLKRYSS ARNHHRVSAWSNSFKRPRDFLASEGPIKSRRSDGSHPFVAKRPDELNGITIGVKGPRDELGIGIVPGKPRDFE LENNKRRAFVFRIGKSTGTEDEINATKRASSFVIRGKSLKDSTDNFNSQKRASSFVRIGKSSLSQDNSALVSPADSSKLSESFWDETL EDNZPKNYKTVRGSAPFRIGKPSSAPFVRIGKDNFLDFAEGNRNTGRSQSSSVFRIGK</td>
</tr>
<tr>
<td><strong>Tpi-AAP5</strong> Unigene90516_All</td>
</tr>
<tr>
<td>MKFLPVKPEQOPFLYPDNGKSLMDKTDVIRLKNHANASFRHITCTGTVKSLEQHIDSGQESNSKNEIFGLLELRENELKQIRDGTLLAK RAKKSKTKEQNIAFFEMKMYKMYNEREPLLDDNYHEMTLMKLKRMEEASAVSKKEHEKTEKPFGMELQOKLSEKDSTLLLPCG YPSAPPGFLNPSLIRLHVPQAPQTELKLDTPYKPELETSQKLAELRIKQGTSGDYEDVDSALIRSMRTPYDEHSLRLEVELT SDDEEKRKDHLNDGHEPESLNNKRSDTAFETVRTPTVQKGRKDPLSEDVNAVMSQQLDLASSMLAKIQOLHKPSKS ESELEDKSIVKLTATNSVYTRHQPAPPRVAKLISPLENVNEVASAPRSRPTVPSTPLFNRILFVTTEFHPFKJNCPNCTPLEK PKW TSDNFNMICLKLRIADWDGKVNVDMIALYYTEEGELLESVTADVERTLVAVLNNHASPCSVTEQKFKILSALKALTITFEKE AVVEMMVFQLABILITYTDKSIWALDMNNMLGVDHRPQFQKELSDWDWWEAEALRNKHLMCSEWLEWMWTSYKRLQDSVLL NHHGHSLOQRSRMKSVGGLSKRGKAPARAEASALPQSTKVSSPYTFADVRTSRQDSSATHGSMTVMINNLHSTLSKASYSVDD</td>
</tr>
<tr>
<td><strong>Tpi-AAP4</strong> Unigene90869_All</td>
</tr>
</tbody>
</table>
| MTSSAIALLAPVLFCHVKWADQPSIQMLTSSPGELEKDALA KAIRKDREEISPSYSIUNSWSIEDALRFLDEOPVERTPRESSSDD SSDRRIRKEIKNLPIQTHFNFIRFGKRFHQLDIPSTSISLEDSDTSGGLKLRSNRLKRTDSYKPKETPSSDLRAIRG 

**Table S3.1.** List of sequences used for analysis.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tpi-AAP1</td>
<td>CL16802.Contig2_All</td>
</tr>
<tr>
<td>Tpi-AAP2</td>
<td>Unigene50087_All</td>
</tr>
<tr>
<td>Tpi-LASGLV_like</td>
<td>Unigene5024_All</td>
</tr>
<tr>
<td>Tpi-AAP5</td>
<td>Unigene90516_All</td>
</tr>
<tr>
<td>Tpi-AAP4</td>
<td>Unigene90869_All</td>
</tr>
<tr>
<td>Tpi-AAP3</td>
<td>CL12909.Contig2_All</td>
</tr>
</tbody>
</table>

**Table S3.1.** List of sequences used for analysis.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tpi-AAP1</td>
<td>CL16802.Contig2_All</td>
</tr>
<tr>
<td>Tpi-AAP2</td>
<td>Unigene50087_All</td>
</tr>
<tr>
<td>Tpi-LASGLV_like</td>
<td>Unigene5024_All</td>
</tr>
<tr>
<td>Tpi-AAP5</td>
<td>Unigene90516_All</td>
</tr>
<tr>
<td>Tpi-AAP4</td>
<td>Unigene90869_All</td>
</tr>
<tr>
<td>Tpi-AAP3</td>
<td>CL12909.Contig2_All</td>
</tr>
</tbody>
</table>
>Tpi-enterin Unigene 71958_All
SFVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFSVGRAPFGHFS

>Tpi-buccalin-1 CL7089.Contig4_All
MPRNQTLQQLLLALVLTNLQSLASDVADTADPDAFGDNSAVNDQVNLEHSAALAPECQDSEPDVEEKRLDGYYGFGQIGKRRGDYVSDFTGKQGRDQPPFSGAGHLGPKVRICTEYFDSDDKRISKDFGFAAGIKGKRLDLRDFGAFQLQGKRLRLDYTEGAGIGKRRDASYGFAGIGKRRPLTDYGAGIGKRRLDFGAGIGKRRKIDKFAGIGKRRSEDDFTYDDDVKVASENDSHVDKRASQLDKQVGSIS

>Tpi-AAP8 Unigene121884_All
IFDDRVSLRPYKWEMLHLHNVKVSDAGYQYCQINTDPVQIREVLHNVNEPPKIMDSLSSADTDTVREGTDVTVELCVNVTGLPPPSVTWYRTGFLDHEDDM

>Tpi-sensorin-A CL5976.Contig1_All
MSGLQFVNHMSILAVVLTVSTCLFEVHSAGPASMEHREEHHRPARKYRVGYMFKRASSASASASEPSALDIISRVDITKEELES

>Tpi-AAP7 Unigene55744_All
MAGVISPAAVACLWGFESLQGVEQNPPVKNENNNTNLAKHGRDIDKLSDKRLRLDSIAHGHSFGAFGRQGGGVPSTVAIDD

>Tpi-pleurin Unigene77450_All
MHHQASVFMVALLAVSGCVCAVFYTKSTDNDYPRIRRTFTYGGNGIHYPERGRRAVTSQLLLPGDLSMDSTFEKRGVFTQAGH

>Tpi-buccalin-4 Unigene77331_All
MFKTTHLMITLCLALFWSDIAIAIDSHSĐTNTGAFDEVDEVDKEADSQVYTGIPESDEYFDDDEKLKRKDPYFYGFGQGIKRVDKFGFAGIKKRELPYFGYSWARIGKRRSDNIGLSSDLKRXLDKLFAGFLGKKRRKIDKFFAGGIFGKKRVRDFSGAAGGIIKRPMYDGAFAGIGKKRDEDFRFDAYEDQSAALESSNPVSEKVQKLQTEQNSLQSSP

>Tpi-AAP9 Unigene29221_All
YFVTTFPSSTNRLENKPLQESTLPTDFPKQLSQGNPVNGCLKRTIPFSSFRQNSFSKMEIQISIIFLVSIFCVSAAPANDRQRQDLQHLL

>Tpi-AAP14 CL12069.Contig2_All
MRGLSRYAKKDEKFRMGRDLRDNQFDDIQNKFRMGRDLRNVDTEDKEFRIVGLNOLLARLDKKFMRGRFNKRHFHLENY

>Tpi-LFRF Unigene21004_All
MVNGSCCTGVALLAVSCVQSEETHETHTGLHSEDVeHSHTAHQAQKRNAFTTPQEEDDGLILTNDGLADPLASNEFLDKRSLFR

108
Chapter 4: Differential peptide expression in the central nervous system of the land snail *Theba pisana*, between active and aestivated

Published in Peptides 2015, S0196-9781(15)00232-6

**Abstract**

Hypometabolism is a physiological state of dormancy entered by many animals in times of environmental stress. There are gaps in our understanding of the molecular components used by animals to achieve this metabolic state. The availability of genomic and transcriptome data can be useful to study the process of hypometabolism at the molecular level. In this study, we use the land snail *Theba pisana* to identify peptides that may be involved in the hypometabolic state known as aestivation. We found a total of 22 neuropeptides in the central nervous system (CNS) that were differentially produced during activity and aestivation based on mass spectral-based neuropeptidome analysis. Of these, 4 were upregulated in active animals and 18 were upregulated in aestivation. A neuropeptide known to regulate muscle contractions in a variety of molluscs, the small cardioactive peptide A (sCAPA), and a peptide of yet unknown function (termed Aestivation Associated Peptide 12) were chosen for further investigation using temporal and spatial expression analysis of the precursor gene and peptide. Both peptides share expression within regions of the CNS cerebral ganglia and suboesophageal ganglia. Relative transcript abundance suggests that regulation of peptide synthesis and secretion is post-transcriptional. In summary, we provide new insights into the molecular basis of the regulation of aestivation in land snails through CNS peptide control.

**Key words:** Snail, *Theba pisana*, peptides, neuropeptides, aestivation, sCAP, central nervous system
Background

Many animals enter a state of suspended animation through slowed metabolism (hypometabolism). This can be for a short period of time (daily torpor) or for an extended period (hibernation), when conditions get too cold to ensure an adequate food supply and survival (Wilz and Heldmaier 2000). Animals, both vertebrate and invertebrate, may enter a similar hypometabolic state called aestivation, when conditions become too hot or dry (Storey and Storey 2012). Similar to hibernation, aestivation helps the animal to conserve energy and moisture supplies, allowing the body to withstand temperatures well outside the normal range, thus ensuring survival (Storey and Storey 2012).

Vertebrates that undergo aestivation include amphibians (Armentrout and Rose 1971, Barry and Shaffer 1994, Hudson, Lonhienne et al. 2008), reptiles (Christian, Webb et al. 2003, Winne, Willson et al. 2006, Firth, Christian et al. 2010) and fish (Frick, Bystriansky et al. 2008). Amongst invertebrates, aestivation is seen in pulmonate land snails (Storey 2002), some earthworms and insects (Bayley, Overgaard et al. 2010, Benoit 2010) sea sponges (Loomis 2010) and sea cucumbers (Yang, Zhou et al. 2006), as well as the similar hypometabolic states of dauer or diapause being found in the insects (Hahn and Denlinger 2011, Padilla and Ladage 2012) and nematodes. To ensure survival in harsh environmental conditions, aestivating animals must conserve both water and energy. To retain water within the body, animals such as frogs and lungfish form a cocoon (Withers 1998, Chew, Chan et al. 2004, Tracy, Reynolds et al. 2007) whereas land snails typically elevate themselves above the ground to avoid excessive ground-level temperatures. Alternatively, snails may shelter in crevices or under logs (Storey 2002) and form an epiphragm consisting of a specialised mucus that seals the opening in their shell, yet still allowing for some gas exchange (Jokinen 1978, McNair, Kier et al. 1981). During times of aestivation, respiration decreases to between 14% and 38% of normal and discontinuous breathing patterns prevail, which further minimise water loss (Gehlbach, Gordon et al. 1973, Horne 1973, Barnhart and McMahon 1987, Rees and Hand 1990, Withers 1993). Also, urea synthesis continues without excretion, thereby assisting with water retention by increasing the osmolality of body fluids to high concentrations (Rees and Hand 1993, Withers and Guppy 1996, Storey 2002). Heart rate is also reduced during aestivation (Horne 1973, Storey 2002), where reductions of from 45% to 60% have been observed in some land snails (Dale 1973, Rizzatti and Romero 2001).

The mechanisms of metabolic rate depression include suppressing fuel catabolism, ion channel arrest and reducing protein synthesis, gene expression and other energy
expensive activities such as cell division (Churchill and Storey 1989, Rees and Hand 1991, Storey and Storey 2012). The goal is to reduce net ATP turnover and reprioritize ATP use to sustain functions that are vital for long term survival (Ramnanan and Storey 2006, Ramnanan and Storey 2006). Global control is clearly necessary to regulate and coordinate the diverse range of biological functions altered during aestivation, yet little is known about the molecular components involved. However, such wide ranging control is typically achieved by chemical signalling using neurohormones derived from the central nervous system (CNS), most of which are peptides (Kiss 2011).

To date, differential gene expression in hibernating and aestivating animals has not been studied extensively, particularly within the CNS and with respect to neuropeptide responses. However, the land snail Helix pomatia has provided some insight, demonstrating that land snails may represent excellent experimental models for studies of the molecular components of aestivation. In the H. pomatia, changes in the levels of the peptides have been profiled for the CNS and haemolymph during hibernation (Pirger, Lubics et al. 2010). Approximately 19 CNS (including some known molluscan neuropeptides) and 10 haemolymph peptides were significantly higher in active snails. Other peptides (11 in the brain and 13 in the haemolymph) were present only in hibernation samples, while several peptides showed no difference between metabolic states.

Theba pisana (the white Italian snail) is native to the Mediterranean region but has now become established as a major pest in grain crops, pastures and vineyards throughout the world, including southern Australia (Baker 1991, Baker 1998, Baker 2008). It is predominantly active during autumn, winter and spring, and aestivates in early summer when temperatures increase (Baker and Vogelzang 1988). Extending upon the inventory of known neuropeptides in terrestrial snails (Kiss, Pirger et al. 2006), the in silico neuropeptidome of T. pisana has been described, with many neuropeptides supported by mass spectrometry of the CNS (Adamson, Wang et al. 2015). In this study, we have analysed the differential expression of CNS peptides during periods of snail aestivation and activity. Of those differentially expressed, two peptides were further investigated by characterization of their temporal and spatial expression.
Materials and methods

Animals and tissue collection for proteome analysis

_Theba pisana_ were collected from agricultural sites surrounding Warooka, located on the Yorke Peninsula, South Australia in early spring (September). Snails were transported to the University of the Sunshine Coast (USC) and housed within purpose-built enclosures. Snails were provided with water and food (cucumber and carrot) _ad libitum_ and maintained at room temperature. Active snails were conditioned by feeding and supplied with water daily over 7 days. Snails to be used for aestivation experiments were placed into glass jars without food or water and kept in an incubator on a cycle of 12 h at 30°C with light, 12 h at 20°C dark, and humidity at approximately 40%, to emulate South Australian summer conditions (www.bom.gov.au). The positions of the snails were marked on the jars after 14 days. Any snails that had not moved following a further 21 days were deemed to be in aestivation.

Protein isolation from CNS and nanoHPLC, mass spectrometry LC-ESI-QTOF peptide identification

Active moving _T. pisana_ were quickly killed upon removal of CNS while aestivated snails were rapidly killed by immersion in liquid nitrogen prior to CNS removal. CNS were dissected out, and separately pooled before freezing in liquid nitrogen prior to storage at -80°C until use. Frozen samples of CNS were ground to a powder under liquid nitrogen in a mortar, then quickly weighed and homogenized in extraction buffer (90% methanol, 9% glacial acetic acid in deionized water) in a 1:5 w:v ratio. Crude extracts were then sonicated with three pulses, 30 s each, and centrifuged for 20 min (16,000 x g, 4°C). Supernatant was collected and lyophilised.

The CNS extracts were analyzed by LC-MS/MS on a Shimadzu Prominence Nano HPLC (Japan) coupled to a Triple-ToF 5600 mass spectrometer (ABSCIEX, Canada) equipped with a nano electrospray ion source. Aliquots (6 µl) of each extract were injected onto a 50 mm x 300 µm C18 trap column (Agilent Technologies, Australia) at 30 µl/min. The samples were de-salted on the trap column for 5 minutes using solvent A [0.1% formic acid (aq)] at 30 µl/min. The trap column was then placed in-line with the analytical nano HPLC column, a 150 mm x 75 µm 300SBC18, 3.5 µm (Agilent Technologies) for mass spectrometry analysis. Peptide elution used a linear gradient of 1-40% solvent B [90:10...
acetonitrile: 0.1% formic acid (aq)] over 35 min at 300 nl/minute flow rate, followed by a steeper gradient from 40% to 80% solvent B over 5 min. Solvent B was then held at 80% for 5 min to wash the column and then returned to 1% solvent B for equilibration prior to the next sample injection. The ionspray voltage was set to 2400V, declustering potential (DP) 100V, curtain gas flow 25, nebuliser gas 1 (GS1) 12 and interface heater at 150°C. The mass spectrometer acquired 500ms full scan TOF-MS data followed by 20 by 50 ms full scan product ion data in an Information Dependent Acquisition (IDA) mode. Full scan TOFMS data was acquired over the mass range 350-1800 and for product ion ms/ms 100-1800. Ions observed in the TOF-MS scan exceeding a threshold of 100 counts and a charge state of +2 to +5 were set to trigger the acquisition of product ion, ms/ms spectra of the resultant 20 most intense ions. The data was acquired and processed using Analyst TF 1.5.1 software (ABSCIEX, Canada).

Fragmentation data was analysed by PEAKS v7.0 (BSI, Canada) software. Sequences of peptides were determined manually and/or by comparing the fragmentation patterns with those predicted from the T. pisana CNS transcriptome (Adamson, Wang et al. 2015). Search parameters were as follows: no enzyme was used; variable modifications included methionine oxidation, conversion of glutamine/glutamate to pyroglutamic acid, deamidation of asparagine and peptide amidation. Precursor mass error tolerance was set to 20 ppm and a fragment ion mass error tolerance was set to 0.05 Da. Maximum expectation value for accepting individual peptide ion scores [-10*Log(p)] was set to ≤0.01, where p is the probability that the observed match is a random event. Proteins and their supporting peptides were obtained and analysed. The quantitative analysis of proteins was carried out using the label-free quantification module (PEAKS Q) of PEAKS v7.0, and relative concentrations of proteins at two stages were compared. Biological triplicates [i.e., (A1, B1), (A2, B2) and (A3, B3)] of each stage were used in tandem repeats, and the average values were calculated as the final results. Extracted peptides were quantified based on absorbance at 280 nm using a NanoDrop spectrophotometer for later normalization. For each run, about 1.5 µg of the mix was then analyzed via liquid chromatography combined with mass spectrometry on a Triple-TOF. In addition to the protein identification method mentioned above, peptide feature fold change and protein unique peptide were set to 2, peptide feature significance filter was set to 0.01 and other parameters were adjusted according to the ‘volcano plot’ generated by PEAKS.
**Gene analysis**

A BLAST analysis was performed against an in-house derived *Theba pisana* neuropeptide database (http://thebadb.bioinfo-minzhao.org/) to characterize which of the significantly up- and down-regulated peptides identified from MS/MS corresponded to protein precursors. Schematic diagrams of protein domain structures were prepared using Domain Graph (DOG, version 2.0) software (Ventura Garcia, Wajnberg et al. 2002). Sequences were derived by BLASTp analysis of the NCBI database and through analysis of land snail and slug assemblies derived from SRA file SRP056280, then protein sequences were aligned using the MEGA 5.1 (Tamura, Peterson et al. 2011) platform with the clustalW protocol utilising the Gonnet protein weight matrix. Neighbour-joining trees were generated based off these alignments. Gene expression levels were determined from RNA-seq data (NCBI Genbank under SRA file SRP056280) using the FPKM method, which represents the fragments per kilobase of transcript per million fragments mapped. The formula used to assign FPKM was:

\[
FPKM = \frac{10^6 C}{NL/10^3}
\]

Where \(C\) represents the number of fragments that uniquely aligned to gene \(g\); \(N\) is the total number of fragments that uniquely aligned to all genes; and \(L\) indicates the number of bases on the gene.

**Absolute qPCR**

RNA was extracted from tissue using TRIzol Reagent (Invitrogen), as per the manufacturer’s protocol. Each sample contained CNS from 3 *T. pisana*. A total of 15 aestivating and 14 active snail samples were prepared. Following extraction, RNA was assessed for quality by visualization on a 1.2% denaturing formaldehyde agarose gel, quantified using a Nanodrop spectrophotometer (Thermo Scientific). Approximately 500 ng of total RNA were reverse-transcribed using QuantiTect Reverse Transcription kit (Qiagen) as per the manufacturer’s protocol. Products were visualized by agarose gel electrophoresis, then cDNA was stored at -20°C until use.

Absolute qPCR was performed using a SensiFast HRM kit (Bioline) as per the manufacturer’s protocol, on a Rotor-gene 6000 cycler (Corbett research) using Rotor-gene 1.7.87 software. A total volume of 10 µl per reaction was used, containing 1x SensiFast, and
400 nM of each primer, with 1 µl of template cDNA. All reactions, including a no template control (NTC) for each primer pair, were done in duplicate.

Cycling parameters used were an initial hold at 95°C for 2 min followed by 40 cycles of denaturation at 95°C for 5 s, annealing at 60°C for 10 s and extension at 72°C for 20 s acquiring to cycle A (Green). Melt curve analysis was performed rising by 1°C per cycle from 70°C to 95°C and was held for 90 s on the first step and 5 s on the next steps. Quantitative PCR primers for the genes of interest were designed from transcriptome sequences using Primer3 software. Primers were: sCAP forward GGGGCCTCTTTTCTTTAGTG, reverse GTGCGCTTCAGTGCTCCTACA; and AAP12 forward TGACGTGCAAGCTCCAATAG, reverse ATTTCCATCGGCATCTTTCA. A standard curve of known concentration was prepared for the genes of interest. Complementary DNA was amplified using the primers described and the PCR products were used as template for the standard curves. Following PCR the amplified products were visualised to check for correct size and then purified using the QIAquick kit (Qiagen). Purified products were quantified three times and mean concentrations calculated. From this and the known length of the products, the copy number was calculated using the calculator of URI Genomics & Sequencing Centre (http://www.uri.edu/research/gsc/resources.cdna.html). A DNA concentration of 2 x 10^9 was then prepared and serially diluted six times into 1:10 dilutions. A standard curve for the gene using the serially diluted PCR product was prepared (in triplicate) using the reaction conditions described above. No template control (NTC) was carried out in duplicate. Target gene expression levels of the CNS tissue from active and aestivating animals was quantified using the qPCR described above, samples were run in duplicate using cDNA synthesised from CNS RNA.

**Immunolocalisation**

Rabbit polyclonal antibodies were generated against SGYLAFPRM-amide that is present within the Tpi-sCAP precursor, and the peptide EEMLSVKGDGKVQ (AAP12) (Genscript). CNS were isolated from active and aestivated *T. pisana* and then fixed in 4% paraformaldehyde overnight at 4°C, before dehydration in ascending concentrations of ethyl alcohol for 30 min each, cleared in xylene three times, infiltrated, and embedded in paraffin. Serial transverse sections of the tissues were cut at 5 µm thickness using a microtome. Sections were then deparaffinized in xylene and rehydrated in a descending concentration of
ethanol. Subsequently, sections were incubated in 0.1% glycine in 0.1 M PBS for 30 min, and washed three times with 0.1 M PBS (phosphate buffered saline) containing 0.4% Triton X-100 (PBST). Non-specific binding was blocked in 4% normal goat serum in PBST for 2 h, followed by incubation in the primary antibody (anti-peptide) at a dilution of 1:1000 in blocking solution, at 4°C overnight. Sections were then washed 3 times with PBST, and incubated for 2 h with Alexa Fluor 488 conjugated goat anti-rabbit IgG (Santa Cruz, USA) at room temperature. After washing with PBST, nuclei were stained with DAPI (Santa Cruz, USA) for 10 min. Finally, sections were washed with PBST and mounted with VECTA shield fluorescent mounting medium (Molecular Probes) before viewing under a confocal laser-scanning microscope (Nikon). In negative controls, tissues were processed by the same protocol, but preimmune mouse serum was used instead of primary antibody.

Results

Identification of differentially expressed peptides in active and aestivated Theba pisana CNS

Peptide extraction of the T. pisana CNS during periods of normal activity and aestivation followed by LC-MS/MS analysis was used to identify those peptides of different abundance. To ensure CNS samples accurately reflected the natural deep aestivation state and that aestivated snails did not begin arousal, aestivated snails were rapidly killed by immersion in liquid nitrogen prior to CNS removal, while active moving T. pisana were quickly killed upon removal of CNS. Using this method, we have identified 22 peptides that were significantly differentially expressed between normal active and aestivated metabolic states (Fig. 4.1, Table 4.1 and Table S4.1).
Hierarchical clustering of peptide expression in the *Theba pisana* CNS of active versus aestivated snails based on MS/MS analysis. The 22 differentially expressed peptides can be divided into two clusters.

The majority match to precursor sequences that resemble either uncharacterised neuropeptides or unknown peptides (Aestivation Associated Peptides, AAPs). Those identified as uncharacterised neuropeptides, AAP1, AAP3, AAP4, AAP7, AAP11 and AAP13, were defined as such due to the presence of either or both a signal sequence and multiple sites for cleavage of peptides, some of which would be amidated (Fig. 4.2). Of the 22 peptides differentially expressed, 4 peptides were up-regulated in active snails, Leu-Phe-Arg-Phe (LFRF), and 3 peptides that have not been previously identified in any species (Tpi-AAP11-13).

**Fig. 4.1.** Precursor protein sequences with neuropeptide-like characteristics that have no defined function. Highlights: Yellow, signal peptide; red, cleavage site; green, amidation; blue, peptide match obtained following mass spectrometry.

The remaining 18 peptides were up-regulated in aestivation, including the molluscan neuropeptides previously identified...
in the T. pisana neuropeptidome analysis; the small cardioactive peptide A (sCAPₐ), feeding circuit-activating peptide (FCAP), enterin, pleurin, sensorin-A, a LSGLV-like peptide and buccalin.

**Table 4.1.** List of peptides up- and down-regulated during aestivation in Theba pisana CNS as determined by mass spectrometry; the ratio aestivated:active peptide levels is also shown. The database used for this analysis was derived from the T. pisana CNS (Adamson, Wang et al. 2015).

<table>
<thead>
<tr>
<th>Neuropeptide</th>
<th>Predicted avg mass</th>
<th>m/z</th>
<th>[z] charge</th>
<th>Aestivated</th>
<th>Active</th>
<th>Ratio</th>
<th>Sequence</th>
<th>Precursor length AA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tpi-LSGLV-like</td>
<td>977.46</td>
<td>489.2</td>
<td>2</td>
<td>8.32E+03</td>
<td>0.00E+00</td>
<td>1.00E+00</td>
<td>RPEDEL6SG6</td>
<td>286 (ent FL)</td>
</tr>
<tr>
<td>Tpi-pleurin</td>
<td>1474.64</td>
<td>492.2</td>
<td>3</td>
<td>8.20E+03</td>
<td>0.00E+00</td>
<td>1.00E+00</td>
<td>SVEGTGDSG/SID6-APamide(-58)</td>
<td>179</td>
</tr>
<tr>
<td>Tpi-pleurin**</td>
<td>1513.78</td>
<td>759.9</td>
<td>2</td>
<td>4.39E+03</td>
<td>0.00E+00</td>
<td>1.00E+00</td>
<td>TTYSTQGS/TIPamide(-58)</td>
<td>139</td>
</tr>
<tr>
<td>Tpi-sCAPA**</td>
<td>932.51</td>
<td>417.5</td>
<td>3</td>
<td>1.08E+00</td>
<td>1.00E+00</td>
<td>1.00E+00</td>
<td>SYKELPAP6</td>
<td>120</td>
</tr>
<tr>
<td>Tpi-FCAP</td>
<td>1072.86</td>
<td>520.77</td>
<td>2</td>
<td>6.26E+03</td>
<td>2.93E+03</td>
<td>1.00E+00</td>
<td>SCYAPPAMamide(-58)</td>
<td>130</td>
</tr>
<tr>
<td>Tpi-FCAP</td>
<td>1072.86</td>
<td>520.77</td>
<td>2</td>
<td>6.26E+03</td>
<td>2.93E+03</td>
<td>1.00E+00</td>
<td>SCYAPPAMamide(-58)</td>
<td>130</td>
</tr>
<tr>
<td>Tpi-sensorin-A</td>
<td>1181.37</td>
<td>507.80</td>
<td>2</td>
<td>7.62E+00</td>
<td>0.00E+00</td>
<td>1.00E+00</td>
<td>AKYTV/S6-Aamide(-48)</td>
<td>120</td>
</tr>
<tr>
<td>Tpi-sensorin-A</td>
<td>1181.37</td>
<td>507.80</td>
<td>2</td>
<td>7.62E+00</td>
<td>0.00E+00</td>
<td>1.00E+00</td>
<td>AKYTV/S6-Aamide(-48)</td>
<td>120</td>
</tr>
<tr>
<td>Tpi-sensorin-A</td>
<td>1181.37</td>
<td>507.80</td>
<td>2</td>
<td>7.62E+00</td>
<td>0.00E+00</td>
<td>1.00E+00</td>
<td>AKYTV/S6-Aamide(-48)</td>
<td>120</td>
</tr>
<tr>
<td>Tpi-buccalin-1</td>
<td>1181.37</td>
<td>584.81</td>
<td>2</td>
<td>1.21E+04</td>
<td>9.73E+03</td>
<td>1.00E+00</td>
<td>NLDYVS+AFG6G-amide(-58)</td>
<td>165</td>
</tr>
<tr>
<td>Tpi-buccalin-4</td>
<td>1181.37</td>
<td>584.81</td>
<td>2</td>
<td>1.21E+04</td>
<td>9.73E+03</td>
<td>1.00E+00</td>
<td>NLDYVS+AFG6G-amide(-58)</td>
<td>165</td>
</tr>
<tr>
<td>Tpi-AAP1</td>
<td>980.05</td>
<td>460.273</td>
<td>2</td>
<td>4.46E+02</td>
<td>0.00E+00</td>
<td>1.00E+00</td>
<td>G6OQSVFRV-amide(-96)</td>
<td>465</td>
</tr>
<tr>
<td>Tpi-AAP2</td>
<td>2208.58</td>
<td>757.078</td>
<td>3</td>
<td>5.21E+00</td>
<td>0.00E+00</td>
<td>1.00E+00</td>
<td>KL6ALR6AIPK7Y6SH6L6PA</td>
<td>108</td>
</tr>
<tr>
<td>Tpi-AAP3</td>
<td>1936.01</td>
<td>641.967</td>
<td>3</td>
<td>5.57E+00</td>
<td>0.00E+00</td>
<td>1.00E+00</td>
<td>PHY6LK6EE6ES+AF6G6G</td>
<td>486</td>
</tr>
<tr>
<td>Tpi-AAP4</td>
<td>1152.21</td>
<td>369.323</td>
<td>3</td>
<td>4.12E+02</td>
<td>1.32E+03</td>
<td>1.00E+00</td>
<td>PH6M+SSAF6+V6G-amide(-78)</td>
<td>264</td>
</tr>
<tr>
<td>Tpi-AAP5</td>
<td>1152.21</td>
<td>369.323</td>
<td>3</td>
<td>4.12E+02</td>
<td>1.32E+03</td>
<td>1.00E+00</td>
<td>PH6M+SSAF6+V6G-amide(-78)</td>
<td>264</td>
</tr>
<tr>
<td>Tpi-AAP6</td>
<td>1195.31</td>
<td>371.597</td>
<td>3</td>
<td>7.40E+00</td>
<td>3.40E+00</td>
<td>1.00E+00</td>
<td>AH6HGL6TAKY</td>
<td>384 (ent FL)</td>
</tr>
<tr>
<td>Tpi-AAP7</td>
<td>1195.31</td>
<td>371.597</td>
<td>3</td>
<td>7.40E+00</td>
<td>3.40E+00</td>
<td>1.00E+00</td>
<td>AH6HGL6TAKY</td>
<td>384 (ent FL)</td>
</tr>
<tr>
<td>Tpi-AAP8</td>
<td>1195.31</td>
<td>371.597</td>
<td>3</td>
<td>7.40E+00</td>
<td>3.40E+00</td>
<td>1.00E+00</td>
<td>AH6HGL6TAKY</td>
<td>384 (ent FL)</td>
</tr>
<tr>
<td>Tpi-AAP9</td>
<td>1195.31</td>
<td>371.597</td>
<td>3</td>
<td>7.40E+00</td>
<td>3.40E+00</td>
<td>1.00E+00</td>
<td>AH6HGL6TAKY</td>
<td>384 (ent FL)</td>
</tr>
<tr>
<td>Tpi-AAP10</td>
<td>1195.31</td>
<td>371.597</td>
<td>3</td>
<td>7.40E+00</td>
<td>3.40E+00</td>
<td>1.00E+00</td>
<td>AH6HGL6TAKY</td>
<td>384 (ent FL)</td>
</tr>
<tr>
<td>Tpi-AAP11</td>
<td>1195.31</td>
<td>371.597</td>
<td>3</td>
<td>7.40E+00</td>
<td>3.40E+00</td>
<td>1.00E+00</td>
<td>AH6HGL6TAKY</td>
<td>384 (ent FL)</td>
</tr>
<tr>
<td>Tpi-AAP12**</td>
<td>1195.31</td>
<td>371.597</td>
<td>3</td>
<td>7.40E+00</td>
<td>3.40E+00</td>
<td>1.00E+00</td>
<td>AH6HGL6TAKY</td>
<td>384 (ent FL)</td>
</tr>
<tr>
<td>Tpi-AAP13</td>
<td>1195.31</td>
<td>371.597</td>
<td>3</td>
<td>7.40E+00</td>
<td>3.40E+00</td>
<td>1.00E+00</td>
<td>AH6HGL6TAKY</td>
<td>384 (ent FL)</td>
</tr>
</tbody>
</table>

**To which antibodies were generated.** FL: full-length precursor; (-98): amidation; (+15.99): methylation

**Characterization of differentially expressed sCAP**

We identified that the Tpi-sCAPₐ, SGYLAFPRM-amide is up-regulated in aestivation (Table 4.1). T. pisana contains 3 sCAP precursor isoforms which all contain 2 sCAP neuropeptides that immediately follow the signal sequence. In Tpi-sCAP1 and 2, sCAPₐ and sCAPₐ are separated only by an Arg residue that would act as a cleavage site, while in Tpi-sCAP3 the two neuropeptides are 31 amino acids apart. We found that the second sCAP neuropeptide in the precursor, MNLYAFPRM-amide (sCAPₐ), was not significantly different between metabolic states in T. pisana. Mass spectral fragmentation data for the sCAP precursor peptides is shown in Table S4.2. Besides the sCAP neuropeptides, all sCAP
precursors have a region containing a total of 6 spatially conserved cysteine residues whose function has not yet been defined and were not detected by our MS analysis (Fig. 4.3A).

**Fig. 4.3.** Analysis of the *Theba pisana* sCAPs. A) Multiple sequence alignment of Tpi-sCAP precursors with sCAP of other molluscs. Blue shading shows amino acids that are highly conserved. Sequence log above alignment shows amino acid conservation, represented by letter size. B) Unrooted phylogenetic tree with corresponding mollusc sCAP precursor protein schematic. Bootstrap values are shown and scale bar represents amino acid differences. All sequences used for analysis can be found in Table S4.1.

Since sCAP\textsubscript{A} is distributed widely throughout the CNS in other molluscs, and is known to regulate muscle contraction, we speculated that sCAP\textsubscript{A} could regulate metabolic events in aestivating snails. For that reason, Tpi-sCAP was further investigated via transcriptome screening and comparative sequence analysis, followed by spatial expression analysis of sCAP\textsubscript{A} within the CNS. Based on its relative expression profiles derived from
next-generation sequence data of the CNS (Danks 2002), Tpi-sCAP is one of the most highly expressed transcripts (combined active and aestivated adult animals; FPKM= 732) in the CNS, with relatively lower expression in the hepatopancreas (FPKM=0.32) and muscle tissue (FPKM=0.2) (Table 4.2). Phylogenetic tree analysis shows that gastropod sCAP clearly branch separately to the non-gastropods, despite demonstrating a similar precursor organisation (Fig. 4.3B).

Table 4.2. List of abundance of differentially expressed gene transcripts of peptides in CNS, hepatopancreas and muscle tissue in Theba pisana.

<table>
<thead>
<tr>
<th>peptide name</th>
<th>peptides</th>
<th>peptide sequence</th>
<th>CNS FPKM</th>
<th>CNS Rank/250849</th>
<th>Hepatopancreas FPKM</th>
<th>Hepatopancreas Rank/250849</th>
<th>Muscle FPKM</th>
<th>Muscle Rank/250849</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tpi-LASGLV-LIKE 1</td>
<td>RPFDELGSG</td>
<td>1</td>
<td>1985</td>
<td>14</td>
<td>12956</td>
<td>20.15</td>
<td>6602</td>
<td></td>
</tr>
<tr>
<td>Tpi-enterin 1</td>
<td>GPFIVG(198)</td>
<td>1</td>
<td>679</td>
<td>174</td>
<td>5301</td>
<td>109.9</td>
<td>907</td>
<td></td>
</tr>
<tr>
<td>Tpi-sCAP** 1</td>
<td>SGYLAFPRM</td>
<td>1</td>
<td>732</td>
<td>157</td>
<td>166685</td>
<td>0.2</td>
<td>189403</td>
<td></td>
</tr>
<tr>
<td>Tpi-sensorin-A 1</td>
<td>KMRFGMYM(198)</td>
<td>1</td>
<td>51</td>
<td>2244</td>
<td>111511</td>
<td>7</td>
<td>2244</td>
<td></td>
</tr>
<tr>
<td>Tpi-buccalin-1 1</td>
<td>RUDYFGFAG</td>
<td>2</td>
<td>74800</td>
<td>0</td>
<td>212103</td>
<td>0.11</td>
<td>204144</td>
<td></td>
</tr>
<tr>
<td>Tpi-buccalin-4 1</td>
<td>RVDKFGF +/-</td>
<td>1</td>
<td>2785</td>
<td>18</td>
<td>82660</td>
<td>8.3</td>
<td>18598</td>
<td></td>
</tr>
<tr>
<td>Tpi-LFRF 1</td>
<td>GSAEFAPK</td>
<td>454</td>
<td>366</td>
<td>0.00</td>
<td>300159</td>
<td>140.1</td>
<td>691</td>
<td></td>
</tr>
<tr>
<td>Tpi-AAP1 1</td>
<td>GSSFV(198)</td>
<td>1</td>
<td>54</td>
<td>2264</td>
<td>201297</td>
<td>3.58</td>
<td>44600</td>
<td></td>
</tr>
<tr>
<td>Tpi-AAP2 1</td>
<td>KEAMAPVFYSAKLEAKA</td>
<td>1</td>
<td>470</td>
<td>252</td>
<td>116491</td>
<td>2.4</td>
<td>64478</td>
<td></td>
</tr>
<tr>
<td>Tpi-AAP3 1</td>
<td>WYGLKFIPVESNYGAD</td>
<td>6</td>
<td>24113</td>
<td>0.01</td>
<td>209654</td>
<td>0.6</td>
<td>130035</td>
<td></td>
</tr>
<tr>
<td>Tpi-AAP4 1</td>
<td>AVVINTPL(198)</td>
<td>67</td>
<td>184</td>
<td>1.09</td>
<td>109866</td>
<td>24.2</td>
<td>5275</td>
<td></td>
</tr>
<tr>
<td>Tpi-AAP5 1</td>
<td>AIVNGGFAK</td>
<td>2.7</td>
<td>55715</td>
<td>10.83</td>
<td>4339</td>
<td>97.83</td>
<td>1041</td>
<td></td>
</tr>
<tr>
<td>Tpi-AAP6 1</td>
<td>KASKVRKHKTEK</td>
<td>0.12</td>
<td>196440</td>
<td>0.92</td>
<td>115983</td>
<td>3.85</td>
<td>43808</td>
<td></td>
</tr>
<tr>
<td>Tpi-AAP7 1</td>
<td>NDESHELKDIAGG</td>
<td>54</td>
<td>2279</td>
<td>0.09</td>
<td>194107</td>
<td>1.64</td>
<td>88434</td>
<td></td>
</tr>
<tr>
<td>Tpi-AAP8 1</td>
<td>NDHSPEEFVGD</td>
<td>61</td>
<td>182</td>
<td>1.09</td>
<td>109866</td>
<td>24.2</td>
<td>5275</td>
<td></td>
</tr>
<tr>
<td>Tpi-AAP9 1</td>
<td>AIVNGGFAK</td>
<td>2.7</td>
<td>55715</td>
<td>10.83</td>
<td>4339</td>
<td>97.83</td>
<td>1041</td>
<td></td>
</tr>
<tr>
<td>Tpi-AAP10 1</td>
<td>KASKVRKHKTEK</td>
<td>0.12</td>
<td>196440</td>
<td>0.92</td>
<td>115983</td>
<td>3.85</td>
<td>43808</td>
<td></td>
</tr>
<tr>
<td>Tpi-AAP11 1</td>
<td>EEMMLSVKGDGKVVQ</td>
<td>405</td>
<td>320</td>
<td>3.57</td>
<td>38191</td>
<td>40</td>
<td>2912</td>
<td></td>
</tr>
<tr>
<td>Tpi-AAP13 1</td>
<td>TQNIENVKDEREE</td>
<td>18.3</td>
<td>7441</td>
<td>0.08</td>
<td>199014</td>
<td>0</td>
<td>237148</td>
<td></td>
</tr>
</tbody>
</table>

**To which antibodies were generated

Polyclonal antibodies generated to the Tpi-sCAP SGYLAFPRM-amide were used for spatial immunolocalization within the CNS of aestivated (Fig. 4.4) and active (Fig. S4.1) T.pisana. Qualitatively, similar overall localization was found between aestivated and active states. Immunoreactivity was widely distributed throughout the CNS, within regions of the cerebral ganglia, pleural ganglia, parietal ganglia, pedal ganglia and visceral ganglia (Fig. 4.4A-L). Within the cerebral ganglia, immunopositive staining was clearly observed within all regions (Fig. 4.4A,B,D,E,G,H,J,K). Staining was also observed within the pedal ganglia (Fig. 4.4A,C) and in the pleural, parietal, visceral region (Fig. 4.4D,F,G,I,J,L), where individual neurons can be seen, particularly within the region of the visceral ganglia. Nerve fibres can be seen in close proximity to the pedal, pleural, parietal, visceral regions (Fig. 4.4A,C,D,F,G,I).
Fig. 4.4. Immunolocalisation analysis of *Theba pisana* sCAP. (A-L) Localisation of peptide within sections of aestivated CNS. Arrow heads show regions of nerve fibre immunoreactivity. PaG, parietal ganglia; PeG, pedal ganglia; PIG, pleural ganglia; RCG, right cerebral ganglia; VG, visceral ganglia. Immunolocalisation within active CNS is shown in Figure S4.1.
Characterization of differentially expressed AAP12

One peptide that was significantly down-regulated in *T. pisana* CNS during aestivation and matched with a previously uncharacterised precursor sequence, Tpi-AAP12, was further analysed. The peptide identified by MS was EEMLSVKGDGKVQ, encoded by a transcript that is also one of the more highly expressed transcripts (combined active and aestivated adult CNS; FPKM= 405) in the CNS, with relatively lower expression in the hepatopancreas (FPKM=3.57) and muscle tissue (FPKM=40) (Table 4.2). Mass spectral fragmentation data for the AAP-12 precursor peptide is shown in Table S4.2. BLASTp analysis indicates that there is a similar protein precursor present in *Aplysia californica*, *Lottia gigantea* and *Crassostrea gigas*. We further investigated the transcriptomes of several land gastropods available within the sequence read archives of NCBI where we identified additional homologous proteins. A multiple sequence alignment and phylogenetic tree groups the three land gastropods (*T. pisana*, *Cornu aspersum*, *Cochlicella acuta*) together, most closely related to the marine nudibranch, *A. californica* (Fig. 4.5A). Outside of the land gastropods, only the *A. californica* precursor had a region with similarity to the Tpi-AAP12 (Fig. 4.5B).

---

**Fig. 4.5.** Analysis of the *Theba pisana* AAP12. A) Phylogenetic tree of Tpi-AAP12 precursor protein with homologues of other species. Bootstrap values are shown and scale bar represents amino acid differences. B) Multiple sequence alignment of the AAP12 with other species. Blue shading shows amino acids that are highly conserved. Sequence log above alignment shows amino acid conservation, represented by letter size. All sequences used for analysis can be found in Table S4.1.
Polyclonal antibodies generated to the Tpi-AAP12 EEMLSVKGDGKVQ were used for spatial immunolocalization within the CNS of aestivated *T. pisana*. Expression of Tpi-AAP12 can be observed within regions of the cerebral ganglia, pleural ganglia, parietal ganglia, pedal ganglia and visceral ganglia (Fig. 4.6A-L). Within the cerebral ganglia, immunopositive staining is clearly observed within all regions including cells surrounding neuron clusters (Fig. 4.6A,B,D,E,G,H) with individual neurons visible (Fig. 4.6J,K). Staining is seen in the pedal ganglia and left pleural ganglia (Fig. 4.6A,C) and in the pleural, parietal, visceral regions (Fig. 4.6D,F,G,I), with individual neurons observed in the visceral ganglia (Fig. 4.6J,L). As was also seen for sCAP, no differences in spatial peptide expression of Tpi-AAP12 were observed between active and aestivated CNS using this method. Overall, Tpi-AAP12 expression intensity was less than that observed for Tpi-sCAP.

**Fig. 4.6.** Immunolocalisation analysis of *Theba pisana* AAP12. (A-L) Localisation of peptide within sections of aestivated CNS. Arrow heads show regions of nerve fibre immunoreactivity. PaG, parietal ganglia; PeG, pedal ganglia; PlG, pleural ganglia; RCG,
right cerebral ganglia; VG, visceral ganglia. Immunolocalisation within active CNS is shown in Figure S4.2.

**Quantitative PCR analysis of sCAP and AAP12 transcripts**

Absolute qPCR was performed on CNS samples from active and aestivated snails to determine relative gene expression of the Tpi-sCAP and Tpi-AAP12 transcripts. There was a trend for Tpi-sCAP to be up-regulated in aestivated CNS, however this was not statistically significant (student t-test P (T<=t) two-tail 0.18949; Fig. 4.7). Expression of the Tpi-AAP12 transcript was significantly up-regulated in aestivated snails (student t-test P (T<=t) two-tail 0.03561; Fig. 4.7)

![Quantitative PCR expression analysis of Tpi-sCAP and Tpi-AAP12 in CNS of active and aestivated Theba pisana. Asterisk indicates significantly different (P<0.05).](image)

**Fig. 4.7.** Quantitative PCR expression analysis of Tpi-sCAP and Tpi-AAP12 in CNS of active and aestivated Theba pisana. Asterisk indicates significantly different (P<0.05).

**Discussion**

There are as many as 5504 precursor proteins derived from the *T. pisana* CNS transcriptome that are destined for secretion, based on the presence of an N-terminal signal sequence (Danks 2002). Of these, 35 were known molluscan neuropeptide precursors. In this study, we have investigated the relative abundance of two *T. pisana* CNS peptides during aestivation and normal activity.
Abundance of CNS peptides is modified during aestivation

Amongst those peptides that were identified as differentially expressed in the *T. pisana* CNS based on mass spectral analysis were the mollusca neuropeptides FCAP, enterin, pleurin, sensorin-A, LFRF, buccalin and sCAP, which have been identified in several aquatic molluscs (Cropper, Miller et al. 1988, Sweedler, Li et al. 2002, Moroz, Edwards et al. 2006, Veenstra 2010, Stewart, Favrel et al. 2014), and from the *T. pisana* neuropeptidome (Adamson, Wang et al. 2015). We have previously found that similar neuropeptides are differentially present in the CNS in aestivating *T. pisana* using a slightly different preparation, thus supporting the current findings (Adamson, Wang et al. 2015). In this study, we’ve found that some precursor proteins contain neuropeptide-like characteristics, yet have never been described, including the Tpi-AAP1 and Tpi-AAP3 which contain multiple variable repeats of C-terminal FVRIa and N-terminal FMRGL, respectively. These would appear to be a variation on the FMRFa precursor protein that has previously been identified in *T. pisana* (Adamson, Wang et al. 2015), and is known to have a cardioexcitatory effect in molluscs (Higgins, Price et al. 1978, Price and Greenberg 1989). Another neuropeptide-like Tpi-AAP, Tpi-AAP13, does BLASTp most closely with tachykinin precursors, although does not contain the C-terminal sequence, Gly-Leu-Met-NH2, which is most typical for tachykinins (Satake, Kawada et al. 2003).

Analysis of Tpi-sCAP and Tpi-AAP12 in the CNS

We found that the neuropeptide Tpi-sCAP\textsubscript{A} is up-regulated in the CNS of aestivating snails which contrasts with a study of hibernating *H. pomatia*, where the sCAP\textsubscript{A} showed higher expression in active animals (Pirger, Lubics et al. 2010). This may be due to the slightly different extraction methods used or the different physiological states, aestivation versus hibernation. The neuropeptide sCAP\textsubscript{A} was first identified in active *Aplysia*, where it is highly concentrated in the buccal ganglia, implying an important role in feeding (Lloyd, Mahon et al. 1985). It is also present in fibres and neuronal bodies throughout the CNS, suggesting involvement in more widespread functions (Lloyd, Mahon et al. 1985, Santama, Brierley et al. 1994, Li, Moroz et al. 1998). Also, in *Lymnaea stagnalis*, sCAP\textsubscript{A} is present in all ganglia and in nerve fibres that link many of them (Santama, Wheeler et al. 1994).

The Tpi-sCAP precursors have a general organisation similar to that of other molluscan sCAP precursors (Dobbins, Schofield et al. 1999), including signal peptides, 2
sCAP neuropeptides and an sCAP-associated peptide, with the exception of the cephalopods. This general precursor organisation resembles that of the gonadotropin-releasing hormone (GnRH) precursor, which also contains a signal sequence, a small GnRH peptide, and a GnRH associated peptide (GAP) which may have some function in the bioactivity of the GnRH peptide (Bigot, Zatllyn-Gaudin et al. 2012).

The high levels of Tpi-sCAPA found throughout the T. pisana CNS, including the cerebral, pleural, parietal, pedal and visceral ganglia is consistent to its observed widespread distribution in Aplysia CNS (Lloyd, Kupfermann et al. 1987). While its function in T. pisana requires further study, it is known to regulate muscle contraction in Aplysia and the heart beat amplitude in H. pomatia (with approximately 3 times more potency than the sCAPB, neuropeptide) (Price, Lesser et al. 1990). We expect that a similar function in T. pisana would be advantageous during arousal from aestivation and thus it was targeted for further investigation via transcriptome screening, comparative sequence analysis, and spatial expression analysis within the CNS. The observed distribution of Tpi-sCAPA within CNS regions provides some insight into its physiological function. Within the snail cerebral ganglia, olfactory function is controlled by the procerebrum region, reproduction by the mesocerebrum, and the metacerebrum regulates motor actions (Chase and Tolloczko 1993). Within the pleural-parietal-visceral-pedal ganglia, the pairs of pedal ganglia control the foot, the pleural controls the mantle, parietal controls the pallial cavity, and a single visceral ganglia regulates the visceral organs (Chase and Tolloczko 1993). Based on the widespread distribution of Tpi-sCAP within each of the aforementioned regions, the neuropeptide could potentially have multifactorial roles, perhaps specifically in the regulation of muscle contractions in the heart (visceral ganglia) and in breathing (pleural ganglia). A clearer understanding of the bioactivity of Tpi-sCAP could be achieved by further investigation of the function of Tpi-sCAP-containing neurons.

Our qPCR data showed that there was no significant differential gene expression of Tpi-sCAP between active and aestivated CNS, suggesting that the bioactive form of this neuropeptide is up-regulated in aestivation through some form of post-translational modification (Mann and Jensen 2003). Similar results have been observed in ground squirrels, where a variety of proteins are over or under expressed during hibernation, with no changes in mRNA levels (Shao, Liu et al. 2010). Alternatively, as happens with many hormones which are stored in large dense-core vesicles (secretory granules), the peptides could be stored until cells are stimulated to secrete the granules, ensuring that large amounts of peptide are available when needed (Kiss, Pirger et al. 2006). This is known to occur with Aplysia
egg-laying hormone (Denlinger 1972). Similarly, the sCAP precursor may be transcribed and post-translationally processed into its multiple peptides and stored within secretory vesicles. This would be beneficial considering the speed with which Theba returns from aestivation to full activity (as little as 5 min, personal observation) and also in conservation of energy required for protein production during aestivation.

In contrast, we found that the Tpi-AAP12 was of significantly lower abundance in the CNS of aestivating snails, compared to active snail CNS. However, the Tpi-AAP12 precursor transcript was significantly up-regulated during aestivation, suggesting regulation of the peptide could be post-transcriptional or post-translational, whereby the peptide would be rapidly secreted from the CNS during aestivation. This explanation is plausible if this peptide is required extracellularly to maintain aestivation or accumulates to function during reanimation. Gene expression would then be down-regulated during normal activity. This is typical for some antioxidant peptides found in snails that are required when oxygen levels increase dramatically during arousal (Nowakowska, Świderska-Kołacz et al. 2009); however, in some animals, antioxidants increase during aestivation in preparation for reanimation (Hermes-Lima, Storey et al. 1998). A challenge to aestivating animals is the damage to cells that can be caused by reactive oxygen species (ROS) exceeding the capacity of antioxidant defenses. While reduced oxygen consumption and blood oxygen content during aestivation should lower production of ROS, levels of antioxidant enzymes have been observed to increase in aestivating Otala lactea snails (Hermes-Lima, Storey et al. 1998), while they remain at constant high levels during normal activity and aestivation in Helix pomatia snails (Nowakowska, Świderska-Kołacz et al. 2009). This is presumably necessary to remove high levels of ROS generated upon initiation of arousal. Animals that are frequently subjected to oxidative stress prefer constant high antioxidant levels to counteract ROS damage, whereas when the stress is infrequent, as in most cases of aestivation, then antioxidant levels increase only during low oxygen periods and in preparation for high ROS during arousal (Storey 1996).

Antioxidant peptides may take a wide variety of unique forms, and as Tpi-AAP12 does not match with any known peptides, further work is necessary to confirm our speculation. A comparison of the Tpi-AAP12 with other helicid snails shows high conservation, yet there is little conservation with other molluscs apart from Aplysia, which may imply some association with hypometabolic-related processes.
Conclusions

In summary, we have performed an investigation of CNS peptides associated with aestivation by comparative mass spectral analysis, followed by spatial gene and peptide expression of sCAP and AAP12. This work sets a foundation to clarify precisely what these gene and extracellular CNS peptides are and what, if any involvement they have in triggering maintaining aestivation in land snails, and possibly other molluscs.

Abbreviations
AAP: aestivation associated peptide; CNS: central nervous system; FPKM: Fragments per kilobase of exon region in a given gene per million mapped fragments; GO: gene ontology; MS: mass spectrometry; FCAP: feeding circuit-activating peptide; ROS: reactive oxygen species; sCAP: small cardioactive peptide.

Competing interests
The authors declare that they have no competing interests.

Author contributions
KJA carried out the experimental analysis, constructed figures, tables and drafted the manuscript. BR analysed immunocytochemical tissues. TW carried out mass spectral proteome work. AVK helped with qPCR and to draft the final manuscript. KBS and SFC conceived the idea and obtained funding for the experiments and drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements
This work was supported by grants from the Australian Research Council (SFC) and the Grains Research Development Corporation (KA). We thank Dr Alun Jones (Institute for Molecular Bioscience, the University of Queensland) for advice and assistance with the tandem mass spectrometry
Additional Figures and Tables
Figure S4.1. Immunolocalisation analysis of *T. pisana* sCAPA in active CNS.

Figure S4.2. Immunolocalisation analysis of *T. pisana* AAP12 in active CNS.
Table S4.1. List of all protein sequences used in this study.

<table>
<thead>
<tr>
<th>Protein Family</th>
<th>Unigene/Contig</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tpi-LASGLV-like</td>
<td>Unigene422_All</td>
<td>Tpi-LASGLV-like Unigene422_All</td>
</tr>
<tr>
<td>Tpi-pleurin</td>
<td>Unigene77450_All</td>
<td>Tpi-pleurin Unigene77450_All</td>
</tr>
<tr>
<td>Tpi-enterin</td>
<td>Unigene71958_All</td>
<td>Tpi-enterin Unigene71958_All</td>
</tr>
<tr>
<td>Tpi-sCAP</td>
<td>CL6374.Contig2_All</td>
<td>Tpi-sCAP CL6374.Contig2_All</td>
</tr>
<tr>
<td>Tpi-FCAP</td>
<td>Unigene93548_All</td>
<td>Tpi-FCAP Unigene93548_All</td>
</tr>
<tr>
<td>Tpi-sensorin</td>
<td>CL5976.Contig1_All</td>
<td>Tpi-sensorin CL5976.Contig1_All</td>
</tr>
<tr>
<td>Tpi-buccalin-1</td>
<td>CL7089.Contig4_All</td>
<td>Tpi-buccalin-1 CL7089.Contig4_All</td>
</tr>
<tr>
<td>Tpi-buccalin-4</td>
<td>Unigene77331_All</td>
<td>Tpi-buccalin-4 Unigene77331_All</td>
</tr>
<tr>
<td>Tpi-LFRF</td>
<td>Unigene21004_All</td>
<td>Tpi-LFRF Unigene21004_All</td>
</tr>
<tr>
<td>Tpi-AAP1</td>
<td>CL16802.Contig2_All</td>
<td>Tpi-AAP1 CL16802.Contig2_All</td>
</tr>
</tbody>
</table>
>Tpi-AAP2 Unigene50087_All

MPSIAVCFPLVFAILHAVPAEEGSSLAGASPRHGTGLLLRKRSELAALRARPSISYSEAIAPAKRDDGYWIWMPAAGYMPIP
TDDAPEQQGGRDSMGNLLRHYG

> Tpi-AAP3 CL12909.Contig2_All

MRGLHJKYDNEKDRKRTGTYLNLNENTLDQSQGSLQSDSFENNEEEDIKRFRMGRGLKKVNEENDQVSVDKRFRMFGRMLYSKKESEDD
TSTRDTRKRFMFRGFKSKKDNLTDQMNIMNSANSFDAYETDENFLKQFKLQGKLRYANYAKDHVMGESVAD
KRFMIKLVRKLVKQSDLDQFEKDKYKVLKKDSQKETDSTSNDFKDFAAKNDSLQDKMRFLVRKESQGPLRKEKETRFDDR
DPDLRLRTLEKNGKFQGLSFLETSNDAIWHLPAYDSFEHTHNFMRGLGRYSKRKDYDLDSSNEFMDIDQFSEFENQQDISRPFMK
GFLHYSNGKFRHGYLTKPEESSSSNPGADKRFKGLGLMYTRPVGKFRMGLQQYGRNSDFTWEEKFRMGRMLPNKFKMLEKRF
MRGLGRFAKRPQDGESQSYESYHRKKPSDYLNVFTERQILSPRSYHQSFISSNSQDQRTQLVL

> Tpi-AAP4 Unigene90869_All

MTSSAIALLPVLLFCHLVKADSQPSIQLTLSGGLEKDKHALAIRKDEIESPSDYLISNLSWIIDEDDLARFLDEQPVERTPEKSSS
SSDRIRHKHEKNKLPTQFNHQIFGKRHFQNLOSTSISLSSEDTSSGLKRLSIRNLKTDYVSVPETDSSDLYAIRGTFQSEGGRD
RFKRRMHNFVFRGKRYSEDNTNPQVQIFGFPVGDSDNQAVGKVRYSFKVFRIGFL

> Tpi-AAP5 Unigene50081_All

MGVDVHSDWVKMFKEPKSVSLQLAQAFDQDLVSSHPVYPVGPDEINEIFDTISYAAGSIRMRFMRGLFTDFQGRTIYLRSSRF
GAADFHHDLWLSATQALSEKSLPKNVSKIMDTLTQPLMNFPVTERTDNLLNMIRATQARYLDPLATDPGYVSFENYTVWVPLT
FTSSNLSEFNQTSDVVRWMNSQIRERTFSTSDNLPSISDPNWLAVNQQQYGYRYNVHNDNLALSQRTLHTNLVPISTNAQIN
BDWNLKAHHILGTLAERKHYEDYWYARSELYWNESMLDSLGHYFPQAFMRKLHRATFDNLLGLANISTDHNVNSTQ
ALIAQASQYGISCLDEVSRQYLWNPNPHNPINNIK

> Tpi-AAP6 Unigene59163_All

MKFLPVKPEQQPPLYDGKLNSLMDKTDVIRLKNAHFRSHCITGVTKLQSHQDQSESNNKKEIFGLLERELNFKQIRDGTKLAK
RKAKSTKETQNIQAFREYMKMYEEREKRLPPDNDNYHEMLMTKLKREESAASVIDTNTETKQTFPGFMELGQKSLKSDDLLPCG
YPSAPSGFLPSVLKLLHPVPAQTELKDLTPLKYPPETLSQIALIISYKVRQGTQYDSQYDEVDASLRSRMYTDFDDLHRSLVLEVELT
SDDEEKRKDLKLDNGLVHEPEKLSNKRSTAFEVTTRTVPTVKQGRRKIDLPSEDVNEADAVSSQDLASKSSLMAKIQHLHPSSS
ESELEDISKSCPDKLATNSVYTRHPYPRFAPKAKLISPELVNEVASQPSRPSTLPPSTLPNFLRFFVTWEHFKFYPNCTNLPKPW
TSDNFINMCMLRIADWDGDVKVQVIAWYTYEDEGSLAMESDQVATDARVAVLNLHNASAPCSVQKFKQLALSALALTFLKEKE
AVVEMVMQLFGDLTDISRVLADAMNLLIGDPHRUQKFQELDSDWVWIEEALNRNKLHMCSEWLEWTMSYKIRLQDSDYL
NHGHSQGLRSMKVSQGSLRQRQAPEARASALQPSVTKSPPTVAQVDTRSDRQDSSAHTSMVNLHSLTSLKASVYD

> Tpi-AAP7 Unigene55743_All

MGAVISVPAAAVLACLWLFYESLQVEQVNPVPNKENNUNNNTLKHAGHRDKSTLDLRKLDLSIAGHSSFGAFGRQCGGSPGRPAZD
EESEFQKRFDRSDNRSQSGFAPREEFLRKKSDYDFLADKDNVYENIDSEPPSEFLILETNGSLREVPSIDANDLNEAVGFTRRGNF
DGNFNTDTDEAATLSLTPFSSAQNLQIOTIPTSSAESTVESVAREEQEPQLPPVPLRLKRHKVFEDENTASHIRKFKDISSNHFAAF
GKRPLTENNEYLLPDNLNYSQNPSSPESPLPEDSRVQNVNHSYQHDFQGKRROFKDQGSKAFAFDGFKDROQNNSNVNKRRFDSI
AGDSSGRFGRFKQDEERKRFDSISGHSFGSLKGRKDDVEKRRKRFDSISGHSFGSLKGRKDDVEK

> Tpi-AAP8 Unigene121884_All

IFDDRVSLEPYKVWNLHNVKVDAGVYQCQINTDPVQREVVVLHVNEPPKIMDSLSSADSTTVREGDTVELVNVTGLPPS
VTWYTVNTGFNDHEDDM

> Tpi-AAP9 Unigene29221_All

YFVTPFPSSTRNLLENKPLEQESTLTDPFKQLSSPNVGCLELRTKFTPSFSSFROQNSKMEIQUISIFILVSCIFVSAAAPHRQRQDLQLH
SLVNIKRDLNRHISPSLNTNDTENKVAHLNLNSDALVEAAEDEKPEQKKGVMKSAQGLPKRQGAWSDYGLGGRRFGRK
NYGDYIYGIGGFGRDQVHVHDLSNDADVSS

> Tpi-AAP10 Unigene2101_All

DVCEKGEICTNDLFSGCSTIFSRQHRIRSRESYELGTRIERLDEERRLLIKAGFTWKDAYQTCTQNLTLLFESQGRIFNRKLCCEALV
STDWQTQSANEAIRVVLQALASAGAVDRDRDEAANQNEEEGEEEEELNLSNKRDEIVOYHKYFKDSTGQKLRDPDDESGYENSSDDP
KYTVVYEPQLVQVEKNTYPPASENDGDLREEEPSEPLARDLIFLQKNNFPDQAQKSQAKQDVQKSEGRPNNFIDJENLVLTPE

133
DELNLEGLEGTISSPKDLDGQQQRQLVFIKGLIMGLSSNSKTFNLINDDDEEADAAHSHTPTLDHNDIAEDPDGPDGDVDVLPLVGETTPADOQILDDVIDDLNAERLKLAIKNPDLQVQKTDIDHVDTEKNEGDDHADTHYMELYSLKEVNGSSATQFVDNLTEI
LRLPPGLVQMLDVGPRMISLKLVRTDPDGTVDAEIKIAAEQAQNANARLEKIRGITVEGAVGRGDNVRVEKQDNRVYVTVLFLG
TIAVGVVLAVGVYILCKHRHRSQKQLAQLATNGDGEASKDQDCLCQRMRQQGKSSEKPELHLAASRIGVSEQVQVRSPSSRSSSTSSW
SEEEAATNMIDSTGHLSSYMEDHLKQDRDLIREAEASYAVDSFTKQVDATQSMRKNRFADHPYDHRSVLLSSNSNVSNSD
YINANFTTDHPRNPAYIAOGPMLHTVPADVFMWQEMWQSVVVMMLTVLENGTSLCHRYWPEEGSDLHYIHIEVHLVSEIHWCDD
YLVRSSFYKLNQNTIERTTVTFQHFLTPELSVPSAIKALLDFDRKVKNFSFRRSCPVLHCVHSDCGCRTGTYCLIDMLVNLKMAKGT
KEIDMAATLEHIRDQRMAMVKTKEEQFQFSLAASVEEHAILKALPQ

>Tpi-AAP11  Unigene45991_All
MNCLLDTWVCFCTIVLMTGHVFRAEYDSDRDESNKQDNGAVRLGGLQMLRLGKRSTSNVDSTDDHETVPSVBRRSVDPPDNSH
LSPNDQVAVLASIFDQPREESRRQPPLPRYGRDSSSNISGRLLADALSDIADPFIQSSNNARFFFRPAPRGGYRRFSFLGERT
YGDFISQDASDRTFAPFRFQNXDIVPIVHAKAVPRPRIGRFQNDQSGASFQPSA

>Tpi-AAP12  CL19954.Contig1_All
MGSRFFMSLLSMMLVSCQVASFTISSRRHERSTDLAEALAILQRQRRAVTDYLGSDDRLNNDPYALEEADWLTADVGDSNSRN
ALSDWESNRGRFDVQAIPYEVPVVRNIRISPVAIERSQEDLIIDFAEDADGVSQPKINPIPEKKKEEMSLVGDKGVVQQK
SPPPVDIEATVPADSDNSGIDLTITKDEHALKMAVDKLQQAQVFSEKTETPEARTYKAVESNGEPKFNAITITQPSKE
ELQAIFDJEESPLVKEETEVEPNSAGGERIIEQQQVMNSSNPANKEKELKNEELDINIASELEDNIEAEEKHAEERDLIFNAVNLNEA
ERSNGAJIERKWLNRKFPQNPTLKTKRAAUPFVPVQRPYNDDDLATEENAQNNDALDDJSALIPNAEQFVAKILEQDEVDNLKI
VSRLEDLENDVTDALSENATLQKEGSSYTDMEFESLQLQAIREEALQKLNKEEDSPYLYEVEVKGRLPWRKGRQOVDADISANF
YPSQRQASYAVLPEEERSPLGNDIQIIEVEPTDEQPDADVASFSMSNANDECPAQVSTNCELADLVLSPYDEARLSCLNLH
EMCYACGKLSQVQCDLVYREAAASALCGKDNCCVESEIFRLTMKLNRYIPHSQPLCRSTCAAQFLGML

>Tpi-AAP13  CL4923.Contig1_All
QGFHWSGGKRSDANYLDDEKRSKLGFNSWGGRSDDANYLDDEKRSKLGFNSWGKRGAISNNLDDDELQKRKSLGFHSGW
GKRTQNIENFYTDEEREKRSKLGFHSGWGGKRCDENVNLEEIEKRGKLGFSWGKRGDRIEASLPLITRNTHEGLDESNPREGINL
GVDSDDRNVDKHLRKKASLGLNSWGEHEHTFDIDKESKRKNGFGDSVDEEKSRLGFHSGKGK
Table S4.2. Mass spectral analysis of CNS showing MS/MS fragmentation data for sCAP and AAP-12 precursor peptides.

Tpi-sCAP-1

<table>
<thead>
<tr>
<th>Peptide Sequence</th>
<th>MS/MS Fragmentations</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEMSLQBUSI SMSLLVILVC SAAEMNYLAF PPRM SGGYLA PPRM RSQAK SETAAEFGNC CGVGLKSELV IGHDGKEELR</td>
<td>b1, y2, y3, y4, y5, y6, y7, b8, y8, b9, y9, b10, y11, b11</td>
</tr>
<tr>
<td>PVCTVNAECQ QGLRELADQK PGTYIYMCV PDLPFSLESN GSPSEVLKKL KTLMKK</td>
<td>b1, y2, y3, y4, y5, y6, y7, b8, y9, b10, y11, b11</td>
</tr>
</tbody>
</table>

Tpi-sCAP-2

<table>
<thead>
<tr>
<th>Peptide Sequence</th>
<th>MS/MS Fragmentations</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEVSVSRTSL FAAVLLIMVC GAAENMYLAF PPRM SGGYLA PPRM RSQGQ SAAVEGVASC CGLGLKSEFV IGQNGKEELD</td>
<td>b1, y2, y3, y4, y5, y6, y7, b8, y9, b10, y11, b11</td>
</tr>
<tr>
<td>VVCAAPAGCC EGLREIVEQS STGTSYSLCV PDLSFHQDTK ESAEELRKF KVLTRM</td>
<td>b1, y2, y3, y4, y5, y6, y7, b8, y9, b10, y11, b11</td>
</tr>
</tbody>
</table>
Chapter 5: Discussion

The investigation of the transcriptome from the land snail *Theba pisana* as detailed in this study, expands our understanding of gastropod biology and in particular the metabolic components that may regulate aestivation. Initial deep transcript sequencing of the central nervous system (CNS), foot muscle and hepatopancreas, followed by *in silico* predictions and mass spectral analysis enabled the identification of neuropeptides, including those previously unknown. This information greatly enhanced our ability to utilise this species as an experimental model in neuroscience and metabolism research, with the potential for development of molecular tools that could manipulate the physiology and behaviour of this agricultural and horticultural pest species.

Development of a land snail transcriptomic resource

In Chapter 2, comprehensive transcriptome libraries representing the CNS, hepatopancreas and foot muscle were obtained from deep RNA sequencing, using RNA isolated from the tissue of active and aestivated snails. This diversity of physiological stages allowed for variance in gene expression levels which can occur as a result of different animal behavioural and physiological states. A comprehensive analysis of tissue gene transcripts was carried out, with over 300 million raw reads providing approximately 200,000 unigenes for each tissue. As no annotated reference is currently available for *Theba*, or any other helicid snail, *de novo* assembly of the transcriptome sequences (denoted unigenes and contigs) was necessary. The majority of genes most closely matched by BLASTp to oyster, *Crassostrea gigas*, which within the NCBI database has the most comprehensive inventory of molluscan-type genes; primarily due to the recent release of a genome and multi-tissue transcriptomes (Zhang, Fang et al. 2012). With *C. gigas* belonging to a different taxonomic group, the bivalves, the sequences derived from this study and deposited into the NCBI Genbank SRA archive no. SRP056280, ensure a far better representation of helicid gene sequences in the NCBI databases.

Marine and freshwater gastropods, including *Aplysia* and *Lymnaea*, have for around 50-60 years been used for neurological research. One study involving each of *Lymnaea stagnalis* (Feng, Zhang et al. 2009, Sadamoto, Takahashi et al. 2012) and *Aplysia californica* (Moroz, Edwards et al. 2006) have provided a large number of CNS transcripts. Additional gastropod investigations have provided unpublished CNS transcript data for the Queen conch.
(Strombus gigas; SRX265463) and cone snail (Conus consors; SRX980532). And now our study has produced a very comprehensive list of T. pisana CNS transcripts, including 220,602 unigenes.

With the CNS, hepatopancreas and foot muscle likely to regulate critical events in the initiation and maintenance of aestivation, our transcriptomes and associated transcriptome-derived protein databases provided a platform for in-depth gene- and peptide-associated hypometabolic studies, similar to those described in other animals (Bell, Dawson et al. 2012, Ito, Okada et al. 2012, Lama, Bell et al. 2013). Several of those neuropeptides found to be significantly up- or down-regulated in the CNS or hemolymph of hibernating Helix pomatia (Pirger, Lubics et al. 2010) were also identified in our Theba investigation. Additionally, we found that the T. pisana CNS contains as many as 5504 precursor proteins that are destined for secretion, based on the presence of an N-terminal signal sequence and lack of a transmembrane domain. Comparison of these with derived molluscan precursor neuropeptides from Lottia gigantea (Veenstra 2010) enabled us to establish a final list of 35 known neuropeptide precursor genes shared between both Lottia and Theba (Chapter 2).

This Ph.D research was primarily interested in peptides associated with hypometabolism, so transcriptome analysis focused on that area, leaving a vast amount of data to be interpreted. Other T. pisana-related research has utilized the CNS, hepatopancrease and foot muscle transcriptomes, thus demonstrating that such a resource has widespread investigative potential. A greater understanding of the underlying genes has enabled preparation of ‘Multi-tissue transcriptomics for construction of a comprehensive gene resource for the terrestrial snail Theba pisana’ (Appendix I). This provides a more in-depth analysis of identified genes ontology and annotation, as well as providing community access with the online resource termed ThebaDB (http://thebadb.bioinfo-minzhao.org/).

This transcriptome resource has also provided a foundation for exploration into reproductive-associated neuropeptides in T. pisana. For example, using the CNS transcriptome, two conopressin isoforms were identified, followed by expression analysis (see ‘Characterisation of two conopressin isoforms in the land snail, Theba pisana) (Appendix II). Conopressin is one of a number of neuropeptides involved with reproduction in molluscs, and was found in different tissues including the reproductive tissues in both immature and mature Theba. The conopressin transcript was identified within Chapter 2, although due to a low level of overall conservation with other species conopressin, had been overlooked.
Analysis of non-neural tissues

The hepatopancreas and foot muscle are major organs for lipid metabolism and storage in gastropods, and play a crucial role during periods of starvation and hypometabolism (Böer, Graeve et al. 2006, Lama, Bell et al. 2013). Metabolic enzyme genes, as expected, dominated annotated transcripts associated with the hepatopancreas (see Chapter 2), which is a major source of digestive enzymes and is involved in nutrient absorption, food storage and excretion (Barker 2002). Gene transcripts encoding an epiphragmin protein, a major constituent of the epiphragm mucus that seals off the aperture during hibernation and aestivation (Li and Graham 2007) were found in the foot muscle. The muscle transcriptome may be useful in future studies of muscle disuse atrophy which may be a problem during prolonged periods of hypometabolism. We additionally found a number of enzymes key to metabolic depression in other gastropod species present in T. pisana transcriptomes, including pyruvate kinase, phosphofructokinase and glutamate dehydrogenase (GDH) (Brooks and Storey 1992, Ramnanan and Storey 2006). These are likely targets for future metabolic studies in this species.

We have utilised the new hepatopancreas transcriptome to help reveal enzymatic pathway genes considered to be important based on our comparison of hemolymph profiles between active and aestivated snails. In that study, we have found a significant number of metabolites differentially regulated, including specific phospholipids, which may be used as a source of energy during aestivation. The site for phospholipid synthesis is the hepatopancreas, thus we were able to utilise the hepatopancreas transcriptome to establish the presence and relative expression levels for phospholipid synthesis enzyme genes. All findings from the metabolite analysis are presented in ‘Global metabolite analysis of the land snail Theba pisana hemolymph during active and aestivated states’ (Appendix III).

Identification of a land snail neuropeptidome

All peptides present in the T. pisana CNS were identified by extraction followed by liquid chromatography-mass spectrometry (LC-MS) analysis, before identifying those differentially expressed between active and aestivating snails. To extract CNS peptides, all aestivating snails were immersed in liquid nitrogen to ensure they did not begin arousal (potentially altering endogenous peptide levels), while one experimental group of active
snails were also killed by immersion in liquid nitrogen, and another experimental group of active snails were killed by rapid dissection.

We had previously identified 35 neuropeptide precursors in our multi-physiological CNS transcriptome (see Chapter 2), and LC-MS analysis of the CNS confirmed the presence of 13 of these. The absence of the others is most likely due to abundance below the detection limit for the MS. A further 6 were identified as known molluscan neuropeptides, but were not previously identified by us due to the relative divergence with any peptide in Lottia. A further 13 peptides were found which do not correspond to known neuropeptides. Using the CNS derived from snails killed by immersion in liquid nitrogen, we identified 19 peptides that were differentially expressed, with only two peptides being up-regulated in active snails. Our sample set using snails killed by dissection identified a total of 22 peptides differentially expressed, 4 of which were up-regulated in active snails. Those peptides that were up-regulated in aestivating snails included the molluscan neuropeptides for *T. pisana* small cardioactive peptide (sCAP), feeding circuit-activating peptide (FCAP), enterin, pleurin, sensorin-A, a LASGLV-like peptide and buccalin.

The majority of the remaining peptides identified as differentially expressed, besides buccalin, match to precursor sequences that resemble either uncharacterised peptides or unknown proteins thus we named these aestivation associated peptides (AAPs).

In this research, we identified 5 different buccalin peptides that were not present in our original 35 mollusc neuropeptides. These are cleaved from two *T. pisana* buccalin precursor proteins corresponding to Tpi-buccalin-1 and Tpi-buccalin-4, and are up-regulated during snail aestivation (see Chapter 3). The neuropeptide buccalin is involved in feeding in *Aplysia*, but distribution throughout the CNS of a number of molluscs implies that it has pleiotropic functions. Tpi-sCAPA was also found to be up-regulated in the CNS of aestivating snails in our study. This contrasts a study of hibernating *H. pomatia* showing that sCAPA is more highly expressed in active animal CNS (Pirger, Lubics et al. 2010). This could be due to differences in the extraction method used or possible differences in the physiological states of aestivation versus hibernation. Similar to buccalin, the sCAPA has also been found throughout different regions of the mollusc CNS, also implying multiple functions for this neuropeptide (see Chapter 4).
Temporal and spatial expression of Tpi-buccalin-4 and sCAP in the CNS

_Tpi-buccalin-4_ and _Tpi-sCAP_ appear to be two of the most highly expressed transcripts within the CNS based on our RNA-seq. We speculated that with previously reported roles in muscle contraction and other functions, both buccalin and sCAP were excellent candidates for critical roles in the regulation of metabolic events in snails and as such, were targeted for further investigation.

Absolute quantitative PCR (qPCR) was performed on CNS samples from active and aestivated snails to determine relative gene expression of both _Tpi-buccalin-4_ and _Tpi-sCAP_ transcripts. There was a trend for both genes to be up-regulated in CNS of aestivating snail however this was not statistically significant (Chapters 3 and 4). The lack of differential gene expression suggests that these neuropeptides are up-regulated in aestivation not by gene expression, but by some other mechanism, such as post-transcriptional regulation by miRNA (Bartel 2009, Carthew and Sontheimer 2009). Similar results have been observed in ground squirrels, where a variety of proteins are over or under expressed during hibernation, with no changes in gene mRNA levels (Shao, Liu et al. 2010). Alternatively, and perhaps more likely, the peptides may simply accumulate and be stored within neuron secretory vesicles during aestivation but not be utilized until the snail returns to normal activity. There are neuropeptides such as _Aplysia_ egg-laying hormone (ELH) which, in preparation for egg laying, are produced and stored in secretory vesicles of the bag cell neurons (Fisher, Sossin et al. 1988). ELH is then secreted into the hemolymph when stimulated by some other hormone. This may also be the case with the _T. pisana_ neuropeptides we were investigating, as this would ensure rapid release upon reanimation, which _Theba_ can complete in as little as 5 minutes (personal observation). Storage of peptides until required would also save energy whereby rapid protein synthesis would not be required. There are a number of other possibilities including decreased turnover of the mRNA during aestivation, or preferential translation of the transcript while other transcripts are suppressed.

While buccalin and sCAP were up-regulated in aestivation, _Tpi-AAP12_ was significantly down-regulated, compared to active snail CNS. However, qPCR indicated significant up-regulation of the _Tpi-AAP12_ precursor transcript during aestivation (Chapter 4). While this could indicate regulation of the peptide by post-transcriptional or post-translational modification during aestivation, it could also suggest that this peptide is required extracellularly for aestivation maintenance or that it is required during reanimation and accumulates outside the CNS. This would enable the down-regulation of gene expression
during normal activity that we observe. The most likely explanation for this gene/peptide expression pattern is the possibility of this being an antioxidant peptide, required when oxygen levels increase dramatically during arousal, as seen in *Helix* (Nowakowska, Świderska-Kołacz et al. 2009), and therefore the peptide increases during aestivation in preparation for reanimation (Hermes-Lima, Storey et al. 1998).

Spatial expression of the buccalin precursor gene using in situ hybridisation and visualisation of the Tpi-buccalin peptide expression using immunolocalisation, showed high levels of both precursor gene and peptide in a number of areas throughout the CNS (see Chapter 3). The relatively wide distribution in the cerebral ganglia, and the subesophageal ganglia is consistent with the widespread buccalin distribution observed in other molluscs.

Further work would be required to determine exactly how many buccalin-expressing neurons are present and their location within the various ganglia. In addition, it would be of interested to precisely assess buccalin concentration within individual neurons during aestivation and normal activity.

In situ hybridisation of *T. pisana* CNS using a DIG-labeled riboprobe specific for the Tpi-sCAP was unsuccessful after numerous attempts, most likely due to either the low expression level of the transcript or degradation of mRNA during tissue processing. Yet, as with buccalin, immunolocalisation showed high levels of Tpi-sCAP throughout the cerebral, pleural, parietal, pedal and visceral ganglia regions of the CNS (Chapter 4), which is consistent with its observed widespread distribution in *Aplysia* CNS (Lloyd, Mahon et al. 1985, Santama, Brierley et al. 1994, Li, Moroz et al. 1998). Distribution of Tpi-sCAP in the CNS is similar to the aforementioned regions in which buccalin is expressed, and therefore similar multifactorial roles including heart rate control and breathing, are likely. A clearer understanding of the bioactivity of Tpi-sCAP could be achieved by further investigation of the function of Tpi-sCAP-containing neurons.

Of particular interest to us, were the high levels of both sCAP and buccalin peptides in the visceral ganglia region. In the land snails studied, the visceral ganglia is known to regulate the heart (Chase 2000), thus, the presence of buccalin and sCAP within this region is consistent with a functional role in controlling heart rate, which is known to slow dramatically during aestivation (Storey and Storey 1990).
AAP12 and potential antioxidant activity

*In situ* hybridisation using a DIG-labeled riboprobe specific for the Tpi-AAP12 was unsuccessful, also likely due to either low transcript expression level or degradation of mRNA during processing. As with buccalin and sCAP, immunolocalization showed widespread distribution of the Tpi-AAP12 within regions of the cerebral ganglia, pleural ganglia, parietal ganglia, pedal ganglia and visceral ganglia. While this might imply a number of functions, as in the case of buccalin and sCAP, we speculate that AAP12 could be an antioxidant peptide that needs to be widely distributed throughout the body.

It would be of interest to test the AAP12 peptide for antioxidant activity to confirm antioxidant function. There are a number of assays currently available that could be used to test the antioxidant activity of this peptide (Najafian and Babji 2012), mainly based on two methods, hydrogen atom transfer (HAT) reactions and electron transfer (ET) reactions, although these would only test the radical (or oxidant) scavenging capacity of the peptide, and not any preventive antioxidant capacity. These assays include, for HAT, inhibition of induced low-density lipoprotein autoxidation, oxygen radical absorbance capacity, total radical trapping antioxidant parameter, and crocin bleaching assays, and for ET, the total phenols assay by Folin-Ciocalteu reagent, trolox equivalence antioxidant capacity, ferric ion reducing antioxidant power, “total antioxidant potential” assay using a Cu(II) complex as an oxidant, and DPPH (2,2-diphenyl-1-picryl-hydrazyl-hydrate) (Huang, Ou et al. 2005). In some of these, the assay involves adding the antioxidant to an oxidant and observing a colour change that occurs when the oxidant is reduced. Peptides to be tested could be either synthesised or naturally occurring peptides extracted from snail CNS and isolated by size, using high-performance liquid chromatography (HPLC).

A variety of antioxidants, including peptides, are of interest for a number of reasons, including the prevention of medical conditions such as cancer, cardiovascular disease, Alzheimer's disease and ageing, caused by oxidative damage in the human body (Halliwell 1994). Antioxidants are also effective as food preservatives and there is interest in finding and using naturally occurring rather than synthetic products (Meyer, Suhr et al. 2002). These have been investigated from a number of different species including plants. If our peptide is found to have antioxidant properties, given the speed that antioxidants are required to work at in snails during arousal, we believe it is likely to be a very efficient antioxidant.
Involvement of microRNAs in aestivation, including neuropeptide regulation

There is little doubt that miRNAs are involved in gene regulation, and this includes genes involved with hypometabolism (Biggar and Storey 2011), including in hibernating squirrels, (Dubuc and Storey 2008), in wood frogs during freezing (Biggar, Dubuc et al. 2009, Palyga 2011) and of more relevance to this study, in sea cucumbers, where over 40 differentially expressed miRNAs were found during aestivation (Chen, Zhang et al. 2013). While gene regulation by miRNAs involves silencing of target genes, of particular importance is the ability of miRNA regulation to quickly reverse the silencing when animals arouse back to activity (Dubuc and Storey 2008).

As a preliminary investigation towards establishing a role for miRNAs specifically in neuropeptide regulation, I prepared T. pisana CNS tissue from both active and aestivated snails for small RNA analysis. The results showed that the majority of total small RNAs were found in both active and aestivating snails, with only a small percentage unique to each group. On the other hand, almost 50% of unique small RNAs were found only in active and 40% found in aestivated CNS (Figure 5.1A). The majority of small RNAs in both active and aestivating CNS are miRNAs (Figure 5.1B), with the majority being 22 nucleotides in length (Figure 5.1C), a feature also found in aestivation-associated miRNAs (Chen, Zhang et al. 2013). Analysis of miRNA expression showed that a large number are differentially expressed (Figure 5.1D).
Figure 5.1. Small RNA analysis A. Venn diagram showing a comparison of small RNAs in active and aestivated snails. B. Types of small RNAs and their distribution in active and aestivated snail CNS. C. Size distribution of total small RNAs in nucleotides D. Scatterplot of up- and down-regulated miRNAs

From these differentially expressed miRNAs 10 (5 up-regulated in active and 5 up-regulated in aestivated snails) have been chosen for further study; including those with known functions in other species and whose functions appear to be involved in metabolic control (Table 5.1). These preliminary data requires further qPCR to confirm differential expression, and its binding properties to mRNA could establish precisely which gene transcripts the miRNA bind to and regulate. We speculate that some may be important regulators of neuropeptide genes.
Table 5.1. List of most highly differentially expressed miRNAs.

<table>
<thead>
<tr>
<th>miRNA-name</th>
<th>ACTIVE</th>
<th>AEST</th>
<th>fold-change (log2 SRAES1/SRACT1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>upregulated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>let-7d-5p</td>
<td>169</td>
<td>1677</td>
<td>3.315471</td>
</tr>
<tr>
<td>miR-125b</td>
<td>35</td>
<td>3349</td>
<td>6.584931</td>
</tr>
<tr>
<td>lin-4</td>
<td>2335</td>
<td>149168</td>
<td>6.002057</td>
</tr>
<tr>
<td>miR-2a</td>
<td>226</td>
<td>7267</td>
<td>5.011651</td>
</tr>
<tr>
<td>miR-124b</td>
<td>2126</td>
<td>6779</td>
<td>1.677616</td>
</tr>
<tr>
<td>downregulated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-34</td>
<td>3320</td>
<td>840</td>
<td>-1.97804</td>
</tr>
<tr>
<td>miR-84-5p</td>
<td>792933</td>
<td>194</td>
<td>-11.9922</td>
</tr>
<tr>
<td>miR-745</td>
<td>36671</td>
<td>69</td>
<td>-9.04914</td>
</tr>
<tr>
<td>miR-240-5p</td>
<td>17532</td>
<td>8736</td>
<td>-1.00026</td>
</tr>
<tr>
<td>miR-8</td>
<td>2258</td>
<td>2</td>
<td>-10.13645427</td>
</tr>
</tbody>
</table>

Development of a bioassay to test neuropeptide function

Future study will require in vivo bioassays to characterize the functions of peptides identified in this research. It would, therefore, be helpful to define a metabolic marker that could definitively categorize snails as being in aestivation. In this research, we were confident that aestivation had been accomplished prior to analysis, however, further temporal expression of genes and peptides across the continuum between aestivation to activity would be helpful. To potentially achieve this, oxygen consumption (and thus metabolic activity) was measured. The experimental set-up initially involved housing 40 well fed and hydrated, fully active snails individually in 60 ml jars. Aestivation was then induced in 30 snails by withholding food and water, while the remaining 10 control snails continued to receive food and water daily. For each of 30 consecutive days, jars were completely sealed to allow for measurement of oxygen levels within, using a Presens Fibox 3 optical oxygen meter. Snails were left for 1 h before measurements were repeated. In the time that oxygen consumption was monitored we encountered two problems, (1) at times the snail oxygen consumption was very low, outside the accurate range that the equipment could measure, and (2) the movement
involved in sealing the jars and taking the measurements frequently would arouse some snails with the likelihood for a resultant increase in their metabolic rate. The experiment was then repeated with two modifications to potentially overcome these problems. First, jars were modified to allow oxygen measurements with negligible movement, and second, the time between beginning and terminating measurements each day was doubled. However, similar results to the initial experiment were observed, particularly in relation to snails waking, so it appeared that this mechanism for oxygen monitoring was not a suitable system to determine aestivation in *T. pisana*.

Another potential physiological marker to determine the metabolic state could be to determine the level of known metabolic enzymes. There are a number of metabolic processes known to be up- or down-regulated during aestivation and the enzymes controlling two of these processes were selected for testing for use as indicators of aestivation. Protein kinase A (PKA) and protein kinase C (PKC) gene sequences were found in our transcriptome and primers were designed for qPCR. During initial testing of these primers using both PCR and qPCR, we were unable to eliminate contamination in the reactions, which gave us a false positive result in our negative, no template control. Further investigation found that our PKA and PKC were highly conserved across a wide range of organisms, and after extensive experimentation we were unable to identify or eliminate the contamination inherent to our laboratory.

Perhaps the most promising method for measurement of snail change in metabolism is the measurement of heartrate, which is known to decrease dramatically during aestivation. Monitoring the heartrate is possible using an infrared emitter and detector, fixed directly on the shell covering the heart. However difficulties were encountered since the heart moves relative to the sensor, as the animal withdraws deeper into the shell during aestivation, taking the heart outside the physical range of the sensor. We have established that the *Theba* heartbeat is visible upon high intensity light exposure through the animal so another approach was to monitor heart rate manually, through video recording and data analysis. While this method was not suitable for determining aestivation, this could be used to monitor change in heartbeat following administration of target peptides.

**Implications for applied research**

The identification of bioactive peptides, including neuropeptides in this study could lead to development of novel pest control methods for not only *Theba*, but also a number of
other invasive land snails that have become pests. Current pest control methods used in the South Australian grain fields consists of either commercially available snail poison or a regime of ‘burn ‘em, bash’em, bait ‘em’ (http://www.grdc.com.au/GRDC-Snails-BashBurnBait), which involves burning grain stubble (which was common practice, but can increase soil erosion and is no longer recommended other than for snail control), mechanical crushing of snails using a variety of devices towed by tractors during the summer months while snails are aestivating, and also broadcasting poison. While these practices have had some effect in the short term, they do not appear to have substantially decreased the snail population in the long term. Findings from this study could result in design and production of a novel pest control method using mimetic peptides that unlike naturally occurring peptides are not broken down by enzymes internally.

Mimetic peptides bind to the target receptor and so disrupt the action that the peptide would normally regulate (Wrighton, Farrell et al. 1996, Nachman, Kaczmarek et al. 2004). Because these mimetic peptides are not regulated by enzyme degradation, they can inhibit essential cellular functions. Mimetic peptides have been shown to be effective in a number of areas. Cone snails, for example, have thousands of venom peptides (conopeptides) in their venom gland and several have been found to block receptors in pain signalling pathways in humans, relieving pain (Bogin 2005, Vetter and J Lewis 2012). Several analogue neuropeptides have been investigated for use in pest management, with disruption of the insect digestive and diuretic processes (Nachman, Kaczmarek et al. 2004, Nachman and Pietrantonio 2010) and locomotion, reproduction and digestion in locusts (Lange, Orchard et al. 1995) being two such possibilities.

Conclusions

The intent of this research project has been to increase our knowledge and understanding of the control of hypometabolism, specifically aestivation in the pest snail *T. pisana*. This has been accomplished through the completion of five significant aims, including

1) A comprehensive transcriptome resource for *T. pisana* has been established, including annotation. This is currently the most extensive gene resource for a land snail,

2) Neuropeptide genes expressed in *T. pisana* have been identified, as well as an expansive list of candidate neuropeptides.

3) Neuropeptides expressed in *T. pisana* CNS have been identified using liquid chromatography-mass spectrometry of neural ganglia.
4) Those neuropeptides and other peptides in the *T. pisana* CNS up- or down-regulated during aestivation have been identified.

5) Three target peptides have been further investigated by temporal and spatial location of neuropeptide genes and their products.

This study has established a powerful resource and inferred metabolic functions for several neuropeptides and peptides. This forms a comprehensive framework for further research into land snail neuropeptides, some of which has already commenced. The limited time available within a Ph.D project has meant that a more comprehensive study of the entire group of neuropeptides, and in particular those differentially expressed between active and aestivated snails, was not possible.

Should this lead to the discovery of a peptide which controls essential processes during aestivation, then the ability to disrupt the functioning of the peptide by the use of peptide mimetics could prevent *Theba* from entering a state of aestivation during the heat of summer. There is no doubt that this would result in a very high and possibly total death rate as a result of desiccation, and be an effective pest control strategy not only for *Theba*, but also transferable to other pest molluscs.
Literature cited


Barker, G. M. (2002). Molluscs as crop pests, CABI.


Morishita, F., Y. Furukawa, O. Matsushima and H. Minakata (2010). "Regulatory actions of neuropeptides and peptide hormones on the reproduction of molluscs: The present review is one of a series of occasional review articles that have been invited by the Editors and will feature the broad range of disciplines and expertise represented in our Editorial Advisory Board." Canadian Journal of Zoology 88(9): 825-845.


Appendix I  Multi-tissue transcriptomics for construction of a comprehensive gene resource for the terrestrial snail *Theba pisana*

M. Zhao¹, T. Wang¹, K.J. Adamson¹, K. B. Storey², S.F. Cummins¹,§
School of Engineering, Faculty of Science, Health, Education and Engineering, University of the Sunshine Coast, Maroochydore DC, Queensland, Australia, 4558; ²Institute of Biochemistry & Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, ON, Canada K1S 5B6

ABSTRACT
The land snail *Theba pisana* is native to the Mediterranean region. It has become one of the most abundant invasive species worldwide through increased global movement from Mediterranean countries. Here, we present the assembled 250,848 gene transcripts of this well-known agriculture pest derived from three tissues: the central nervous system, hepatopancreas (digestive gland), and foot muscle. Sequencing of the three tissues produced 339,479,092 high quality reads and a global de novo assembly obtained a total of 250,848 unique transcripts (unigenes) with an average length of 557 nucleotides. BLAST analysis mapped 52,590 unigenes to NCBI non-redundant protein databases and further functional analysis annotated 21,849 unigenes with gene ontology. We have found that *T. pisana* transcripts have representatives in all functional classes and a comparison of differentially expressed transcripts amongst all three tissues demonstrates an enormous difference in their potential metabolic activities. Of those genes differentially expressed, they include genes associated with multiple bacterial diseases and neurological diseases. To provide a valuable resource that will assist functional genomics study, we have implemented a user-friendly web interface, ThebaDB ([http://thebadb.bioinfo-minzhao.org/](http://thebadb.bioinfo-minzhao.org/)). This online database allows for complex text queries, sequence searches, and data browsing by enriched functional terms and KEGG mapping.

Key words: *Theba pisana*; central nervous system; hepatopancreas; foot muscle; ThebaDB
Appendix II  Characterisation of two conopressin isoforms in the land snail, *Theba pisana*

M. J. Stewart¹, B. I. Harding¹, **K. J. Adamson¹**, T. Wang¹, K. B. Storey², and S. F. Cummins¹
¹ Genecology Research Centre, Faculty of Science, Health, Education and Engineering, University of the Sunshine Coast, Maroochydore DC, Queensland, Australia, 4558; ² Institute of Biochemistry & Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, ON, Canada K1S 5B6.

**ABSTRACT**

Increased understanding of the molecular components involved in reproduction may assist in understanding the evolutionary adaptations used by animals, including hermaphrodites, to produce offspring and retain a continuation of their lineage. In this study, we did focus on the Mediterranean snail, *Theba pisana*, a hermaphroditic land snail with high promiscuity and fecundity. Our analysis of *T. pisana* CNS tissue has revealed two gene transcripts encoding the molluscan reproduction-associated conopressin precursor protein. *T. pisana* conopressins share high identity with other molluscs within the conopressin peptide, but less within the cysteine-rich neurophysin peptide region. Phylogeny supports the conopressin grouping in close association with the vasopressin peptides found in other animal phyla. Our molecular dynamic models provide a most likely stable structure of Tpi-conopression-I peptide and when associated with neurophysin. RT-PCR shows that this gene is expressed in numerous tissues of mature and immature snails, including the reproductive tissues, supported by whole mount *in situ* hybridisation localisation. In summary, this study represents a detailed molecular analysis of conopressin genes in a land snail and provides a foundation for understanding its function in a reproduction and development context.

**Key words:** conopressin; CNS; mucous gland; dart sac; *Theba pisana*; land snail
Appendix III  Global metabolite analysis of the land snail *Theba pisana* hemolymph during active and aestivated states

Bose, U¹. Centurion, E¹. Adamson, K. J¹, Hodson³, M. Shaw⁴, P. N., Storey, K. B²., S. F. Cummins¹

¹ Genecology Research Centre, Faculty of Science, Health, Education and Engineering, University of the Sunshine Coast, Maroochydore DC, Queensland, Australia, 4558; ²Institute of Biochemistry & Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, ON, Canada K1S 5B6; ³Metabolomics Australia, The University of Queensland, St, Lucia, 4072; ⁴School of Pharmacy, The University of Queensland, St, Lucia, 4072.

ABSTRACT
The state of metabolic dormancy has fascinated people for hundreds of years, leading to research exploring the identity of natural molecular components that may induce this state. Many animals lower their metabolism during extreme high temperatures, called aestivation. The biological significance for this is clear; by strongly suppressing metabolic rate to low levels, animals gain an extension of the time that they can resist stressful conditions. Understanding the pattern of blood or hemolymph metabolite changes during active and aestivated stages can provide crucial information relating to those molecular components that regulate aestivation in animals, and how they adapt to their different environmental conditions. In this study, we have interrogated the hemolymph metabolite composition of the land snail *Theba pisana*, a remarkably resilient mollusc that displays a yearly natural aestivation period during the summer months. We have identified those metabolites that show up- or down-regulation, including specific phospholipids. Further investigation of RNA-seq data demonstrates the expression of phospholipid-synthesis genes in the snail hepatopancreas, responsible for their synthesis. In summary, we have identified a vast number of aestivation metabolites and some molecular pathways for synthesis.

Key words: snail, *Theba pisana*; aestivation; metabolites; hemolymph; phospholipid