Accepted Manuscript

Title: Characterisation of two conopressin precursor isoforms in the land snail, *Theba pisana*

Author: M.J. Stewart B.I. Harding K.J. Adamson T. Wang K.B. Storey S.F. Cummins

PII: S0196-9781(15)30015-2
DOI: http://dx.doi.org/10.1016/j.peptides.2015.12.009
Reference: PEP 69581

To appear in: *Peptides*

Received date: 28-10-2015
Revised date: 22-12-2015
Accepted date: 28-12-2015


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Characterisation of two conopressin precursor isoforms in the land snail, *Theba pisana*


1 Geneecology 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.

*Corresponding Author: Dr. Scott Cummins, Geneecology Research Centre, Faculty of Science, Health, Education and Engineering, University of the Sunshine Coast, Maroochydore DC, Queensland, Australia, 4558, scummins@usc.edu.au
Highlights:

- A conopressin gene encoding a precursor protein for conopressin identified in the land snail *Theba pisana*.
- *T. pisana* conopressin precursor contains a conopressin peptide and neurophysin.
- Conopressin found in CNS, as well as reproductive tissues, including the mucous glands and dart sac.
- Dart sac conopressin peptide may be involved in ‘love’ dart shooting.

Abstract

Increased understanding of the molecular components involved in mollusc reproduction may assist in understanding the evolutionary adaptations used by animals, including hermaphrodites, to produce offspring. The neuropeptide conopressin, a member of the vasopressin/oxytocin-like peptide family, can modulate various reproductive activities in invertebrates. In this study, we used the hermaphroditic land snail, *Theba pisana*, to investigate the presence and tissue-specific distribution of a conopressin gene. Our transcriptomic analysis of *T. pisana* CNS sheath tissue has revealed two conopressin gene transcripts (*Tpi-conopressin-1* and *Tpi-conopressin-2*), each encoding for precursors containing an identical conopressin nonapeptide and a variable neurophysin. *T. pisana* conopressins share high identity with other land snails and slugs, as well as other mollusc and vertebrate vasopressin/oxytocin, supported by phylogenetic analysis. Conserved residues in the *T. pisana* neurophysin are important for peptide binding, and we present molecular dynamic models demonstrating the most likely stable structure of the *Tpi-conopressin-1* peptide when associated with neurophysin. RT-PCR shows that *Tpi-conopressin-1* is additionally expressed in reproductive tissues, including the dart sac, where abundant spatial expression throughout the sac region is found; this implies a role in ‘love’ dart synthesis or dart injection during mating. The presence of a conopressin receptor in the CNS sheath indicates CNS neural excitation. In summary, this study represents a detailed molecular analysis of conopressin in a land snail.

Key words: conopressin; neurophysin; CNS; dart sac; *Theba pisana*; land snail
1. Background

Mollusc gastropod reproduction is regulated by the neuroendocrine system [21]. In molluscan reproduction, there are some neuropeptides that have attracted particular research interest, including the egg-laying hormone, Ala-Pro-Gly-Try-amide (APGWamide), conopressin, and gonadotropin-releasing hormone.

Conopressin was originally identified from the cone snail venom [14], then subsequently described from several other invertebrates, including the oysters, octopus, leeches, insects and earthworms [28, 30, 31, 36, 39]. Similar to the APGWamide, conopressin has been implicated in the regulation of male reproductive function; although, it acts against APGWamide by stimulating the smooth muscles in the vas deferens to spontaneously contract, leading to eventual ejaculation [42]. Bioactive conopressin is cleaved from a precursor protein that also contains the neuropeptide neurophysin; whereby conopressin is the shorter disulphide bonded peptide (CFIRNCPKGamide) compared to the larger neurophysin, of which the latter is thought to be responsible for the folding and transportation of the lead peptide [21, 43]. Analogues of conopressin are vasopressin and oxytocin, both of which are considered as reproduction-associated hormones in humans and other mammals [43].

In the hermaphroditic land snail, Theba pisana (Müller, 1774), numerous neuropeptides have been reported through RNA-seq followed by in silico analysis of the central nervous system (CNS) [1], although conopressin was not described. The land snail CNS is tetra neural, comprised of a buccal ganglion, visceral ganglia, and a cerebral and pedal ganglion that are encased by a connective sheath [1, 9, 12]. The connective sheath can contain dorsal bodies, small endocrine cells that are proposed to be responsible for regulating reproduction [21]. In general, dorsal bodies can take two different forms, dependent on the species; distinct organs, as in the Great pond snail (Lymnaea stagnalis) [21] and the striped false limpet (Siphonaria pectinanta) [35], or as clusters of small cells throughout the cerebral commissure, as described in the land snails Cornu aspersum [48] and Megalobulimus abbreviatus [27]. Dorsal bodies ablation from immature animals results in underdeveloped female accessory sex organs [21]. Although steroids have been implicated in this underdevelopment, but it is likely that neuropeptides also play a role, as they are well known to form key signalling molecules for sexual development and reproduction [21].

In this study, we explored the existence of conopressin within the CNS sheath of the T. pisana. In many area’s of the world this snail has become an introduced pest, where it can multiply at prolific rates causing widespread damage to agriculture [25, 47]. At present we have a limited understanding of the molecular components involved in its reproduction, or that of any other hermaphrodite land snail, particularly since their anatomy involves complex pathways to regulate control over reproductive metabolism.

We take advantage of a de novo transcriptome assembly of gene transcripts present within the T. pisana CNS sheath. Through bioinformatic analysis we have identified two conopressin genes, and
demonstrate that they share high amino acid sequence homology with conopressin genes in other molluscs. We further show tissue and spatial distribution of conopressin-1 within the CNS, dart sac and mucous glands, supporting a role for this peptide in reproduction. A single conopressin receptor was also identified.

2. Materials and methods

2.1. Animals

*T. pisana* were obtained from Warooka, South Australia (34.9900° S, 137.3990° E). The animals were identified as *T. pisana* by the criteria described in Integrated Snail Management in Crops and Pastures [25]. Once in the laboratory, snails were kept in terrarium-meshed pens at 19°C, 30% humidity and a 12:12h light:dark cycle. They were fed weekly with cucumber and carrot. To determine snail maturity, they were classified into three groups based on visual inspection of their reproductive systems upon dissection; juvenile (no reproductive system), immature (underdeveloped reproductive system including small mucous and albumin glands), and mature (fully developed reproductive system including presence of darts, large mucus and albumin glands).

2.2. Theba pisana CNS sheath transcriptome sequencing (RNA-Seq), bioinformatics and phylogenetics

For histological analysis, cryostat sections were prepared from CNS cerebral ganglia and isolated connective sheath fixed overnight in 4% paraformaldehyde. Sections were cut at 10 µm thickness and then stained using Sudan Red IV (adapted from the methods previously described by Bruce-Gregorios [8]). Micrographs were taken with a Nikon C2 confocal microscope fitted with a Nikon DS-Fi2 camera attachment using NIS Elements AR v4 software package. Digital enhancements for brightness and contrast, as well as figure assembly were constructed through Photoshop CS6 (Adobe, CA, USA).

Approximately 130 sexually mature *T. pisana* CNS were dissected, sheath carefully removed, and total RNA was isolated using TRIzol reagent (Invitrogen, CA, USA) as per supplier’s instructions. Total RNA was quantified with a Nanodrop2000 spectrophotometer (Thermo Scientific, MA, USA) at wavelengths of 260 nm and 280 nm. CNS sheath total RNA was provided to BGI (Hong Kong) for library construction for the Illumina HiSeq platform (4 Gigabases, normalised). De novo assembly was performed using the CLC Genomics Workbench v6.9 (CLC Bio, Aarhus, Denmark) with high-quality reads assembled into unigenes. CNS sheath transcriptome protein coding regions were identified using the ORF-predictor (http://proteomics.ysu.edu/tools/OrfPredictor.html) [26] and combined into a database. SignalP (http://www.cbs.dtu.dk/services/SignalP/) [29] was used to determine presence of a signal peptide, and NeuroPred (http://neuroproteomics.scs.illinois.edu/cgi-bin/neuropred.py) [38] to predict cleavage sites, posttranslational modifications, and bioactive peptide products. Transmembrane domains were predicted using the TMHMM tool (http://www.cbs.dtu.dk/services/TMHMM/). Schematic diagrams of protein domain structures were prepared using the Domain Graph (DOG, version 2.0) software [32].
Protein sequences from *T. pisana* aligned against the predicted protein sequences for conopressin and conopressin receptor were obtained from NCBI, as well as from land snail and slug assemblies available from SRA file XXXXXX (to be included), using the MEGA 5.1 platform using ClustalW and the Gonnet protein weight matrix [41]. Multiple sequence alignment and receptor topology schematics were generated using LaTeX’s TeXShade and TEXtopo packages, respectively [4]. Phylogenetic trees constructed using MEGA5.1 and the neighbour-joining method [34] with 1000 bootstrap replicates.

2.3. Peptide models

The initial conformation of conopressin-1 and neurophysin was built using the LEAP module of AMBER version 11 [10]. Then, molecular dynamics (MD) simulation was unrestrained and carried out in the canonical ensemble using the SANDER module. The ff13 force field [16] was employed. Energy minimisation with 2500 steps was first performed to remove unfavourable contacts. The AMBER structure was then heated to 325K over 50 ps to avoid being kinetically trapped in local minima, and cooled down to 298.15K, then subjected to unrestrained MD simulations at 298.15K for the purpose of peptide equilibration. The structural information was sampled every 1 ps (i.e. 345,000 structures were calculated for 345 ns MD simulation). This MD simulation was continued until the root mean square deviation (RMSD) of structures within a reasonable long time range (from about 60 to 345 ns in this study) is stably less than 2Å. Then a lowest energy structure can be determined, and considered as the representative of the conformations simulated over this period. Visualisation of the systems was effected using via VMD software [20].

2.4. Reverse transcription-polymerase chain reaction (RT-PCR) and gene cloning

Total RNA was isolated from eight different *T. pisana* tissues (bursa copulatrix, bursa tract, dart sac, foot muscle, hepatopancreas, mucous glands, ovotestis, and penis) collected from both mature and immature animals using TRIZol (Invitrogen) extraction methods. Total RNA integrity was analysed using a 1.2% agarose gel with formaldehyde and ethidium bromide staining. Total RNA was used as a template for complementary DNA (cDNA) synthesis using a QuantiTect kit (QIAgen, Limburg, Netherlands) as per the supplier’s instructions. PCR was carried out on template cDNA using REDTaq (Sigma-Aldrich) as per supplier’s instructions and including gene specific primers (50 pmol each; Table 1). Cycling parameters were 94°C for 1 min, 45°C for 2 min, and 72°C for 3 min for 30 cycles. PCR products were separated with a 2% agarose gel (0.6x Tris-Boric acid EDTA, TBE; 0.2% ethidium bromide) prior to visualisation (Syngene, Cambridge, England). Amplicons obtained were purified from 2% agarose gels with a QIAquick spin PCR purification kit (Qiagen) as per supplier’s instructions. Purified PCR products were then ligated into a pGEM-T easy plasmid (Promega, WI, USA) and recombinant clones were sequenced at the Australian Genome Research Facility (AGRF, Brisbane).
2.5. Whole mount in situ hybridisation (WMISH)

Digoxigenin (DIG)-labelled riboprobes were designed against the amplicon regions of the *Tpi-conopressin-1* gene for WMISH, as previously described [15]. Dart sac tissue were collected from mature snails and fixed in 4% paraformaldehyde in PBS overnight, then stored in 70% ethanol at 4°C. WMISH was then performed using DIG-labeled riboprobes with modifications according to Cummins *et al* 2009 [15]. For documentation, specimens were dehydrated by stepwise ethanol changes, cleared in benzyl benzoate: benzyl alcohol (2:1 v/v) and mounted in 70% glycerol. Tissues were then examined with a Leica M205A stereoscope and images were captured with a Leica DFC550 digital camera.

3. Results

3.1. Identification of conopressin in *Theba pisana*

CNS of sexually mature *T. pisana* was investigated to confirm the location of sheath. **Fig. 1A** shows small round globular cells (~20 µm in diameter) stained red with Sudan IV, thought to be dorsal bodies, bound to the connective sheath. These cells are also clearly observed within the connective sheath that had been removed from the cerebral ganglia (**Fig. 1A**, inset). **Fig. 1B** shows an illustrated schematic of the *T. pisana* CNS, including functional regions typical of land snails such as the cerebral ganglia, cerebral commissure, the pedal ganglion and location of the dorsal bodies. This CNS sheath region was successfully removed for RNA isolation and RNA-seq.

Approximately 27.1 million Illumina raw sequence reads were obtained from a CNS sheath transcript library. These raw sequences were trimmed of adapters and deposit into the Genbank SRA database accession number: XXXXXXXXXX (to be included upon acceptance). *De novo* transcriptome sequence assembly revealed a total of 36,360 transcripts with an N50 length of 330 bases. The longest transcript was 6783 bp and the shortest was 129 bp.

BLASTn analysis of this transcriptome identified two transcripts with high similarity to known conopressin precursors (e-value <10^-3). Both transcripts encode for two full-length precursors for *T. pisana* conopressin (*Tpi-conopressin-1, -2*). The 166 and 161 amino acid precursors each contain a 32-residue signal sequence and cleavage sites for the release of a bioactive conopressin nonapeptide (C1FIRNCPKG-NH2) and associated cysteine-rich neurophysin peptide (**Fig. 2A**). A lysine (K) at position 8 of conopressin infers that the correct nomenclature for this peptide is Lys-conopressin. BLASTp identification of conopressin in other land snail and slug species databases (**File S1**) and comparative sequence analysis with *Tpi-conopressins*, shows that the conopressin peptide region is most highly conserved, as well as the spatial conservation of cysteine residues within the neurophysin peptide (**Fig. 2B**). Comparison more broadly with other gastropod classes, a bivalve and cephalopod provides a similar pattern of spatial conservation (**Fig. 2C**). Comparative sequence and phylogeny analysis with other animal species indicates that the *Tpi-conopressin-1* precursor shows highest similarity to conopressin of other molluscs, and further supports a distant relationship with the vertebrate oxytocin and vasopressin (**Fig. 3**).
3.2. Molecular dynamic simulation of conopressin-1 peptide and neurophysin-1 peptide

Molecular dynamic simulations were generated that demonstrate the most likely stable structure of Tpi-conopressin-1 peptide and when associated with neurophysin (Fig. 4). Conopressin-1 has a turn structure that is joined by a single 1.99Å sulphide bridge bound by Cys at the first and fourth residue. Neurophysin-1 is more complex, consisting of three beta bridges, four beta sheets, two alpha and two 3-10 helices, while its 14 cysteine residues are staggered throughout the peptide which enable it to make a very stable quaternary structure. The interaction and binding mechanism of conopressin and neurophysin were further explored. Predicted models show conopressin-1 with its own internal cysteine bridge after being cleaved from its immature form (Fig. 4). The potential energy as a function of time during these simulations and the backbone RMSD relative to these structures over time is also shown. The resolved representative structure of Tpi-conopressin-1 occurred at 229.39 ns into the MD simulation, Tpi-neurophysin-1 occurred at 58.98 ns and the Tpi-conopressin-1:neurophysin-1 complex resolved at 409.61 ns (Fig. 4).

3.3. Analysis of Theba pisana conopressin-1 expression

The tissue-specific expression of Tpi-conopressin-1 was further investigated within non-CNS tissues by RT-PCR. Transcript amplicons of the correct size (502 bp) were clearly identified within the hepatopancreas, foot muscle, penis, ovotestis and the dart sac (Fig. 5A). Sanger sequencing of each tissue amplicon confirmed it was the conopressin-1. Positive controls using actin-specific primers demonstrated amplicons of expected size (212 bp) in all tissues. The dart sac of T. pisana is located in the vicinity of the eversible penis and the vagina. Its role is to synthesise the calcareous ‘love’ dart, a needle-like apparatus used during mating. To investigate the spatial expression of Tpi-conopressin-1 within the dart sac, WMISH was performed on dart sacs taken from mature and immature animals (Fig. 5B, C respectively). Gene expression was prominent in the dart sac of mature snails, throughout the length of the muscular capsule (Fig. 5B), but was more concentrated in the area surrounding the ‘love’ dart in both mature and immature individuals (Fig. 5B, C respectively).

3.4 Identification of a Theba pisana conopressin receptor

A single conopressin receptor was identified from our transcriptome assembly. The receptor has an open reading frame containing 2,871 bp and encodes a protein of 957 amino acids (File S2). As shown in Figure 6A, the general organisation is similar to other GPCR, with seven hydrophobic regions representing putative transmembrane domains. The third intracellular loop is considerably longer than that of other conopressin receptors described so far, containing 606 amino acids based on TMHMM analysis and with no significant similarity to any known sequences (Fig. 6B). Numerous putative phosphorylation sites are predicted, including 37 serine, 9 threonine and 12 tyrosine sites. In comparison with several other known
conopressin receptors, most similarity was observed within the region spanning from transmembrane
domains 2 to 5, and within the C-terminal region (Fig. 6B and File S3), some of which are known to
encompass sites that contribute to ligand-binding [2, 19].

4. Discussion
In this study, we have analysed the T. pisana CNS sheath to assess the presence of a conopressin
neuropeptide gene. Prior to RNA-seq, our analysis confirmed the presence of dorsal bodies in the
connective tissue sheath surrounding the cerebral ganglia and cerebral commissure. These morphological
data are consistent with the similar-sized lipid-rich cells of the dorsal bodies found in other terrestrial
gastropods, including Bulinus truncatus [5]. In the M. abbreviates, dorsal bodies appear to secrete
substances implicated in reproduction, such as stimulation of oocyte maturation and development of the
female accessory sex glands [27], yet the precise identity of the dorsal body substances has not been clearly
defined [5, 17, 18, 21, 22].

Our RNA-seq data using RNA obtained from this CNS sheath, helped to reveal transcripts
encoding two conopressin genes and a conopressin receptor gene. The two Theba conopressin genes
encode for two conopressin precursor proteins, each containing a single conopressin peptide that is 100%
conserved with most other known gastropod conopressins. Similarly, two genes encoding conopressin
precursors have been identified from the cephalopod Octopus vulgaris [40]. O. vulgaris bioactive
conopressin peptides, termed octopressin and cephalotocin, share 5 of 9 residues, and likely have different
functions on peripheral tissues [40]. Other gastropods studied to date have described only one CNS
conopressin gene [44, 46]. Our use of a tissue-specific sequence assembly obtained from next-generation
sequencing, likely ensured for excellent transcript coverage.

In contrast to the highly conserved conopressin peptide, the neurophysin-associated primary peptide
sequence varies significantly between molluscs and other species, besides sharing spatial conservation of
cysteine residues. We expect that the T. pisana conopressin Cys residues are disulfide bonded as has been
demonstrated for other neurophysin-like peptides [13, 33]. The role of neurophysin in mammals has been
established, that is, to act as a carrier protein for the octapeptide hormones vasopressin and oxytocin [7].
The crystal structure of the oxytocin in complex with its neurophysin has been determined, showing a
binding site located at the end of a 3 10-helix [33]. Our molecular model demonstrates a stable complex
between T. pisana conopressin-1 and its neurophysin-1, supporting a likely complex during cell secretion.
In support of this, we find conservation of the Arg and Glu residues in neurophysin (see Fig. 2) that are
known to be essential for binding bovine neurophysins to vasopressin and oxytocin peptides [13, 37].

Besides our study, conopressin expression has been demonstrated within the CNS of other
pulmonary mollusc species, such as L. stagnalis [21]. Furthermore, its activity within CNS neurons of the
land snail Otala lactea has been observed following application of vertebrate neurohypophysial hormones
[3]. Its activity on the reproductive tissues is also established, for instance, in Lymnaea its known to be
transported to the \textit{vas deferens} to affect penis ejaculation \cite{42} while its receptors are present in the genital tract \cite{43}. We found that outside of the CNS, \textit{T. pisana conopressin-1} is present within multiple non-neural tissues, including reproduction-associated tissues such as the penis and dart sac, suggesting a role in reproduction. Its widespread distribution throughout around the dart in both mature and immature \textit{T. pisana} may be necessary for love dart synthesis, ejection of the dart at mating, or may even be a component of a love dart allohormone. Investigations have shown in \textit{C. aspersum} that when mucus-covered darts are successfully stabbed into a mating partner, there is a corresponding increase in paternity success \cite{23}.

We have identified a \textit{T. pisana} conopressin receptor from the transcriptome. Although this receptor has distinct similarities with other conopressin receptors, an unusually large third intracellular loop suggests a more complex interaction with intracellular components. This region in \(\beta\)-adrenergic receptors, as well as the carboxyl terminus confers receptor spontaneous activity \cite{11}. The presence of the Tpi-conopressin receptor within the same region as conopressin, supports a role for CNS neuron regulation, a feature similar to that of oxytocinergic neurons of the mammalian hypothalamus \cite{24}. In \textit{Lymnaea}, two conopressin receptors have been identified that localise to the large portion of the anterior lobe neurons \cite{45}, a region that controls the penis nerve \cite{6}. We expect that further study into the spatial expression of the \textit{T. pisana} conopressin receptor will allow for a better understanding of the sites of conopressin activity.

In summary, we have analysed a \textit{T. pisana} CNS sheath transcriptome and report gene transcripts encoding for conopressin neuropeptide and a conopressin receptor. Expression of a conopressin transcript additionally within the reproductive organs, including the dart sac, suggests a role for it in reproductive processes. This information extends upon our understanding of neuropeptides in hermaphroditic land snails and could help to develop ways for managing invasive land snail populations.

\textbf{Authors’ contributions}

MJS carried out the gene analysis, constructed the images and drafted the manuscript. BIH and KJA carried out the RNA isolation, RT-PCR analysis, WMISH, and drafted the manuscript. KBS, TW and SFC designed the study and edited and drafted the manuscript. All authors read and approved the final manuscript.

\textbf{Acknowledgements}

This work is supported by the Grains Research and Development Corporation (GRDC) Innovation and Investment Grants (USU00001; MJS, TW, SFC), GRDC Scholarship (BIH), and the Australian Research Council (SFC). We wish to extend also our gratitude to Mr. Graham Hayes, we thank for supplying the snails used in this study.
References


[40] Takuwa-Kuroda K, Iwakoshi-Ukena E, Kanda A, Minakata H. Octopus, which owns the most advanced brain in invertebrates, has two members of vasopressin/oxytocin superfamily as in vertebrates. Regul Pept 2003;115:139-49.


Figure Legends

Fig. 1
Identification of cerebral ganglia connective sheath of *Theba pisana*. (A) Cryostat section (10 µm thickness) of the cerebral ganglia (CG) stained with Sudan IV. Red indicates location of dorsal bodies (DB) within the connective sheath (CS). Inset, boxed region higher magnification showing image of stained DBs. (B) Schematic of the CNS of *T. pisana*, showing locations of CS and DBs in red. CC, cerebral commissure; PaG, parietal ganglia; PeG/Ped, pedal ganglia; PlG, pleural ganglia; PC, procerebrum; TNv, tentacle nerve; MtC, metacerebrum; PCN, procerebrum, VG, visceral ganglia.

Fig. 2
Molecular characterisation of *Theba pisana* conopressins. (A) Amino acid sequence and schematic showing organisation of the *T. pisana* conopressin-1 and -2. Cys residues are highlighted in black in the sequence and as “S” in the schematic. (B) Multiple sequence alignment between *T. pisana* conopressins and other land gastropod conopressin precursors, and (C) between conopressin precursors of other molluscan groups. Black line shows regions of conopressin peptide. Vertical arrows indicate amino acid residues required for binding of neurophysins to peptides. Species names: *Cochlicella acuta, Cornu aspersum, Deroceras reticulata, Milax gagates, Lottia gigantea, Lymnaea stagnalis, Aplysia californica, Aplysia kurodai, Pinctata fucata, Octopus vulgaris, Euprymna scolopes*.

Fig. 3
Phylogenetic analysis indicates *Theba pisana* conopressin groups within other molluscan conopressin, compared to other phyla, including vasopressin, oxytocin, and mesotocin. Genbank accession numbers are shown in parenthesis.

Fig. 4
Predicted structural models for *Theba pisana* conopressin-1, neurophysin-1 and conopressin-1-neurophysin-1 complex. Shown below each model are the potential energies of conopressin-1, neurophysin-1 and of the neurophysin-1:conopressin-1 complex as a function of time during MD. Also, the backbone RMSD during the same MD, compared to the lowest-energy conformation (the representative structure).

Fig. 5
Tissue-specific and spatial expression of conopressin-1 gene in *Theba pisana*. (A) Amplicons for conopressin-1 gene from *T. pisana* tissues derived from mature animals. Tpi-actin was used as a positive control. (B, C) Whole-mount *in situ* hybridisation localisation of *T. pisana conopressin-1* within the
mature *T. pisana* dart sac (B), and immature *T. pisana* dart sac (C). Blue colour represents gene expression in dart sac.

**Fig. 6**

Molecular characterisation of a *Theba pisana* conopressin receptor. (A) Schematic shows organisation of the *T. pisana* conopressin receptor, including location of seven transmembrane domains 1-7. (B) Schematic shows representation of conopressin receptor in membrane. Asterisks indicate amino-acid regions that are highly conserved in the AVP/OT receptor superfamily that have been suggested to participate and to be important in receptor–ligand interaction. Transmembrane regions are designated as I-VII.
Figure 1

200μm

Figure 2
Figure 5

A

<table>
<thead>
<tr>
<th>bp</th>
<th>Hepatopancreas</th>
<th>Foot muscle</th>
<th>Penis</th>
<th>Ovotestis</th>
<th>Bursa copulatrix</th>
<th>Bursa tract</th>
<th>Dart sac</th>
<th>Mucous gland</th>
</tr>
</thead>
<tbody>
<tr>
<td>502</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>conopressin-1</td>
</tr>
<tr>
<td>212</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>actin</td>
</tr>
</tbody>
</table>

B

distal

proximal

500 µm

C

love dart

distal

proximal

500 µm

Figure 6

A

1 2 3 4 5 6 7

B

H₂N-6⁰

conserved positions

invariable positions

extra

I II III IV V

VI VII

intra

COOH
Supplementary figures

Fig. S1
Amino acid sequences for land gastropod conopressin precursors. Species names: *Theba pisana*, *Cochlicella acuta*, *Cornu aspersum*, *Deroceras reticulata*, *Milax gagates*.

Fig. S2
Nucleotide and amino acid sequence for *Theba pisana* conopressin receptor.

Fig. S3
Multiple sequence alignment between *Theba pisana* conopressin receptor and other invertebrate conopressin receptors.
Table 1. Primers used for RT-PCR.

<table>
<thead>
<tr>
<th>Gene Primer</th>
<th>Sequence (5'‐3')</th>
<th>Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ActinF1</td>
<td>CTCTCTGCTGTGGTGTTGAA</td>
<td></td>
</tr>
<tr>
<td>ActinR1</td>
<td>ATCCAGGCTGTCTCTTTTTCTTT</td>
<td>212</td>
</tr>
<tr>
<td>ConopressinF1</td>
<td>GATGCCATCATATCTTCCAG</td>
<td></td>
</tr>
<tr>
<td>ConopressinR1</td>
<td>TTAATCTCTGATTCTGTTGTTG</td>
<td>502</td>
</tr>
</tbody>
</table>