Myomodulin Gene of *Lymnaea*: Structure, Expression, and Analysis of Neuropeptides

Elaine Kellett, Stephen J. Perry, Niovi Santama, Belinda M. Worster, Paul R. Benjamin, and Julian F. Burke

Sussex Centre for Neuroscience, School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG, United Kingdom

The myomodulin family of neuropeptides is an important group of neural cotransmitters in molluscs and is known to be present in the neural network that controls feeding behavior in the snail *Lymnaea*. Here we show that a single gene encodes five structurally similar forms of myomodulin: GLQMLRLamide, QIPMLRLamide, SMSMLRLamide, SLSMLRLamide, and PMSMLRLamide, the latter being present in nine copies. Analysis of the organization of the gene indicates that it is transcribed as a single spliced transcript from an upstream promoter region that contains multiple cAMP-responsive elements, as well as puta-

tive elements with homology to tissue-specific promoter-binding sites. The presence in nervous tissue of two of the peptides, GLQMLRLamide and PMSMLRLamide, is confirmed by mass spectrometry. *In situ* hybridization analysis indicates that the gene is expressed in specific cells in all ganglia of the CNS of *Lymnaea*, which will allow physiological analysis of the function of myomodulins at the level of single identified neurons.

Key words: myomodulin; neuropeptide; gene; *Lymnaea*; molluscs; neuronal modulation

Studies aimed at understanding the signaling function of neuropeptides have emphasized the value of using model invertebrate systems, such as gastropod molluscs, in which the pattern of gene expression and the physiological function of processed peptides can be studied at the level of single identified neurons (Cropper et al., 1987, 1991; Benjamin and Burke, 1994). A common feature of invertebrate neuropeptide genes is that they encode a variety of structurally related peptides present on one or more precursor proteins. Families of related peptides occur in the same type of animal but, in addition, related families can be recognized in other species as well, resulting in enormous interphyletic diversity of peptide structures. An important example of this structural diversity is the myomodulin family of molluscan neuropeptides. Myomodulin A (PMSMLRLamide) and eight related peptides were first discovered by molecular and biochemical methods in *Aplysia* (Cropper et al., 1987, 1991; Lopez et al., 1993; Miller et al., 1993; Brezina et al., 1995). Myomodulin A is now known to be present in several other molluscs together with at least one more related molecule (Fugisawa et al., 1990), and further types of myomodulin are also likely to be discovered with further investigation.

In the pulmonate snail *Lymnaea*, the presence of myomodulin A was confirmed by HPLC purification and sequencing of CNS extracts (Santama et al., 1994a) and by mass spectrometry of the penis nerve (Li et al., 1994). Application of myomodulin A to the penis enhanced the amplitude of electrically induced contractions

and increased the relaxation rate of the penis muscle (van Golen et al., 1996). This indicated that myomodulin plays a role in modulating neuromuscular transmission. This was also the case in *Aplysia*, in which an important physiological study showed that the myomodulin peptides acted in concert to enhance acetylcholine-induced contractions in a specific muscle of the buccal mass, the feeding apparatus of molluscs (Brezina et al., 1994a,b). The presence of myomodulin-like peptides in the feeding system of *Lymnaea* was also indicated by detailed immunocytochemical studies (Santama et al., 1994b). As well as feeding motoneurons, immunoreactivity was also found in identified modulatory interneurons of the feeding network, suggesting a central function for myomodulin as well as the peripheral neuromuscular function in feeding, already shown for *Aplysia*.

The likely presence of the myomodulins in several behaviorally important networks in *Lymnaea* made it important to carry out a molecular analysis of the *Lymnaea* gene. The aim was to discover the full range of myomodulin-like peptides in *Lymnaea* and to provide a more specific molecular probe for *in situ* hybridization analysis of the CNS. We describe the structure of cDNAs encoding myomodulin A and four structurally related peptides, the organization of the gene and its promoter, and its expression in the CNS (*in situ* analysis). Preliminary mass spectrometric evidence confirmed the presence of diverse myomodulins in nervous tissue.

MATERIALS AND METHODS

Specimens of wild *Lymnaea stagnalis*, collected from freshwater ponds, were supplied by Blades Biological (Edenbridge, Kent, UK). A supply of captive-bred snails was also supplied by the Department of Biology at the Vrije Universiteit (Amsterdam, The Netherlands). Both types of snail were maintained at 20°C in 40 l aquaria in drip-fed tap water, kept under a 12 hr light/dark cycle and fed on washed lettuce *ad libitum*. All molecular biology procedures followed the protocols in Sambrook et al. (1989).

Isolation of cDNA clones. A random-primed [α -³²P]dCTP-radiolabeled cDNA probe encoding a portion of the *Aplysia* myomodulin propeptide (Miller et al., 1993) (kindly donated by K. R. Weiss, Mt. Sinai School of

Received March 22, 1996; revised May 22, 1996; accepted May 24, 1996.

This work was supported by a grant from the Biotechnology and Biological Sciences Research Council. B.M.W. was partly funded by a CASE award from Micromass UK, Ltd. (Wythenshawe, UK). We thank K. Weiss and M. Miller for generously supplying the *Aplysia* myomodulin cDNA clone used in this work.

Correspondence and reprint requests should be addressed to Stephen J. Perry, Sussex Centre for Neuroscience, School of Biological Sciences, University of Sussex, Falmer, Brighton, East Sussex BN1 9QG, UK.

Dr. Kellett's present address: Davidson Building, Division of Biochemistry and Molecular Biology, IBL, University of Glasgow, Glasgow, UK.

Dr. Santama's present address: European Molecular Biology Laboratory, Heidelberg, Germany.

Copyright © 1996 Society for Neuroscience 0270-6474/96/164949-09\$05.00/0

Medicine, New York, NY) was used to screen a *Lymnaea* CNS cDNA library constructed in the vector λ ZAP II (previously described in Vreugdenhil et al., 1988). Hybridization was performed on duplicate filters of 2×10^4 plaques at 65°C overnight in a hybridization solution containing 3 $\times$ SSC, 0.1% (w/v) SDS, and 5 $\times$ Denhardt's solution [0.2% (w/v) polyvinylpyrrolidone, 0.2% (w/v) bovine serum albumin, and 0.2% (w/v) Ficoll]. After four washes at 65°C in 0.2 $\times$ SSC and 0.1% SDS, the filters were left on preflashed X-ray film overnight at –70°C. Duplicate positive plaques were purified through secondary and tertiary screens, and the cDNA clones were isolated by *in vivo* excision of the pBluescript SK[–] plasmid (Stratagene, La Jolla, CA). In this way, six cDNAs that showed homology to the open reading frame of the *Aplysia* sequence were isolated.

Isolation of genomic clones. EcoRI-digested genomic DNA was used to construct a library in the vector λ gt10. Duplicate filters of 5×10^5 plaques were screened using randomly primed probes of both EcoRI fragments of the *Lymnaea* cDNA clone pMM421. Hybridization was performed in the same manner as described for isolation of the cDNA clones, but washes were performed at 65°C with 0.1 $\times$ SSC and 0.1% SDS. Overnight exposures of the filters revealed four positively hybridizing plaques. After tertiary screens, purified plaques were isolated and the inserts were isolated by restriction digestion with EcoRI and subcloned into pUC19.

Sequencing of clones. Restriction digestion of isolated cDNA and genomic clones and subsequent subcloning of the fragments into pUC19 were performed. All subclones were sequenced by the Sanger method of chain termination from double-stranded DNA using fluorescent dye-primer cycle sequencing kits (ABD, Warrington, UK) and analyzed on a 373A automated DNA sequencer (ABD).

5'-Exon extension using PCR. The only region of the cDNA clones that was not present in the previously isolated fragments of the gene (see Results) was confirmed as being the 5' end of the second exon by PCR. A primer designed for the very 5' end of the absent region (MidS1: AATACAGAAGAATCCGGTGGCCAG) was used in conjunction with three primers designed for regions of the reverse complement strand both in the large exon and in the 3' region of the absent fragment (MidA1: AGCAAGCTCAAAATGTCGTCTATGT; MidA2: GAATTCGGGT-TCTGCTCTT-GGTA; GMM36A2: ACCTCCTGAACCTGTGGTGT-TCAG). Reactions were carried out in 50 μ l volumes using 20 ng of genomic DNA as the template, 50 pmol of each primer, 200 μ M dNTPs (Pharmacia, Uppsala, Sweden), 5 μ l of 10 $\times$ PCR buffer (500 mM KCl, 20 mM Tris, pH 8.4, and 20 mM MgCl₂). Cycling conditions were as follows: three cycles of 94°C for 2 min, 65°C for 2 min, followed by 27 cycles of 94°C for 30 sec, 65°C for 30 sec, and 72°C for 1 min. One unit of AmpliTaq DNA polymerase (ABD) was added at the annealing temperature of the first cycle as in a hotstart protocol. The sizes of the PCR products were determined by electrophoretic separation on a 1.2% agarose gel and were cloned into the pCR II cloning vector (R&D Systems), and the sequences of both strands were determined as described previously for the library clones.

Transcriptional start-site mapping by primer extension. Primer MMRTA1, complementary to the coding strand of pGMM28A between 1254 and 1276 bp (ATGTTCTTGACGTATGTTGGCGT), was 5'-radiolabeled with [γ -³²P]ATP for use in primer extension after hybridization to CNS total RNA. Two hundred nanograms of primer were incubated at 37°C for 1 hr in a 20 μ l reaction mix with 2 μ l of 10 $\times$ kinase buffer [700 mM Tris, pH 8.0, 100 mM MgCl₂, and 5 mM dithiothreitol (DTT)], 30 μ Ci of [γ -³²P]ATP, and 10 U of T4 polynucleotide kinase. After precipitation, 0.5 fmol of primer was incubated with 20 μ g of CNS total RNA in a volume of 11 μ l at 70°C for 10 min, followed by rapid cooling to 40°C for 10 min to allow primer annealing. The mixture was placed on ice, and 4 μ l of 5 $\times$ first-strand synthesis buffer (250 mM Tris-HCl, pH 8.3, 375 mM KCl, and 15 mM MgCl₂), 2 μ l of 0.1 M DTT, 1 μ l of 10 μ M dNTP stock solution, and 200 U of MMLV reverse transcriptase were added (Superscript II, Life Technologies). The reaction was incubated for 1 hr at 45°C and stopped by the addition of 4 μ l of 0.5 M EDTA. The reverse transcription products were size-fractionated on a 6.5% polyacrylamide gel, using a cytosine sequencing reaction of clone pGMM28A from the same primer site as a molecular weight marker (standard sequencing reaction using the Sequenase Kit, Amersham Life Sciences). The gel was fixed and dried before autoradiography for 24 hr.

Southern blot analysis. Genomic DNA Southern blots were performed to elucidate the size of the intervening sequence in the coding region of the gene. In each restriction digest, 10 μ g of genomic DNA extracted from the buccal mass, brain, and reproductive organs of 10 snails was digested overnight at 37°C and size-fractionated on a 0.6% agarose gel.

After depurination, denaturation, and neutralization, the DNA was transferred onto a nylon filter (Amersham) and permanently cross-linked by exposure to ultraviolet light. Hybridization was performed under the same conditions as described for the screening of the genomic library, using the randomly primed PCR product pMID2 (Fig. 3b) as a hybridization probe. The filter was autoradiographed overnight at –80°C with preflashed film and an intensifying screen.

Northern blot hybridization. Total RNA was isolated from the CNS of 50 *Lymnaea*, and 20 μ g was size-fractionated on a 1% agarose gel in 10 mM sodium phosphate running buffer (method described by Pellé and Murphy, 1993). The RNA was transferred to a nylon membrane and permanently cross-linked to it by exposure to ultraviolet light. Hybridization with the randomly primed [α -³²P]dCTP-labeled cDNA clone pMM421 was performed overnight at 68°C with the same buffer used for library screening. Washes were performed at 68°C with 0.1 $\times$ SSC and 0.1% SDS before overnight autoradiography.

In situ hybridization. The entire CNS was isolated, and nerves and connective tissue sheaths were removed before snap freezing in liquid freon and immersion in liquid nitrogen. The tissue was freeze-dried for 24 hr, fixed in paraformaldehyde vapor (60°C, 90 min), and embedded in wax. Sections were cut (7 μ m) and arranged on gelatin/chrome alum-coated slides and dried overnight at 37°C. Before hybridization, the sections were dewaxed (xylene, 20 min), washed in methanol (2 min), permeabilized with 0.1% pepsin (10 min), fixed in 2% paraformaldehyde in PBS (4 min), and blocked by 1% hydroxylammonium chloride (15 min). Hybridization, using an oligonucleotide complementary to a coding region of the gene (P25: GAAGCTCGTCCACGTCTCCGTA), 3'-tailed by terminal transferase with digoxigenin-11-UTP, was carried out overnight in hybridization buffer (25% deionized formamide, 3 $\times$ SSC, 500 μ g/ml sheared salmon sperm DNA, 500 μ g/ml yeast tRNA, and 1 $\times$ Denhardt's solution) at 55°C. Washes were carried out the next day (3 $\times$ SSC, 20 min, 25°C; 3 $\times$ SSC, 20 min, 55°C; 3 $\times$ SSC, 10 min, 25°C) before nonspecific binding site blocking [1% blocking reagent (Boehringer Mannheim) in 100 mM maleic acid, pH 7.5, 150 mM NaCl, and 100 mM Tris, pH 7.5, 10 min] and incubation with anti-digoxigenin/alkaline phosphatase-conjugated Fab fragment (Boehringer Mannheim; 1 hr in blocking buffer). After washing to remove unbound antibody (twice for 5 min each in 100 mM Tris, pH 7.5, and 150 mM NaCl), the slides were covered with the substrate mixture (10 mM Tris, pH 9.0, 10 mM MgCl₂, 165 μ g/ml BCIP, and 330 μ g/ml NBT) and the color reaction was allowed to develop in a light-proof box until a sufficient level of staining was reached. Slides were washed (100 mM Tris, pH 7.5, and 150 mM NaCl, 5 min) and overlaid with Immumount (Shandon) and coverslips.

Mass spectrometric analysis of nervous tissue. Small lengths of the external right parietal nerve were dissected out and placed on the mass spectrometer target in 5 μ l of 0.5% dihydroxybenzoic acid (DHB). The tissue was crushed and ripped with fine forceps to release peptides into the DHB matrix solution. The sample was allowed to dry and was analyzed on a Micromass Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometer (MALDI-TofSpec, Micromass Organic, Wythenshawe, UK). Short laser bursts (3 nsec, 337 nm wavelength) were used to ionize the peptides, which were focused using a reflectron system to enhance the resolution of the mass peaks.

RESULTS

Characterization of three classes of cDNA clones encoding a myomodulin prepropeptide

To isolate a cDNA copy of the myomodulin gene transcript, a *Lymnaea* total CNS cDNA library was screened with the *Aplysia* myomodulin cDNA clone. Six positively hybridizing clones were produced, and each was digested with the restriction enzymes *Eco*RI or *Bam*HI to release the inserts from the vector DNA. *Eco*RI produced two fragments for each clone, and *Bam*HI produced a single fragment. The *Bam*HI restriction fragments varied in size for each clone, suggesting that the clones were not identical (data not shown). Sequencing of the clones, however, showed that all six originated from the transcription of a single gene. The observed variation in the lengths of the clones was shown to be attributable to differential use of three polyadenylation signals (indicated in Fig. 1a) and varying degrees of 5' end truncation. The sequence of the longest cDNA clone, pMM421, is shown in

Figure 1b, with 32 bp at the 5' end derived from three shorter cDNAs, which exhibited less 5' end truncation (EMBL database accession number X96933).

Computer-based analysis of the 5' end of the cDNA sequence revealed a methionine start codon (AUG) at position 526 followed by a short (7-amino-acid) open reading frame and a stop codon (UGA). The 526 bp of sequence upstream of this is punctuated with stop codons, and has a high (A+T) base content, indicative of noncoding DNA in *Lymnaea*. A second start codon at position 588 begins a 350 codon open reading frame, ending at position 1640 with a UGA stop codon. This is the major open reading frame in all six cDNA clones, and translation of its sequence is shown in Figure 1*b*. The sequence surrounding this start codon (AGACCAAUGC) compares well with the Kozak consensus, whereas the equivalent sequence around the first start codon at position 526 does not match the consensus (Kozak, 1987). It seems most likely, therefore, that translation initiation occurs at the second start codon.

The stop codon at position 1640 is followed by a variable length of untranslated DNA, which contains two perfect polyadenylation signals (AAUAAA; Proudfoot, 1991) at positions 1972 and 2426, whereas clone pMM421 (the longest, a class III clone in Fig. 1a) ends 18 bp after a possible imperfect polyadenylation signal (AAUAGA) at position 2747, but lacks a polyA tail. The shortest clones (class I in Fig. 1a: pMM52, pMM53, and pMM511) all end 16–19 bp after the first polyadenylation signal; clones pMM52 and pMM53 both end with polyA tails (60 and 11 nucleotides, respectively), whereas pMM511 lacks a polyA tail. The class II clones, pMM46 and pMM54, utilize the second polyA signal at position 2426 and end 14–16 bp 3' to this; both have 10 nucleotide polyA tails. Both the 5'- and the 3'-untranslated regions (587 and 356–1130 bp, respectively) are unusually long compared with other molluscan neuropeptide genes, including the equivalent transcripts isolated from *Aphysia* (Lopez et al., 1993; Miller et al., 1993). Functional significance for the alternate use of different polyadenylation sites and the extraordinary length of the untranslated regions has yet to be determined.

Structure of the deduced prepropeptide encoded by the cDNAs

Translation of the open reading frame between 588 and 1640 bp in Figure 1*b* would produce a 350 residue polypeptide with a predicted molecular mass of 40.4 kDa. The full complement of peptides predicted from the deduced amino acid sequence of the myomodulin cDNA is given in Figure 2. Analysis of the N-terminal 30 amino acids of the polypeptide shows properties of a signal peptide with a highly hydrophobic core (residues 4–17), after which the polypeptide becomes hydrophilic. A cleavage site

Figure 1. Three classes of myomodulin cDNA from *Lymnaea stagnalis*. *a*, Schematic diagram of a class III myomodulin cDNA, indicating the position of the untranslated regions (*white bars*) and the open reading frame (*hatched bar*). The position of the *EcoRI* restriction site (*E*) at 826 bp is shown. The three alternative polyadenylation signals used in the three classes of cDNA are marked with *arrows* and the base position in the complete nucleotide sequence (*b*). *b*, The complete nucleotide sequence of a class III myomodulin cDNA. The deduced sequence of the precursor protein encoded by the open reading frame (588–1637 bp) is also given; the amino acid residue positions are *underlined* on the right. The stop codon is shown as an *asterisk*. The three polyadenylation signals are *bold* and *underlined*, and the polyadenylation sites are marked by *vertical lines*. The *arrowhead* at 645 bp marks the position at which the intron lies within the gene. Potential peptides within the opening reading frame are *underlined*.

Figure 2. The organization of the myomodulin prepropeptide. The putative peptide coding regions are shown in white and are numbered. The hydrophobic leader sequence is labeled S, and the spacer regions are stippled bars. The predicted leader sequence cleavage site at residue Gly²⁰ is marked with an arrow. All putative endopeptidase cleavage sites are shown as black lines, and the sequence of the sites is given above each line. The amino acid sequence of the peptides is given with the number of copies of each encoded within the precursor given in parentheses.

after Gly²⁰ is predicted by the $-1, -3$ rule (von Heijne, 1986) and is indicated by a drop in the hydrophobicity of the polypeptide beyond this point (Kyte and Doolittle, 1982; Hopp, 1986). Cleavage at this point would produce a 330-amino-acid propeptide that encodes five different myomodulin-like peptide sequences. These are flanked at the N terminus by lysine–arginine basic amino acid pairs (sites of endoproteolytic cleavage), with the exception of the first predicted peptide sequence, GLQMLRLG, which is preceded only by a single arginine residue; endoproteolytic cleavage at such sites is also well documented (Newcomb et al., 1987; Linacre et al., 1990). All of the putative peptides are flanked at the C terminus by glycine residues followed by a pair of basic amino acids. Glycine residues at the predicted C termini of peptides are enzymatically processed to amide groups in the mature peptides (Bradbury et al., 1982). Each peptide is expected to be amidated, and one peptide, QIPMLRLamide, may also be cyclized at the N terminus to give a pyroglutaminyl group (pQIPMLRLamide). The copy number of the different classes of peptides varies, with nine tandemly arranged copies of myomodulin A (PMSMLRLamide; peptide 5 in Fig. 2) situated near the C terminus, two copies of SLSMLRLamide, and one copy each of the other three peptides located nearer the N terminus. The C-terminal four amino acids (MLRL) are conserved in all five peptide structures, as is the glycine amidation signal. All of the variation in the peptide structures occurs in the N-terminal three residues. Myomodulin A (PMSMLRLamide) and SMSMLRLamide only vary by a single amino acid at position 1 (Pro→Ser); SLSMLRLamide is structurally identical to these but has an additional conservative substitution at position 2 (Met→Leu). Position 2 is occupied by an apolar residue in all five peptides (Leu, Met, and Ile), and most variation occurs at positions 1 and 3, where charged, polar, and apolar residues are found.

Of the 350 amino acids within the prepropeptide, 183 (52%) are not within the myomodulin-like peptides, their presumed processing sites or the signal peptide, but are found in five highly negatively charged spacer regions (33% of the residues are Asp or Glu) between the peptide-encoding regions. The most C-terminal of the spacers (residues 339–350) contains a dibasic endopeptidase cleavage site which, if cleaved, would release two pentamers: one highly negatively charged, EDDEE, and the other with the sequence

SLAMS. Neither of these peptides shows homology to any known neuropeptide. The two largest spacers (residues 21–53 and 65–178) are bisected by furin-like endopeptidase cleavage sites [RX(K/R)R-residues 41–44 and 114–117] that may act as sites of primary endoproteolytic cleavage within the *trans*-Golgi (Hosaka et al., 1991).

Genomic organization of the myomodulin gene

Four hundred thousand plaques of a genomic library constructed from *Eco*RI-digested *Lymnaea* genomic DNA were screened using *Eco*RI fragments of the class III cDNA clone pMM421. Two positively hybridizing plaques were isolated, and restriction digestion with *Eco*RI revealed that they had insert sizes of ~3.6 and ~2.8 kb (pGMM36A and pGMM28A, respectively). Alignment of the sequence of clone pGMM36A with that of the cDNAs showed that it began at the *Eco*RI site at 861 bp of the cDNA clones (see Fig. 1b) and represented the entire uninterrupted sequence of cDNA clone pMM421 3' to this restriction site (1909 bp) and another 1435 bp of novel sequence past its end. Within this novel region are located a putative downstream RNA 3' cleavage signal (CATGTTTC; Birnstiel et al., 1985) and A/T- and G/T-rich tracts that may act as transcriptional termination signals (Birnstiel et al., 1985; Guntaka, 1993).

Clone pGMM28A (EMBL database accession number X96934) contained uninterrupted sequence identical to bases 1–645 of the cDNA sequence in Figure 1b, after which point the sequence diverged at the sequence TAGGTGAGT, which follows the consensus for the mammalian 5' donor splice site almost exactly (MAG|GTRAGT; Jacob and Gallinar, 1989). Analysis of the sequence 3' to this putative splice site did not reveal a 3' acceptor splice site, a branch sequence, or an open reading frame, indicating that all of the sequence beyond the 5' splice site represents part of an intron within the coding region of the gene. The 5' splice junction falls within the codon for Gly²⁰ of the major open reading frame, the last codon of the signal peptide. Therefore, the first exon of the myomodulin gene contains the 5'-untranslated region and the first 20 codons of the precursor protein. The significance of the positioning of this intron with respect to possible alternate splice forms of the gene is discussed later.

The two genomic DNA fragments isolated included 2556 bp of the 2770 bp of cDNA sequence given in Figure 1b. The 214 bp of sequence not included in these clones were shown to lie at the 5' end of the sequence included in genomic clone pGMM36A by PCR amplification of this region from genomic DNA. Primer pairs designed against the cDNA sequence across this region (as described in Materials and Methods and represented by arrows in Fig. 3a) produced three PCR products that were isolated, cloned, and sequenced. In this way, the entire coding region and the 5'- and 3'-untranslated regions were shown to lie in two exons separated by an intron of unknown length. The length of this intron was revealed by hybridization of clone pMD2 (the PCR product generated using primers midS1 and midA1) to genomic DNA digested with the restriction enzymes *Eco*RI, *Pst*I, and *Sal*I, and in combinations of pairs of these enzymes followed by size fractionation by gel electrophoresis (Fig. 3b). A single hybridizing band was detected in all lanes except the *Sal*I-digested DNA, in which two bands of 25 and >30 kb were observed. All digests with *Eco*RI, whether singularly or in combination with other enzymes, produced a band of 2.7 kb, whereas the *Pst*I and *Pst*I/*Sal*I digest produced a single band of 20 kb. These data, in combination with the restriction maps of the two genomic clones previously isolated, predict an intron length of ~19 kb, in which there are no *Pst*I or

Figure 3. The genomic organization of the myomodulin gene. *a*, PCR amplification of the 5' end of the second exon using primers specific for the cDNA sequence (arrows labeled *GMM36A2*, *midS1*, *midA1*, and *midA2*) produced three products (*pMID1*, *pMID2*, and *pMID3*). Sequencing of these confirmed that the region of the cDNAs absent from genomic clones *pGMM28A* and *pGMM36A* is uninterrupted and lies at the 5' end of the second exon. *b*, Southern blot analysis of restriction-digested *Lymnaea* genomic DNA using the PCR product *pMID2* as a hybridization probe. Restriction digests with *EcoRI* (lanes 2, 3, 7) produced a single band of 2.7 kb, whereas restriction digests with *PstI* (lanes 4, 5) produced a 20 kb band (indicated with arrows). Because of the high-stringency washing of the filter, the second *PstI* restriction fragment was not detected by the probe because of its short length of overlapping sequence (30 bp). 1, Undigested DNA; 2, *EcoRI*; 3, *EcoRI/PstI*; 4, *PstI*; 5, *PstI/SalI*; 6, *SalI*; 7, *EcoRI/SalI*. This photograph has been electronically enhanced for improved clarity. *c*, Deduced map of the myomodulin gene with a single large intron of ~19 kb (black line). All restriction fragment sizes are given in kb. The black region indicates the position of the PCR product *pMID2* used as the hybridization in Figure 5*B*; the positions of genomic clones *pGMM28A* and *pGMM36A* are indicated by diagonally striped boxes.

SalI sites, but two *EcoRI* sites 836 bp 3' to the 5' donor splice site, and ~2.5 kb 5' to the 3' acceptor splice site (Fig. 3*c*).

Mapping of the transcriptional start site

To confirm that the cDNA clones isolated represented almost complete transcript copies, the transcriptional start site of the myomodulin gene was identified by primer extension mapping. Twenty micrograms of CNS total RNA were reverse-transcribed from the primer

MMRTA1 and size-fractionated by PAGE. The C reaction of a sequencing reaction of clone *pGMM28A* from the same primer site was used as a size marker (Fig. 4*a*). Two high-intensity reverse-transcribed products and a weaker band were observed, corresponding closely to the region in which some of the cDNA clones have their start points. The positions of transcription initiation within the CNS, therefore, are very close to the 5' ends of some of the cDNAs. Therefore, all of the sequence of the genomic clone *pGMM28A* 5' to

Figure 4. *a*, Transcriptional start site mapping of the myomodulin gene. Three products of reverse transcription of total CNS RNA from the radiolabeled primer MMRTA1 were separated by PAGE (lane *i*) using the ddCTP reaction of a manual sequencing reaction as a size marker (lane *C*). The sequence of genomic clone pGMM28A to which the marker corresponds is shown on the right (cytosines are underlined). The nearest adenosines to which the reverse transcription products could correspond are marked with arrows (the faint largest product is marked by a gray arrow). The first base of the least truncated cDNA clone isolated is marked with an asterisk. *b*, Northern blot of total CNS RNA hybridized with a radiolabeled class III myomodulin cDNA. Six bands of 3.8, 3.2, 2.8, 2.6, 2.1, and 1.5 kb (indicated with arrows) were detected. Three of these correspond to sizes predicted for the three classes of myomodulin cDNA isolated (2.1, 2.6, and 2.8 kb).

these sites corresponds to the promoter and upstream control regions of the myomodulin gene.

Structure of the myomodulin gene promoter

Genomic clone pGMM28A contains up to 1104 bp of DNA positioned 5' to the point of transcriptional initiation. This is likely to form part or all of the promoter region of the gene. Possible control elements within this region were identified by comparison of its sequence with the Eukaryote Promoter Database (EPD), available as part of the EMBL DNA sequence database, manual searches for the sequences of known control elements (Locker and Buzzard, 1990), and direct sequence comparisons to three published *Aphysia* neuropeptide gene promoter sequences (DesGroseillers et al., 1987). These searches revealed a number of putative sequence elements that may be involved in the control of transcription initiation (Fig. 5), including an imperfect TATA box sequence (GATAAA) at –28, and cAMP-responsive elements (CRE and AP-2) at –133, –445, –839, and –1021. Two AP-5 elements, also present in the *Aphysia* L11 neuropeptide gene promoter (at –54 and –285), are present as well as a number of

sequences with homology to tissue-specific elements (liver-specific promoter element LF-A1 at –442 and –687; immunoglobulin promoter elements μ E1 at –204; and μ E4 at –297). Thus, the promoter region appears to contain a number of different types of control elements that may allow the rate of transcription initiation to be controlled in response to a variety of cellular factors. This would be expected of a gene that encodes peptides expressed in several behaviorally important neural systems.

Expression of the myomodulin gene in the CNS

The presence of the myomodulin transcripts within the CNS of *Lymnaea* was confirmed by hybridization of the cDNA clone pMM421 to size-fractionated total CNS RNA (Fig. 4*b*). After a 12 hr exposure of X-ray film, four strongly hybridizing bands of 2.6, 2.8, 3.2, and 3.8 kb were visible. The two smaller bands appeared as a single intense band, but shorter exposures and densitometer readings allowed them to be distinguished. After longer exposures (36–72 hr), two additional bands of 2.1 and 1.5 kb were also observed. The 2.1, 2.6, and 2.8 kb species of transcript detected correspond well with the mRNA lengths expected for the three classes of myomodulin cDNA isolated from the cDNA library. The 3.2, 3.8, and 1.5 kb transcripts may represent other mRNA species transcribed either from the gene described here, utilizing cryptic promoter or polyadenylation signals, or from a second gene with high sequence homology. The latter explanation seems unlikely in view of the Southern hybridization data. The relative amounts of the different transcripts can also be estimated from the intensity of the bands, which suggests that the class II and III transcripts are of approximately the same abundance in the CNS, whereas the shorter class I transcripts are 10-fold less abundant.

Myomodulin gene transcripts were also directly detected in *Lymnaea* nervous tissue by *in situ* hybridization of an antisense digoxigenin-labeled 25 nucleotide oligomer to fixed 7 μ m sections of the CNS. Transcripts were detected in the cytoplasm of ~1000 neurons located in all ganglia of the brain, suggesting that the gene is widely and abundantly expressed (Fig. 6). Variation in transcript concentration within neurons was indicated by differences in the intensity of staining observed. Complementary strand probes and RNase-treated sections both failed to show any staining, indicating that the observed hybridization with the antisense probe was specific for myomodulin gene transcripts (data not shown).

Detection of myomodulin peptides in nervous tissue

Mass spectrometric analysis of a small piece of the external right parietal nerve (Fig. 7) revealed the presence of a number of peptides, including two with masses identical to some of the predicted myomodulin peptides: 829.8 (GLQMLRLamide) and 846.7 (PMSMLRLamide). The detection of the peptide GLQMLRLamide confirms that the single basic amino acid at its N terminus, predicted from the cDNA sequence, is used as an endoproteolytic cleavage site. The three other myomodulin-like peptides encoded within the gene were not detected (SMSMLRLamide, SLSMLRLamide, and QIPMLRLamide), but the ionization capabilities of these peptides may be lower than those of the other peptides, rendering them undetectable at low concentrations. Control experiments carried out in this laboratory have shown that the myomodulin peptides vary in their detectability by MALDI-MS, which are affected by a number of factors including the concentration of other peptides within the mixture (data not shown). However, these experiments showed that all of the different members of the *Lymnaea* myomodulin family are capable

Figure 5. Analysis of potential promoter elements in the myomodulin gene promoter. Genomic clone pGMM28A is an *EcoRI* restriction fragment (*E*) consisting of 1100 bp of DNA 5' to the transcriptional start sites (white bar and horizontal arrow labeled +1), a 645 bp exon (*Exon I*: shaded bar) and 842 bp of intron (line). The sequence of the 5' splice junction is shown. The promoter region is expanded to show the predicted promoter elements and their positions with respect to the transcriptional start point (see Results).

of detection by MALDI-MS, and the lack of detection of some of them in the present tissue probably reflects their low concentration.

DISCUSSION

The results presented here describe the isolation of cDNA and genomic DNA clones encoding the myomodulin family of neuropeptides in *Lymnaea*. Expression of the gene in the nervous system is confirmed by RNA blot analysis, transcriptional start site mapping, and *in situ* hybridization. Translation of transcripts and processing of the propeptide are also confirmed by detection of some of the predicted myomodulin peptides within nervous tissue.

Myomodulin gene organization

The myomodulin gene contains a single intron of ~19 kb that lies within the 20th codon of the major open reading frame. An intron of unknown size has also been detected in the same position of the *Aplysia* myomodulin gene (Miller et al., 1993). The positioning of the intron suggests that it may play a role in the control of gene expression because failure to remove the intron from transcripts would disrupt the open reading frame and prevent synthesis of the myomodulin prepropeptide. The 20th codon of the open reading frame is predicted to encode the last amino acid of the hydrophobic leader sequence of the peptide precursor. It is possible, therefore, that there is alternative splicing between the first exon and alternative downstream exons similar to the mechanism that is known to exist for the FMRFamide neuropeptide gene in *Lymnaea* (Saunders et al., 1991, 1992). The large size of the intron theoretically could accommodate other exons; however, no alternatively spliced exons or cDNAs have been characterized for the myomodulin gene despite extensive efforts to identify them (our unpublished observations).

Transcriptional control of the myomodulin gene

Isolation of the promoter region of the gene has revealed several promoter elements that may play roles in the control of transcription of the gene. A putative TATA sequence 28 bp upstream of the transcriptional start site has been identified, but other com-

Figure 6. Localization of myomodulin transcripts to neurons in the CNS of *Lymnaea*. *In situ* hybridization with a 25 mer oligonucleotide complementary to part of the myomodulin cDNA sequence revealed the presence of transcripts in the cytoplasm of a subset of neurons in all of the ring ganglia and the buccal ganglia. *a*, Neurons in the left and right parietal ganglia (LPG and RPG, respectively) and the visceral ganglia (VG) exhibit intense staining. *b*, Neurons in both the left and the right pedal ganglia (PEG) show a range of staining intensities, indicating different levels of expression in neighboring cells. Scale bar, 100 μ m.

Figure 7. MALDI-MS analysis of a section of the external right parietal nerve reveals the presence of two of the myomodulin peptides predicted from the cDNA sequence. Mass peaks corresponding to *PMSMLRLamide* (846.7 Da) and *GLQMLRLamide* (829.8 Da) are labeled.

mon transcription promoter elements, CAAT box and SP-1, for example, are absent. The *in situ* hybridization data revealed a cell-specific pattern of gene expression that discriminated between neighboring cells within the same ganglia in the brain. Production of such a specific pattern of transcription would require the absence of these nonspecific transcriptional inducers and the presence of cell-specific DNA binding factors and the sequence elements that they occupy. Such sequence elements have been identified in the myomodulin gene promoter (Fig. 5), supporting the suggestion of an array of cell-specific transcription factors controlling the expression of the gene.

Two types of cAMP-responsive element have been identified: CRE and AP-2. Both of these elements can act as basal enhancers and as rapid-response cAMP-dependent enhancers (Imagawa et al., 1987; Quinn et al., 1988; Roesler et al., 1988). These elements may influence the level of myomodulin gene expression within neurons by increasing the rate of transcription when cAMP levels rise. Such increases could be produced by activation of adenylate cyclases by neurotransmitters binding to G-protein-coupled receptors on the surface of the cell. Therefore, the rate of transcription could be influenced by the level of agonist stimulation of the cell. Vertebrate neuropeptide and neurohormone gene promoters contain these sequence elements (somatostatin: Montminy et al., 1986; VIP: Tsukada et al., 1987; proenkephalin: Comb et al., 1986), suggesting a role for cAMP in the control of gene expression for a wide range of neural signaling pathways.

Transcript and prepropeptide structure

Three different classes of myomodulin cDNA were isolated, each characterized by the position of the polyadenylation signal used for correct 3' end processing. Although the use of alternative polyadenylation signals does not alter the structure of the prepropeptide encoded by the transcripts, it may have significant effects on the stability of the mRNA and its ability to be trans-

lated. Sequence elements that affect both of these properties of transcripts have been reported to exist in the 3'-untranslated regions of other gene transcripts (for review, see Klausner and Harford, 1989; Jackson and Standart, 1990). The different quantities of the three classes of transcript detected by Northern hybridization could be accounted for by these factors, or simply by a bias in the use of the sites.

The predicted structure of the prepropeptide encodes a 350-amino-acid polypeptide encoding 14 putative myomodulin-like peptides of five different structures. Each peptide is flanked by endoprotease cleavage sites and has a glycine residue at its C terminus, suggesting that the peptides are all amidated after cleavage from the propeptide. The detection of two of these peptides by mass spectrometry confirms that both cleavage and amidation of the peptides do occur. Interestingly, detection of the peptide *GLQMLRLamide* also confirms that the single basic amino acid positioned at its N terminus within the propeptide does act as a functional endoproteolytic cleavage site. This peptide is not found in the *Aplysia* precursor cDNA, but it has been biochemically isolated from the ARC muscle (Brezina et al., 1995). Failure to detect the other three myomodulin-like peptides predicted from the propeptide structure (*SLSMLRLamide*, *SMSMLRLamide*, and *QIPMLRLamide*) may imply that these peptides are not processed from the polypeptide. This seems unlikely because all three peptides are flanked by endoproteolytic cleavage sites. The quantities of the detected peptides were very low (<10 fmol), and the quantities of the other three peptides may have been too low even for detection by the sensitive technique of MALDI-MS. The ability of a peptide to be ionized by laser desorption mass spectrometry is dependent on its structure. Even structurally similar peptides can have very different ionization capabilities; therefore, the three peptides not detected may be less easily ionized than the other two myomodulins. It is likely that all five peptide types are cleaved from the propeptide and, further, biochemical and mass spectrometric analysis of nervous tissue would probably detect all five myomodulins.

Two putative furin endoprotease cleavage sites (residues 41–44 and 114–117) within the propeptide may act as sites of primary processing in the *trans*-Golgi network. The most likely cleavage site, RFRR (residues 114–117), is found within the spacer sequence between the first two myomodulins, *GLQMLRLamide* and *QIPMLRLamide*. This putative cleavage site is similar to the tetrabasic cleavage sequence observed in both the *Lymnaea* *FMRFamide* (Linacre et al., 1990) and the *Aplysia* *ELH* precursor proteins, where it has been demonstrated to be the first site of processing, allowing the differential sorting and trafficking of peptides encoded at opposite ends of the same precursor protein (Sossin et al., 1990). A possible cleavage site is also present at an equivalent position in the *Aplysia* myomodulin precursor (Lopez et al., 1993; Miller et al., 1993).

This study has revealed the presence of the myomodulin-like peptides in the nervous system of *Lymnaea*. The elucidation of the cDNA sequences and genomic organization has revealed some similarities to the myomodulin gene of *Aplysia* (for example, three of the putative peptide sequences are identical), but there are also several novel features of this gene and the polyprotein it encodes. The gene has a promoter containing sequence elements that may confer tissue specificity and cAMP induction of expression. Alternative use of polyadenylation sites during post-transcriptional processing of the transcript occurs, with three different lengths of 3'-untranslated region existing in detectable quantities within the nervous system. Three novel peptide structures are described

here, and evidence for the translation and post-translational processing of the polypeptide is presented.

The myomodulins are an important class of neuropeptides within the nervous system of *Lymnaea* and, coupled with the detection of myomodulin immunoreactivity in several identified neurons involved in well studied behavioral networks (Santama et al., 1994b), these peptides will allow the characterization of a central role for a family of neuropeptides in the mediation of behavior.

REFERENCES

- Benjamin PR, Burke JF (1994) Alternative mRNA splicing of the FMRFamide gene and its role in neuropeptidergic signalling in a defined neural network. *Bioessays* 16:335–342.
- Birnstiel ML, Busslinger M, Strub K (1985) Transcription termination and 3' processing: the end is in site! *Cell* 41:349–359.
- Bradbury AF, Finnie MDA, Smyth DG (1982) Mechanism of C-terminal amide formation by pituitary enzymes. *Nature* 298:686–688.
- Brezina V, Evans CG, Weiss KR (1994a) Enhancement of Ca current in the accessory radula closer muscle of *Aplysia californica* by neuromodulators that potentiate its contractions. *J Neurosci* 14:4393–4411.
- Brezina V, Evans CG, Weiss KR (1994b) Activation of K current in the accessory radula closer muscle of *Aplysia californica* by neuromodulators that depress its contractions. *J Neurosci* 14:4412–4432.
- Brezina V, Bank B, Cropper EC, Rosen S, Vilim FS, Kupfermann I, Weiss KR (1995) Nine members of the myomodulin family of peptide cotransmitters at the B16-ARC neuromuscular junction of *Aplysia*. *J Neurophysiol* 74:54–72.
- Comb M, Birnberg NC, Seasholtz A, Herbert E, Goodman HM (1986) A cyclic AMP- and phorbol ester-inducible DNA element. *Nature* 323:353–356.
- Cropper EC, Tenenbaum R, Kolks MAG, Kupfermann I, Weiss KR (1987) Myomodulin: a bioactive neuropeptide present in an identified cholinergic buccal motor neuron of *Aplysia*. *Proc Natl Acad Sci USA* 84:5483–5486.
- Cropper EC, Vilim FS, Alevizos A, Tenenbaum R, Kolks MAG, Rosen S, Kupfermann I, Weiss KR (1991) Structure, bioactivity and cellular localization of myomodulin B: a novel *Aplysia* peptide. *Peptides* 12:683–690.
- DesGroseillers L, Cowan D, Miles M, Sweet A, Scheller RH (1987) *Aplysia californica* neurons express microinjected neuropeptide genes. *Mol Cell Biol* 7:2762–2771.
- Fugisawa Y, Kubota I, Kanda T, Kuroki Y, Muneoka Y (1990) Neuropeptides isolated from *Mytilus edulis* (bivalvia) and *Fusinus ferrugineus* (prosobranchia). In: *Comparative pharmacology of neuropeptides* (Stefano GB, Florey E, eds). Manchester: Manchester UP.
- Guntaka RV (1993) Transcription termination and polyadenylation in retroviruses. *Microbiol Rev* 57:511–521.
- Hopp TP (1986) Protein surface analysis: methods for identifying antigenic determinants and other interaction sites. *J Immunol Methods* 88:1–18.
- Hosaka M, Nagahama M, Kim WS, Watanabe T, Hatsuzawa K, Ikemizu J, Murakami K, Nakayama K (1991) Arg-X-Lys/Arg-Arg motif as a signal for precursor cleavage catalysed by furin within the constitutive secretory pathway. *J Biol Chem* 266:12127–12130.
- Imagawa M, Chui R, Karin M (1987) Transcription factor AP-2 mediates induction by two different signal-transduction pathways: protein kinase C and cAMP. *Cell* 51:251–260.
- Jacob M, Gallinar H (1989) The 5' splice site: phylogenetic evolution and variable geometry of association with U1RNA. *Nucleic Acids Res* 17:2159–2180.
- Jackson RJ, Standart N (1990) Do the polyA tail and 3' untranslated region control messenger-RNA translation. *Cell* 62:15–24.
- Klausner RD, Harford JB (1989) *Cis-trans* models for post-transcriptional gene-regulation. *Science* 246:870–872.
- Kozak M (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger-RNAs. *Nucleic Acids Res* 15:8125–8148.
- Kyte J, Doolittle RF (1982) A simple method for displaying the hydrophobic character of a protein. *J Mol Biol* 157:105–132.
- Li KW, van Golen FA, van Minnen J, van Veelen PA, van der Greef J, Geraerts WPM (1994) Structural identification, neuronal synthesis, and role in male copulation of myomodulin-A of *Lymnaea*: a study involving direct peptide profiling of nervous tissue by mass spectrometry. *Mol Brain Res* 25:355–358.
- Linacre A, Kellett E, Saunders S, Bright K, Benjamin PR, Burke JF (1990) Cardioactive neuropeptide Phe-Met-Arg-Phe-NH₂ (FMRFamide) and novel related peptides are encoded in multiple copies by a single gene in the snail *Lymnaea stagnalis*. *J Neurosci* 10:412–419.
- Locker J, Buzzard G (1990) A dictionary of transcription control sequences: DNA sequence. *J DNA Sequencing Mapp* 1:3–11.
- Lopez V, Wickham L, DesGroseillers L (1993) Molecular cloning of myomodulin cDNA, a neuropeptide precursor gene expressed in neuron L10 of *Aplysia californica*. *DNA Cell Biol* 12:53–61.
- Miller MW, Beushausen S, Vitek A, Stamm S, Kupfermann I, Brosius J, Weiss KR (1993) The myomodulin-related neuropeptides: characterization of a gene encoding a family of peptide cotransmitters in *Aplysia*. *J Neurosci* 13:3358–3367.
- Montminy MR, Sevarino KA, Wagner JA, Mandel G, Goodman RH (1986) Identification of a cyclic-AMP-responsive element within the rat somatostatin gene. *Proc Natl Acad Sci USA* 83:6682–6686.
- Newcomb R, Fisher JM, Scheller RH (1988) Processing of the egg-laying hormone (ELH) precursor in the bag cell neurons of *Aplysia*. *J Biol Chem* 263:12514–12521.
- Pellé R, Murphy NB (1993) Northern hybridization: rapid and simple electrophoretic conditions. *Nucleic Acids Res* 2:2783–2784.
- Proudfoot N (1991) Poly(A) signals. *Cell* 64:671–674.
- Quinn PG, Wong TW, Magnuson MA, Shabb JB, Granner DK (1988) Identification of basal and cyclic AMP regulatory elements in the promoter of the phosphoenolpyruvate carboxykinase gene. *Mol Cell Biol* 8:3467–3475.
- Roesler WJ, Vandenbark GR, Hanson RW (1988) Cyclic AMP and the induction of eukaryotic gene transcription. *J Biol Chem* 263:9063–9066.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Santama N, Wheeler CH, Burke JF, Benjamin PR (1994a) Neuropeptides myomodulin, small cardioactive peptide, and buccalin in the central nervous system of *Lymnaea stagnalis*: purification, immunoreactivity and artifacts. *J Comp Neurol* 342:335–351.
- Santama N, Brierley M, Burke JF, Benjamin PR (1994b) Neural network controlling feeding in *Lymnaea stagnalis*: immunocytochemical localization of myomodulin, small cardioactive peptide, buccalin, and FMRFamide-related peptides. *J Comp Neurol* 342:352–365.
- Saunders SE, Bright K, Kellett E, Benjamin PR, Burke JF (1991) Neuropeptides Gly-Asp-Pro-Phe-Leu-Arg-Phe-amide (GDPFLRFamide) and Ser-Asp-Pro-Phe-Leu-Arg-Phe-amide (SDPFLRFamide) are encoded by an exon 3' to Phe-Met-Arg-Phe-amide (FMRFamide) in the snail *Lymnaea stagnalis*. *J Neurosci* 11:740–745.
- Saunders SE, Kellett E, Bright K, Benjamin PR, Burke JF (1992) Cell specific alternative RNA splicing of a FMRFamide gene transcript in the brain. *J Neurosci* 12:1033–1039.
- Sossin WS, Sweet-Cordero A, Scheller RH (1990) Dale's hypothesis revisited: different neuropeptides derived from a common prohormone are targeted to different processes. *Proc Natl Acad Sci USA* 87:4845–4848.
- Tsukada T, Fink JS, Mandel G, Goodman RH (1987) Identification of a region in the human vasoactive intestinal polypeptide gene responsible for regulation by cyclic-AMP. *J Biol Chem* 262:8743–8747.
- van Golen FA, Li KW, Chen S, Jimenez CR, Geraerts WPM (1996) Various isoforms of myomodulin identified from the male copulatory organ of *Lymnaea* show overlapping yet distinct modulatory effects on the penis muscle. *J Neurochem* 66:321–329.
- von Heijne G (1986) A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res* 14:4683–4690.
- Vreugdenhil E, Jackson JF, Bouwmeester T, Smit AB, van Minnen J, van Heerikhuizen H, Klootwijk J, Joosse J (1988) Isolation, characterization, and evolutionary aspects of a cDNA clone encoding multiple neuropeptides involved in the stereotyped egg-laying behaviour of the fresh-water snail *Lymnaea stagnalis*. *J Neurosci* 8:4184–4191.