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*

A. Linacre,² E. Kellett,¹ S. Saunders,¹ K. Bright,² P.R. Benjamin,² and J.F. Burke¹

Departments of ¹Biochemistry and ²Sussex Invertebrate Neuroscience Group, School of Biological Sciences, Falmer, Brighton, East Sussex, BN1 9QG, United Kingdom

The neuropeptide Phe-Met-Arg-Phe-NH₂ (FMRFamide) is a potent cardioactive neuropeptide in *Lymnaea stagnalis*. Isolation and sequencing of 2 cDNAs and a genomic clone shows that a single gene encodes a precursor protein which contains 9 copies of the FMRFamide peptide, 2 copies of the related peptide Phe-Leu-Arg-Phe-NH₂ (FLRFamide), and single copies of the putative pentapeptides Gln-Phe-Tyr-Arg-Ile-NH₂ (posttranslationally modified to pQFYRIamide) and Glu-Phe-Leu-Arg-Ile-NH₂ (EFLRIamide). The gene is transcribed in the CNS and gives rise to a single RNA of 1.7 kb in size. The organization of the *Lymnaea* gene is significant with respect to the evolution of FMRFamide and related peptides in other organisms.

The discovery that neuropeptides can act as neurotransmitters, neuromodulators, or hormones has had a great impact on the field of neurobiology. We are interested in understanding the role of identified neuropeptides in neuronal circuits of known function using the pulmonate mollusc, *Lymnaea stagnalis*. This snail has previously been used as a model system for analyzing the way in which peptides mediate simple behaviors (Vreugdenhil et al., 1988) and for studies on the structure and organization of invertebrate neuropeptide genes (Smit et al., 1988). In this study we describe the structure of the molluscan gene encoding the neuropeptide Phe-Met-Arg-Phe-NH₂ (FMRFamide).

FMRFamide was first discovered as a cardioexcitatory agent in the venus clam *Macrocallista nimbosa* (Price and Greenberg, 1977), and although the distribution of FMRFamide itself appears to be limited to molluscs (Price, 1986), immunoreactivity to FMRFamide has been reported in all the major animal phyla and a number of FMRFamide-like peptides have been isolated from coelenterates (Grimmelikhuijzen and Graff, 1985), arthropods (Boer et al., 1980), and chordates (Dockray et al., 1983; Yang et al., 1985), indicating that FMRFamide is one of a family of related neuropeptides. However, the antibody used in many of these surveys does cross-react with other Arg-Phe-NH₂ peptides.

FMRFamide-like immunoreactivity has been found in a number of neurons in the CNS of *Lymnaea* (Schot and Boer, 1982), and recent work using immunocytochemistry and radioimmunoassay has identified a pair of heart motoneurons whose excitatory effects on the heart appear to be due to the release of FMRFamide or a related peptide (Buckett, 1987; Benjamin et al., 1988). The electrophysiological evidence combined with pharmacology strongly indicated that FMRFamide is the principal excitatory neurotransmitter peptide used by *Lymnaea* heart motoneurons. In addition, a uniquely identifiable interneuron (VWI) has been found in the *Lymnaea* CNS which has widespread synaptic connections with other central neurons (Benjamin, 1984) and also contains FMRFamide-like peptides (Benjamin et al., 1988). It has been proposed that FMRFamide-like peptides might mediate these synaptic connections. While a physiological role for FMRFamide in cardioexcitation in *Lymnaea* is now well established, FMRFamide has a variety of actions in other systems. It has been shown to modulate the activity of specific central neurons (Stone and Mayeri, 1981), as well as the L2 neuron (Thompson and Ruben, 1988) and R14 neuron (Ichinose and McAdoo, 1988) in *Aplysia*. FMRFamide causes several kinds of membrane conductance changes in the CNS (Cottrell et al., 1984; Boyd and Walker, 1985) and shortens the presynaptic action potential in sensory neurons (Abrahams et al., 1984). FMRFamide has also been shown to act in peripheral systems by inhibiting spontaneous activity of the gut (Austin et al., 1982) and salivary gland (Bullock et al., 1988). It also stimulates contractions of the buccal mass (Richmond et al., 1984) and gill (Weiss et al., 1984) in *Aplysia*.

FMRFamide is not the only member of this family of neuropeptides to have been shown to be present in *Lymnaea*. Ebberink et al. (1987) have isolated 2 heptapeptides from the central ganglia, SDPFLRFamide (Ser-Asp-Pro-Phe-Leu-Arg-Phe-NH₂) and GDPFLRFamide (Gly- . . . etc.). In previous investigations of the genetic organization of families of structurally related neuropeptides, various forms have been found to be encoded by the same gene and are present on the same precursor. The question therefore arises in *Lymnaea* as to whether the 3 known forms of FMRFamide-related peptides are present on the same gene and/or different transcripts.

We report here on the isolation of 2 cDNA clones and a genomic clone that encode 9 copies of FMRFamide in addition to novel forms of FMRFamide-like peptides. The *Lymnaea* genome contains a single copy of the gene that gives rise to a unique transcript within the CNS. Our data from *in situ* hy-

Received June 9, 1989; revised July 19, 1989; accepted July 28, 1989.

This work was supported by MRC project grant G8622759N and SERC studentships to A.L. and K.B. We thank Dr. Mike Powell for help with the RNA analysis and Nicola Ford for preparation of the manuscript.

Correspondence should be addressed to Dr. J. F. Burke at the above address.

Copyright © 1990 Society for Neuroscience 0270-6474/90/100412-08\$02.00/0

bridization suggest that the gene is expressed in specific neurons throughout the *Lymnaea* CNS.

Materials and Methods

Molecular procedures. General procedures such as growth of phage, Southern transfers, and plasmid preparations were carried out as described by Maniatis et al. (1982).

Construction of oligonucleotide. Mixed oligonucleotides were synthesized on an Applied Biosystems 381A DNA synthesizer according to the manufacturer's instructions. The sequence of the oligonucleotide was (5' to 3') (I/C)TTNCC(A/G)AANC(G/T)CAT(A/G)AA (where N represents A, G, C, or T).

Extraction of genomic DNA. The ring of ganglia that forms the CNS (the "brain") was dissected from animals and non-neuronal tissue and mucus strands carefully removed from the ganglia using forceps. The brains were then washed in *Lymnaea* saline (Benjamin and Winlow, 1981) (3 changes). The tissues (10 brains/100 μ l) were gently homogenized in 1 $\times$ PK buffer (Burke and Ish-Horowicz, 1982) containing 2% SDS. Two drops of Antifoam A (Sigma) were added per 100 μ l of buffer. This was followed by phenol extraction and ethanol precipitation of the DNA, which was found to be of high molecular weight and free from contaminating nucleases.

Isolation of FMRFamide-encoding cDNA and genomic clones. Approximately 20,000 plaques from an amplified λ gt10 *Lymnaea* CNS cDNA library (Vreugdenhil et al., 1988) were bound to Hybond-N (Amersham) and screened with the mixed oligonucleotides labeled with γ - 32 P-ATP using T4 polynucleotide kinase. Hybridizations were carried out at 37°C in a solution containing 6 $\times$ SSC, 5 $\times$ Denhardt's solution, and 0.5% SDS. Filters were washed at 42°C in 6 $\times$ SSC, 0.5% SDS, dried, and autoradiographed. Two positive clones, designated D1 and D3, were obtained. When the entire *Eco*RI insert of D1 was subcloned into either M13mp18 or M13mp19, it was found to be unstable and readily deleted. The *Eco*RI inserts of D1 and D3 were subcloned into Bluescribe (Stratagene) and subsequently digested independently with the restriction enzymes *Taq*I and *Sau*3A. The smaller fragments that this generated were ligated into M13mp18 and M13mp19 for single-strand sequencing.

To isolate a genomic clone the D3 insert was hexamer labeled (Feinberg & Vogelstein, 1983) and used to screen 20,000 clones of an *Eco*RI genomic library in λ gt10. A positive clone containing a single insert of 4.1 kb was identified and isolated. The DNA from this clone was digested by *Eco*RI and *Bam*HI, and the resulting fragments were ligated into M13mp18 and M13mp19.

Nucleotide sequence analysis. Phage inserts, or fragments derived from them, were subcloned into either Bluescribe or M13mp18/19 vectors for sequence analysis in both orientations and were also double-strand sequenced. Sequencing using α - 35 S-dATP was performed as described by Sanger et al. (1977) or by using a commercially available kit (Sequenase, U.S. Biochemicals). The samples were separated on a 7 M urea, 6% acrylamide gel. Ambiguities were resolved by resequencing using synthetic oligonucleotides derived from known sequences, 5' CCA-AGCCGCCAGACTCGTTGGCC 3' and 5' GGAAGAGATAT-GAGTGACGTGGAC 3'.

DNA blot hybridization. Genomic DNA, 5 μ g, was fully digested with a restriction enzyme and separated on a 0.5% agarose gel. The DNA was then transferred to Hybond-N (Amersham). The D3 cDNA insert was hexamer labeled in the presence of α - 32 P-dCTP and hybridized at 65°C overnight. The filter was washed in 0.3 $\times$ SSC (Maniatis et al., 1982), 0.5% wt/vol SDS, at 65°C, and autoradiographed.

RNA blot hybridizations. Total RNA was isolated from 10 CNS by the standard method of homogenizing in 2 $\times$ PK buffer followed by phenol extraction. Some 10 μ g of total RNA was run on a formaldehyde gel essentially as described in Maniatis et al. (1982) and transferred to Hybond-N. The D3 cDNA insert clone was hexamer labeled in the presence of α - 32 P-dCTP and hybridized to the filter at 42°C overnight in 4 $\times$ SSC, 10 μ g/ml denatured salmon sperm DNA, 0.1 $\times$ Denhardt's solution, 12.5% formamide, 10% wt/vol Dextran sulfate. The filter was washed in 2 $\times$ SSC at 60°C and then autoradiographed.

In situ hybridizations. CNS from *Lymnaea* were dissected, fixed, and sectioned essentially as described by Dirks et al. (1989). Sections were hybridized with α - 35 S-dATP nick-translated D3 cDNA insert overnight at 37°C. Slides were exposed for autoradiography for 3 weeks.

Protein sequence comparisons. Comparisons and alignments were carried out using the Lipman and Pearson (1985) FASTP program.

Results

Isolation of FMRFamide cDNA clones

A mixture of 256 different degenerate oligonucleotides, complementary to the coding sequence for Phe-Met-Arg-Phe-Gly plus the first 2 complementary bases for lysine, was used to screen a *Lymnaea* CNS cDNA library (described in Vreugdenhil et al., 1988, and kindly donated by E. Vreugdenhil). The codon for Gly was included as this amino acid is thought to be required for amidation (Loh and Gainer, 1983) and Lys is an amino acid commonly found at a site of proteolytic cleavage within hormone precursor molecules (Douglass et al., 1984). Approximately 20,000 recombinant clones were screened with the radiolabeled oligonucleotide from which 2 independent clones were isolated. Subsequent restriction analysis indicated that the 2 cDNAs contained inserts of 1.2 and 0.9 kb. The 2 clones were designated D1 and D3, respectively. Nucleotide sequence data revealed that both D1 and D3 contained sequences that encode FMRFamide. As neither cDNA included an initiator methionine that was in frame with the FMRF sequences, it was concluded that both D1 and D3 were partial cDNA clones. An additional 20,000 clones were screened with hexamer labeled D1 but no further positives were identified.

Isolation of a FMRFamide genomic clone

Southern blot analysis (Fig. 3) had indicated the presence of a single *Eco*RI genomic fragment of approximately 4.1 kb in size containing FMRFamide sequences. A *Lymnaea* genomic library was constructed from *Eco*RI digested DNA in λ gt10 from which approximately 20,000 clones were screened. This resulted in the isolation of one positive phage which hybridized with D1 and D3 and was found to contain an insert of the expected size (4.1 kb). A 0.9 kb *Eco*RI/*Bam*HI restriction fragment from this genomic clone was found by hybridization and sequence analysis to encode the 5' end of the cDNAs including an in frame AUG start codon.

Deduced structure of the FMRFamide mRNA preprohormone

The D1 and D3 cDNAs were completely sequenced by the strategy outlined in Figure 1. This showed that the D3 sequence was contained within D1 and was identical except that D3 contained a run of 17As (position 986, Fig. 2), whereas D1 contained 14. The overlap between the cDNA clones D1, D3, and the genomic clone is also illustrated in Figure 1. The longest open reading frame (ORF) in the 2 cDNA clones contained the FMRFamide sequences (Fig. 2). There was no N-terminal methionine in either cDNA, indicating that both were partial at the 5' end. The cDNA clone D1 contains 11 nucleotides at the extreme 5' end which do not correspond to the genomic sequence. This could be a polymorphism or evidence of splicing or a cloning artifact. The first stop codon (TAA) in the FMRFamide reading frame is at position 942–944. This is followed by multiple stop codons in all 3 reading frames and a stretch of 14As starting at nucleotide 986. This is unlikely to be the posttranslational polyA tail as extra nucleotides are present 3' to this. Furthermore, from Southern analysis (data not shown) these 3' sequences are contiguous in the genome with sequences in the FMRFamide ORF. The FMRFamide ORF in the λ gt10 genomic clone extends 5' to a stop codon at nucleotide 51–53, suggesting that the methionine at 82–84 is the initiating methionine. As is typical of many membrane or secreted proteins, a signal sequence containing, in this case, 6 hydrophobic

Figure 1. Organization of cloned sequences. *A*, Alignment of λ gt10 genomic clone and cDNAs showing regions sequenced (—). *B*, Part of the strategy for sequencing the insert of cDNA D1. Arrows indicate direction and extent of sequencing runs for *TaqI* (T) and *Sau3A* (S) fragments cloned into M13mp18/19 (B, *BamHI*).

amino acids follows the methionine codon. This N-terminal sequence is significant (Kyte and Doolittle, 1982) as it is the only hydrophobic region in the whole precursor protein. Interestingly, 6 hydrophobic amino acids makes the *Lymnaea* signal sequence the least hydrophobic of over 300 eukaryotic signal sequences recently surveyed (von Heijne, 1985, 1986). The hydrophobic leader also conforms to the -1, -3 rule for signal peptide cleavage. Residue -1 must be small and residue -3 should not be aromatic, charged, or large and polar. In the precursor, Thr (-1) and Ala (-3) at positions 18 and 16, respectively, fulfill these criteria, predicting cleavage between Thr and Lys at amino acids 18–19. It cannot be discounted that the first potential cleavage site of Lys-Arg at positions 19–20 may act as the site for removal of the signal sequence.

Nine copies of Phe-Met-Arg-Phe (FMRF) are contained on the precursor (Fig. 2). Eight of the FMRF sequences are flanked on their amino termini by Lys-Arg, the other by Lys-Ser. The carboxy termini of the 9 FMRF sequences contain the residues Gly-Arg (4 out of 9), Gly-Lys-Ser (4 out of 9), and Gly-Lys-Arg (1 out of 9). The Gly is required for amidation of the peptides and the basic residues, which serve as proteolytic cleavage signals, are similar to those found on the *Aplysia* FMRFamide precursor (Taussig and Scheller, 1986). Spacer peptides separating the 9 FMRF sequences vary in size and composition of amino acids and probably do not have any biological function other than to separate the FMRFamide sequences.

The precursor protein also contains 2 Phe-Leu-Arg-Phe (FLRF) sequences at positions 21–24 and 57–60. These sequences are flanked by a probable cleavage signal Lys-Arg on the amino termini and Gly-Arg on the carboxy termini and would therefore be amidated if cleaved from the precursor.

Other putative peptides may be contained on the precursor protein. These included the pentapeptides Gln-Phe-Tyr-Arg-Ile (QFYRI) at position 41–44 which was bounded by Arg-Arg and Gly-Arg at its amino and carboxy termini, respectively, and Glu-Phe-Leu-Arg-Ile (EFLRI) at position 262–266 and bounded by the cleavage signals Arg-Arg on the amino side and Gly-Arg on the carboxy side. These 2 pentapeptides would be amidated if processed from the precursor and glutamine if present at the amino terminal of a peptide can circularize to form pyroglutamine. This would produce pQFYRIamide and EFLRIamide.

A single gene encodes the FMRFamide precursor

Genomic DNA was digested with various restriction enzymes, blotted, and probed with hexamer labeled D3 cDNA insert. The

result of this Southern blot is shown in Figure 3. Single digests with *DraI*, *BamHI*, and *EcoRI* when hybridized with D3 give single fragments of 1.1, 3.4, and 4.1 kb. A double digest with *DraI/BamHI* gives a single fragment of 1.0 kb and a double digest with *BamHI/EcoRI* gives 3.2 kb fragment. The observation that the *BamHI* site is asymmetric in the 4.1 kb *EcoRI* fragment which is present only once in the genome and that *BamHI*, which also cuts in the cDNA, produced a single fragment that hybridizes with D3, suggests that the sequence encoding D3 is present only once in the genome. Furthermore, the size of the bands was consistent with the size of the genomic fragments prepared from the *EcoRI* λ gt10 genomic library, suggesting that there are no large introns in the genomic sequence.

Transcription of the FMRFamide precursor

Northern blot analysis (Fig. 4) was used to indicate the number and size of transcripts produced by the FMRFamide gene in the CNS. Total RNA was isolated from 10 complete CNS, fractionated according to size, transferred to Hybond-N, and hybridized to D3 cDNA insert. The result shown in Figure 4 indicates that this gene gives rise to a unique 1.7 kb mRNA species within the CNS.

The FMRFamide gene is expressed throughout the Lymnaea CNS

In order to determine more precisely the location of the FMRFamide transcripts in the CNS, *in situ* hybridization was performed (Fig. 5). Preliminary analysis shows that at least 100 neurons express the FMRFamide gene and these neurons are distributed throughout the CNS.

Discussion

We have characterized 2 cDNA clones and a genomic clone that encodes 9 copies of FMRFamide, 2 copies of FLRFamide, and 2 novel pentapeptides (summarized in Fig. 6). We have yet to determine how many of these individual peptides are released from the precursor protein. However, the basic amino acids that serve as recognition sites for cleavage are similar to those of the *Aplysia* FMRFamide precursor with the exception of the Lys-Ser seen on the amino side of one FMRFamide. *In vivo* labeling experiments using ^3H -phenylalanine have indicated that most of the 28 FMRFamides encoded by the *Aplysia* FMRFamide gene are cleaved from the precursor (Taussig and Scheller, 1986). Therefore, it is a reasonable assumption that most of the peptides on the *Lymnaea* precursor will also be cleaved.

The presence of the highly related peptide FLRFamide has

```

1
ATTGTAACTTTACTTTAGGGTTGTAATATGTATGTATGCAAGCTGTTTTTAATTTAAAA
60

GTCTTTCCAGTTTTTTTGTATTGTATTTCCCAACACTGATAGTTTGTTCCTTTCTTT
120
MetTyrSerProThrLeuIleValCysLeuSerPhePhe
1
CACTCTGCAGTGACCAAAACGCTTTTGTAGGTTTGGCCGAGCTCTGGACACTACGGATCCT
180
HisSerAlaValThrLysArgPheLeuArgPheGlyArgAlaLeuAspThrThrAspPro
20
TTCATACGGTTGAGAAGGCAGTTCTATCGAATTGGCCGAGGGGGCTATCAGCCTTACCAG
240
PheIleArgLeuArgArgGlnPheTyrArgIleGlyArgGlyGlyTyrGlnProTyrGln
40
GACAAGAGATTCTTGGGTTCTGGCCGGTCTGAGCAACCGGATGTCGATGACTACCTTACA
300
AspLysArgPheLeuArgPheGlyArgSerGluGlnProAspValAspAspTyrProArg
60
GATGTTGTCTCCAGTCAGAGGAGCCGCTGTACAGAAAACGGAGATCCACGGAGGCGGGT
360
AspValValLeuGlnSerGluGluProLeuTyrArgLysArgArgSerThrGluAlaGly
80
GGCCAATCAGAAGAGATGACACACAGAACAGCGAGATCAGCTCCCGAGCCAGCAGCTGAG
420
GlyGlnSerGluGluMetThrHisArgThrAlaArgSerAlaProGluProAlaAlaGlu
100
AACAGAGAGATTATGAAAAGAGAAACCGGTGCGGAAGATCTAGATGAGGAAAAGAGGTTT
480
AsnArgGluIleMetLysArgGluThrGlyAlaGluAspLeuAspGluGluLysArgPhe
120
ATGAGGTTTCGGAAGAGGAGACGAAGAGGCTGAAAAGAGGTTTATGAGGTTTCGGAAGAGT
540
MetArgPheGlyArgGlyAspGluGluAlaGluLysArgPheMetArgPheGlyLysSer
140
TTCATGAGGTTTCGGAAGAGATATGAGTGACGTGGACAAAAGGTTTCATGAGGTTTCGGTAAA
600
PheMetArgPheGlyArgAspMetSerAspValAspLysArgPheMetArgPheGlyLys
160
CGATTTCATGAGGTTTCGACGAGAGCCTGGAACAGACAAGAGGTTTCATGAGGTTTGAAGA
660
ArgPheMetArgPheGlyArgGluProGlyThrAspLysArgPheMetArgPheGlyArg
180
GAACCTGGAGCCGACAAGAGGTTTCATGAGGTTTCGGGAAAAGCTTCGATGCGAGAAGAAG
720
GluProGlyAlaAspLysArgPheMetArgPheGlyLysSerPheAspGlyGluGluGlu
200
AACGACGACGATCTCTACTACAACGAGAGTGACGCGGACTCAAACGATGACGTGGACAAG
780
AsnAspAspAspLeuTyrTyrAsnGluSerAspAlaAspSerAsnAspAspValAspLys
220
CGGTTTATGCGCTTTGGGAAAAGCGCAGAGGAAAAGCGATTTATGAGATTTCGGGAAAAGC
840
ArgPheMetArgPheGlyLysSerAlaGluGluLysArgPheMetArgPheGlyLysSer
240
CAGGACGCTAGCAGAGATAAGAAAGAGTTCTTGCGCATTGGGAAGCGGAGTCAAGGTCG
900
GlnAspAlaSerArgAspLysLysGluPheLeuArgIleGlyLysArgGluSerArgSer
260
GCGGAAGTGGAACAATATCCAAATTGCTGCCAAACAGTCGTAATCAAATTAGAGCTA
960
AlaGluValGluAsnAsnIleGlnIleAlaAlaLysGlnSer***
280
GGGGGGCTTGAAGAGACAGTTGAGTAAAAAAGCCGAAATATCACCGTCTTA
1120
TAGTAATATCATATTTGGCAATCAGTATCGAAGATCGTTGTCAGTTCTAAGTCACAGCGG
1180
GAAACTCGATCTTAAATACTTTTTTCCCTTTAAATATAGCTTTTCATGTCTTAACCA
1240
ACAAATCGAGCGAAAACGAAAAATATGTTGTATCCGCACTTTATTAAATAAAAAGAGTG
1300
ATGACTTTTAGCTGTATTATTGCCCGTTTATTTGTAAATAATATATGTGATCCCAATTC
1360
1382
GGCACACAGAATAGCGGAGATC

```

Figure 2. Compilation of cDNA and genomic sequences encoding the FMRFamide precursor. Nucleotide 1 corresponds to a *Mae*III site in the genomic λ gt10 clone which was sequenced out to nucleotide 178. cDNA D1 extends from nucleotide 135 to 1382 and cDNA D3 extends from 380 to 1352 (a short sequence, 5'-GGCCA-GTATGA3' at the start of the cDNA D1 does not correspond to the genomic sequence and is not included in the compilation).

not been detected by HPLC analysis of the *Lymnaea* CNS (Eberink et al., 1986). A similar situation arose in *Aplysia* where FLRFamide remained undetected by HPLC analysis (Lehman et al., 1984) but was later found to be present on the FMRFamide precursor by sequencing of a cDNA (Taussig and Scheller, 1986).

The unique position (Fig. 6) of the 2 FLRFamide peptides in the *Lymnaea* precursor suggests that they have some biological function and are not just simple methionine-leucine substitutions. The substitution of a leucine for a methionine may confer greater stability on the peptide as leucine is less susceptible to

Figure 3. Southern transfer of genomic *Lymnaea* DNA digested with various restriction enzymes and hybridized with D3 insert. A, Undigested DNA; B, *DraI*; C, *DraI/BamHI*; D, *BamHI*; E, *EcoRI/BamHI*; F, *EcoRI*. Size markers are from a *HindIII* digest of λ DNA and the D3 cDNA insert (0.9 kb).

oxidation than methionine. This greater stability may be important in the potential hormonal actions of FLRFamide (Price, 1986).

Lymnaea has also been shown to possess 2 amino terminally extended forms of FLRFamide, SDP- and GDPFLRFamide. Extensions at the amino terminal may give the peptide even greater stability by protecting against enzymatic degradation. The fact that the N-terminal-extended FLRFamides are not present in the FMRFamide precursor protein supports the proposition of Price (1986) that after a duplication of the ancestral FMRFamide gene, one gene became specialized for the production of transmitters (FMRFamide) and the other for the production of extended FMRFamide hormones.

The 2 pentapeptides, pQFYRIamide and EFLRIamide, are previously unreported. They are bounded by basic amino acid cleavage signals and therefore likely to be processed.

The Northern blot analysis indicated that there is only one mRNA produced. Therefore, any neurons that express the FMRFamide gene are likely to synthesize the precursor, so any differential usage of the peptides probably occurs as a result of posttranslational processing.

Figure 4. Northern transfer of total *Lymnaea* CNS RNA hybridized with D3 cDNA insert. The size is estimated from the migration in adjacent tracks of yeast mRNAs *Ura4* and *hsp70*.

Comparison of the *Aplysia* CRF-like peptide with the corresponding *Lymnaea* sequences

The *Aplysia* FMRFamide precursor was reported to contain sequences that are homologous to mammalian brain peptide sequences, including corticotropin-releasing factor (CRF); α -melanocyte-stimulating hormone (α MSH), and corticotropin-like intermediate-lobe peptide (CLIP) (Taussig and Scheller, 1986).

The *Aplysia* CRF-like peptide consists of 42 amino acids situated between the FLRFamide and the FMRFamides on the precursor. Twelve of these amino acids show homology to the ovine CRF. At a similar position (63–104) on the *Lymnaea* FMRFamide precursor, there is also a sequence of 42 amino acids that is bounded on the amino and carboxy sides by Arg-Ser (62–63 and 105–106). The comparison with the *Aplysia* CRF-like peptide, ovine CRF, frog sauvagine, and carp urotensin-1 is shown in Figure 7. The *Lymnaea* and *Aplysia* genes share 61% homology at the amino acid level with a total of 25

Figure 5. Illustration of the localized expression of FMRFamide encoding mRNA to specific neurons in the CNS of *Lymnaea stagnalis* as shown by *in situ* hybridization. 35 S-labeled probe was prepared by nick translation of the gel-purified insert from the cDNA D3.

positional matches out of 42. However, the degree of homology between the *Lymnaea* sequence and any of the vertebrate peptides is even less than that of the *Aplysia* CRF-like peptide and the vertebrate peptides.

Taussig et al. (1988) concluded that if the *Aplysia* CRF-like peptide was indeed related to the vertebrate CRFs, then there would have been an amino acid substitution every 17.5 million years on the average. On this basis, the degree of divergence between *Lymnaea* and *Aplysia* sequences over the 350 million years since their divergence (Moore and Pitrat, 1960) should be 22 out of 42. The observed difference is 17, indicating that this is a good approximation to what is expected. The degree of homology throughout the 42 amino acids however is not uniform. There is a sequence (79–89) of Ser-Glu-Glu-Pro-Leu-Tyr-Arg-Lys-Arg-Arg that is 100% homologous between the 2 molluscan species and is 79% homologous at the nucleotide level. This exceptionally high level of conservation indicates that this sequence may have some functional importance. The tetra basic sequence Arg-Lys-Arg-Arg is reminiscent of the Arg-Lys-Arg-Arg sequence in the *Aplysia* ELH prohormone (Scheller et al., 1983). The processing of this prohormone has been studied by Fischer et al. (1988), who suggest that this sequence is the first to be recognized and cleaved. The tetra basic cleavage site may not only serve to generate this novel peptide but could also be important in the differential processing of the precursor by separating the FLRFamide and FMRFamides.

The *Lymnaea* precursor, unlike that of *Aplysia*, has no sequences with any significant homology to either vertebrate α MSH or CLIP. Apparent homology between the mammalian CLIP peptide and the *Aplysia* sequence residues 20–42 has been sug-

	* * * * *
Lymnaea	SEQPDVDDYPRDVVLQ- SEEPLYRKR RSTEAGGQSEEMTHRTA
Aplysia	SQEPDIEDYARAIALI EESEPLYRKR SADADGGQSEKVLHRA
Ovine	SQEPPISLDLTFHLLREVLEM-TKADQLAQANSNRKLLDIA
Frog	-QGPPISIDLSLELLRKMIEI-EKQEKEKQAANNRLLDITI
Carp	NDDPPISIDLTFHLLRNMIEM-ARNENQREQAGLNRYLDEV
Dros.	QAEQLPPEGSYAGSDELEGMAKRAAMDYGR

Figure 7. Comparison of *Lymnaea* CRF-like peptide with the analogous sequence in the *Aplysia* FMRFamide precursor, ovine CRF, frog sauvagine, carp urotensin and the CRF-like sequence in the *Drosophila* precursor. **Bold letters** show amino acids conserved between *Lymnaea* and *Aplysia*, and **asterisks** show amino acids in *Lymnaea* CRF which are in common with CRF from at least one other organism other than *Aplysia*. Sequences from Taussig and Scheller (1986) and Schneider and Taghert (1988).

gested (Taussig and Scheller, 1986). Comparison of the same region in *Lymnaea* (amino acids 15–37) shows that only amino acids serine (15), valine (17), threonine (18), and arginine (36) are conserved between *Lymnaea* and *Aplysia*, and only valine (17) is common to all 3 sequences. The 15 amino acid α MSH-like sequence in *Aplysia* (45–59) corresponds to *Lymnaea* 40–48, a 9-amino acid region containing a run of 6 gaps which are required to maximize alignment in the first 100 amino acids of the sequence. There is therefore no sequence of any significant homology of α MSH.

Does the *Lymnaea* FMRFamide represent the ancestral precursor?

Previously, the only other molluscan FMRFamide precursor to be examined at the genetic level is that of *Aplysia* (Schaefer et al., 1985; Taussig and Scheller, 1986). The most striking feature of that precursor protein is its highly repeated nature. This is especially true between the 7th FMRFamide and the 26th, where the FMRFamide and its spacer is composed of 15 or 16 amino acids. Many of these sequences are identical at the amino acid level, and several show high homology at the nucleotide level. The repeated nature of the *Aplysia* FMRFamide precursor led Price et al. (1987) to speculate on the origin of the precursor. They proposed a model to show how the modern opisthobranch gene could have arisen from a hypothetical pulmonate gene. They concluded that these highly repeated sequences could have arisen from an intragenic duplication event; presumably by the sequences encoding the 7th FMRFamide and its spacer being amplified 19 times. With the removal of this iterative portion of the *Aplysia* precursor, a model of the ancestral precursor may be generated. In *Aplysia* this would contain 9 FMRFamides, 1 FLRFamide, and the “vertebrate”-like peptides. In this respect, it should be noted that the *Lymnaea* gene contains 9 FMRFam-

Figure 6. Summary of the organization of putative peptides in the *Lymnaea* FMRFamide precursor. Symbols: ■, hydrophobic signal sequence; ▨, FLRFamide; ▩, pQFYRIamide; ▤, CRF-like peptide; ▦, FMRFamide; and ▧, EFLRIamide. The positions of lysine or arginine residues in the sequence are indicated by asterisks.

Figure 8. Comparison of *Lymnaea* and *Aplysia* sequences surrounding the novel peptides in *Lymnaea*. Alignment is by means of FASTP (Lipman and Pearson, 1985). (:), identical amino acid, (:), conservative substitution.

ides. It may therefore be the case that the *Lymnaea* FMRFamide gene bears a close relationship to the original molluscan FMRFamide gene. The *Lymnaea* FMRFamide gene, like that of *Aplysia*, also shows similarities in organization to the *Drosophila* DPKQDFMRFamide precursor protein (Schneider and Taghert, 1988).

A comparison of the *Lymnaea* and *Aplysia* amino acid sequence around 2 of the novel *Lymnaea* predicted peptides suggests how these novel peptides may have arisen (Fig. 8). The predicted peptide Gln-Phe-Tyr-Arg-Ile-NH₂ corresponds to the region of *Aplysia* encoding the α MSH-like peptide. It is possible that this sequence was present in *Lymnaea* at one time and a deletion and some substitutions gave rise to the new sequence. The other novel predicted peptide Glu-Phe-Leu-Arg-Ile-NH₂ could have arisen from a FMRFamide sequence as a result of substitutions. However, it is difficult to be certain of the alignment of one member of the FMRFamide multi-peptide family with another for the purposes of comparison. It will be interesting to determine the processing and physiological role of the novel peptides in *Lymnaea*, and therefore to determine whether such peptides have been subject to natural selection or are simply neutral changes of little or no physiological consequence.

Note added in proof: We now have evidence that the 11 nucleotides unique to D1 may be the result of RNA splicing.

References

- Abrahams, T. W., V. F. Castellucci, J. S. Camardo, E. R. Kandel, and P. E. Lloyd (1984) Two endogenous neuropeptides modulate the gill and siphon withdrawal reflex in *Aplysia* by pre-synaptic facilitation involving cAMP-dependant closure of a serotonin-sensitive potassium channel. *Proc. Natl. Acad. Sci. USA* 81: 7956-7960.
- Austin, T., S. Weiss, and K. Lukowiak (1982) FMRFamide effects a spontaneous contraction of the anterior gizzard in *Aplysia*. *Can. J. Physiol. Pharmacol.* 61: 949-952.
- Benjamin, P. R. (1984) Interneuronal network acting on snail neurosecretory neurones (Yellow cells and yellow-green cells of *Lymnaea*). *J. Exp. Biol.* 113: 165-185.
- Benjamin, P. R., and W. Winlow (1981) The distribution of three wide-acting synaptic inputs to identified neurons in the isolated brain of *Lymnaea stagnalis*. *Comp. Biochem. Physiol.* 70A: 293-307.
- Benjamin, P. R., K. J. Buckett, and M. Peters (1988) Neurones containing FMRFamide-like peptides in the model, *Lymnaea*. In *Neurobiology of Invertebrates. Transmitters, Modulators and Receptors*, J. Salanki and K.S. Rozsa, eds., pp. 415-431, Akademiai Kiado, Budapest.
- Boer, H. H., L. P. C. Schot, J. A. Vennstra, and D. Reichelt (1980) Immunocytochemical identification of neural elements in the central nervous systems of a snail, some insects, a fish, and a mammal with antiserum to the molluscan cardioexcitatory tetrapeptide FMRFamide. *Cell. Tissue Res.* 213: 21-27.
- Boyd, P. J., and R. J. Walker (1985) Actions of the molluscan neuropeptide FMRFamide on neurones in the suboesophageal ganglia of the snail *Helix aspersa*. *Comp. Biochem. Physiol.* 81: 379-386.
- Buckett, K. J. (1987) *Neuronal and Peptidergic Control of the Heart in the Snail Lymnaea stagnalis*. D. Phil. Dissertation, University of Sussex, England.
- Bullock, A. G. M., D. A. Price, A. D. Murphy, T. D. Lee, and H. N. Bowes (1988) FMRFamide peptides in *Helisoma*: Identification and physiological actions at a peripheral synapse. *J. Neurosci.* 8: 3459-3469.
- Burke, J. F., and D. Ish-Horowicz (1982) Expression of *Drosophila* heat shock genes is regulated in Rat-1 cells. *Nucleic Acids Res.* 10: 3821-3830.
- Cottrell, G. A., N. W. Davies, and K. A. Green (1984) Multiple actions of a molluscan cardioexcitatory neuropeptide and related peptides on identified *Helix* neurones. *J. Physiol. (Lond.)* 356: 315-333.
- Dirks, R. W., D. K. Raap, J. van Minnen, E. Vreugdenhil, A. B. Smit, and M. der Ploeg (1989) Detection of mRNA molecules for neuropeptide hormones of the pond snail *Lymnaea stagnalis* by radioactive and non-radioactive *in situ* hybridization: A model study for mRNA detection. *J. Histochem. Cytochem.* 37: 7-14.
- Dockray, G. J., K. R. Reeve, J. Shively, R. J. Gayton, and C. S. Barnard (1983) A novel active pentapeptide from chicken brain identified by antibodies to FMRFamide. *Nature* 305: 328-330.
- Douglass, J., O. Civelli, & E. Hervert (1984) Polypeptide gene expression: Generation of diversity of neuroendocrine peptides. *Annu. Rev. Biochem.* 53: 665-715.
- Eberink, R. H. M., D. A. Price, H. van Loenenhout, K. E. Doble, J. P. Reihm, W. P. M. Geraerts, and M. J. Greenberg (1987) The brain of *Lymnaea* contains a family of FMRFamide-like peptides. *Peptides* 8: 515-522.
- Feinberg, A. P., and B. Vogelstein (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132: 6-13.
- Fischer, J. M., W. Sossin, R. Newcombe, and R. H. Scheller (1988) Multiple neuropeptides derived from a common precursor are differentially packaged and transported. *Cell* 54: 813-822.
- Grimmelikhuijzen, C. J. P., and D. Graff (1985) Arg-Phe-amide-like peptides in the primitive nervous system of coelenterates. *Peptides* 6: 477.
- Ichinose, M., and D. J. McAdoo (1988) The voltage-dependent, slow inward current induced by the neuropeptide FMRFamide in *Aplysia* neuron R14. *J. Neurosci.* 8: 3891-3900.
- Kyte, J., and R. F. Doolittle (1982) A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157: 105-132.
- Lehman, H. K., D. A. Price, and M. J. Greenberg (1984) The FMRFamide-like peptide of *Aplysia* is FMRFamide. *Biol. Bull.* 167: 460-466.
- Lipman, D. J., and W. R. Pearson (1985) Rapid and sensitive protein similarity searches. *Science* 227: 1435-1441.
- Loh, Y. P., and H. Gainer (1983) Biosynthesis and processing of neuropeptides. In *Brain Peptides*, D. Kreiger, M. Brownstein, and J. Martin, eds., Wiley, New York.
- Maniatis, T., E. F. Fritsch, and J. Sambrook (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring, NY.
- Moore, R. C., and C. W. Pitrat (1960) *Treatise on Invertebrate Paleontology, Part I, Mollusca I*, Geological Society of America and University of Kansas Press.
- Price, D. A. (1986) Evolution of a molluscan cardio regulatory neuropeptide. *Am. Zool.* 26: 1007-1015.
- Price, D. A., and M. J. Greenberg (1977) Structure of a molluscan cardioexcitatory neuropeptide. *Science* 197: 670-671.
- Price, D. A., N. W. Davies, K. E. Doble, and M. J. Greenberg (1987) The variety and distribution of the FMRFamide-related peptides in molluscs. *Zool. Sci.* 4: 395-410.
- Richmond, J. E., A. G. M. Bullock, and K. L. Lukowiak (1984) Peptidergic modulation of a neuromuscular junction in the mollusc *Aplysia*. *Soc. Neurosci. Abstr.* 10: 690.

- Sanger, F., S. Nicklem, and A. R. Coulson (1977) DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74: 5463–5467.
- Schaefer, M., M. R. Picciotto, T. Kreiner, K-K. Kaldnay, R. Taussig, and R. H. Scheller (1985) *Aplysia* neurons express a gene encoding multiple FMRFamide neuropeptides. *Cell* 41: 457–467.
- Scheller, R. H., J. F. Jackson, L. B. McAllister, B. S. Rothman, E. Meyeri, and R. Axel (1983) A single gene encodes multiple neuropeptides mediating a stereotyped behaviour. *Cell* 32: 7–22.
- Schneider, L. E., and P. H. Taghert (1988) Isolation and characterization of a *Drosophila* gene that encodes multiple neuropeptides related to Phe-Met-Arg-Phe-NH₂ (FMRFamide). *Proc. Natl. Acad. Sci. USA* 85: 1993–1997.
- Schot, L. P. C., and H. H. Boer (1982) Immunocytochemical demonstration of peptidergic cells in the pond snail *Lymnaea stagnalis* with an antiserum to the molluscan tetrapeptide FMRFamide. *Cell Tissue Res.* 224: 347–354.
- Smit, A. B., E. Vreugdenhil, R. H. M. Ebberink, W. P. M. Geraerts, J. Klootwijk, and J. Joose (1988) Growth controlling molluscan neurons produce the precursor of an insulin-related peptide. *Nature* 331: 5335–5338.
- Stone, L. S., and E. Mayeri (1981) Multiple actions of FMRFamide on identified neurons in the abdominal ganglion of *Aplysia*. *Soc. Neurosci. Abstr.* 7: 637.
- Taussig, R., and R. H. Scheller (1986) The *Aplysia* FMRFamide gene encodes sequences related to mammalian brain peptides. *DNA* 5: 453–461.
- Taussig, R., J. R. Nambu, and R. H. Scheller (1988) Evolution of peptide hormones: An *Aplysia* CRF-like peptide. In *Neurohormones in Invertebrates*, M. C. Thorndyke and G. J. Goldsworthy, eds., Cambridge U. P., Cambridge, UK.
- Thompson, S., and P. Ruben (1988) Inward rectification in response to FMRFamide in *Aplysia* neuron L2: Summation with transient K current. *J. Neurosci.* 8: 3200–3207.
- von Heijne, G. (1985) Signal sequences: The limit of variation. *J. Mol. Biol.* 184: 99–105.
- von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 14: 4683–4690.
- Vreugdenhil, E., J. F. Jackson, T. Bouwmeester, A. B. Smit, Van Minnen, H. Van Heerikhuizen, J. Klootwijk, and J. Joose (1988) Isolation, characterization and evolutionary aspects of a cDNA clone encoding multiple neuropeptides involved in the stereotyped egg-laying behaviour of the freshwater snail *Lymnaea stagnalis*. *J. Neurosci.* 8: 4184–4191.
- Weiss, S., J. I. Goldberg, K. S. Choham, W. K. Stell, G. I. Drummond, and K. Lukowiak (1984) Evidence for FMRFamide as a neurotransmitter in the gill of *Aplysia californica*. *J. Neurosci.* 4: 1994–2000.
- Yang, H.-Y. T., W. Fratta, E. A. Majane, and E. Costa (1985) Isolation, sequencing, synthesis and pharmacological characterisation of two brain neuropeptides that modulate the action of morphine. *Proc. Natl. Acad. Sci. USA* 82: 7757–7761.