Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 1993–1997, March 1988
Neurobiology

Isolation and characterization of a *Drosophila* gene that encodes multiple neuropeptides related to Phe-Met-Arg-Phe-NH_2 (FMRFamide)

LYNNE E. SCHNEIDER AND PAUL H. TAGHERT*

Department of Anatomy and Neurobiology, Washington University School of Medicine, 660 South Euclid Avenue, Saint Louis, MO 63110

Communicated by Gerald D. Fischbach, December 10, 1987

ABSTRACT **A *Drosophila* gene that encodes neuropeptides related to molluscan Phe-Met-Arg-Phe-NH_2 (FMRFamide) was isolated by screening a genomic library with a fragment of an *Aplysia* Phe-Met-Arg-Phe-NH_2 cDNA and with synthetic oligonucleotides. This gene was used to isolate a cDNA from a *Drosophila* adult head cDNA library. The cDNA was defined by sequence analysis to encode 13 peptides that have Phe-Met-Arg-Phe-NH_2 or related sequences at their carboxyl termini. Other putative neuropeptides, including one that has homology to mammalian corticotropin-releasing factor, are present in the deduced ≈39-kDa precursor. Southern blot analysis suggested the presence of a single Phe-Met-Arg-Phe-NH_2-like gene within the haploid genome. RNA blot analysis indicated the expression of at least two transcripts of ≈1.7 and ≈0.7 kilobases. Both transcripts are evident throughout larval, pupal, and adult developmental stages. *In situ* hybridization was used to localize this neuropeptide gene to band 46C on the right arm of the 2nd chromosome. These data provide the basis for utilizing the advanced genetics and molecular techniques of *Drosophila* to address complex aspects of neuropeptide expression and function.**

The use of molecular genetic methods has greatly increased our knowledge of the structure of neuropeptide precursors (e.g., refs. 1–4), but only recently have these techniques been applied to a mechanistic analysis of neuropeptide gene regulation (5, 6). With the fruit fly *Drosophila*, there exists both a wealth of genetic information and the ability to transform animals with experimentally manipulated genes (7, 8). Further, the *Drosophila* nervous system is physiologically complex but numerically simple. Hence, it is amenable to addressing questions as diverse as the regulation of neuropeptide gene expression and the functions of neuropeptides in both embryonic and post-embryonic animals.

Phe-Met-Arg-Phe-NH_2 (FMRFamide) was originally isolated from molluscan ganglia (9) and is a member of a large family of structurally related peptides that have since been found in invertebrates and vertebrates (e.g., ref. 10). The diverse peptides exhibit a broad range of physiological actions (11–13). In insects, FMRFamide and related peptides have modulatory actions at skeletal neuromuscular junctions (14, 15), and peptides that are immunologically related to FMRFamide are released into the circulation from neurohemal organs (16). Immunocytochemical studies have revealed a small complement of ≈35 neurons in the *Drosophila* central and enteric nervous systems (17). With the long term goal being a genetic analysis of neuropeptide expression and function, we report here the use of molecular genetic techniques to identify and to sequence a gene† that encodes FMRFamide-related peptides in *Drosophila*.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

EXPERIMENTAL PROCEDURES

Materials. Restriction endonucleases, T4 polynucleotide kinase, and the Klenow fragment of DNA polymerase I were purchased from Bethesda Research Laboratories, Promega Biotec, or New England Biolabs. Hexamer primer kits for DNA labeling were purchased from Pharmacia; Sequenase kits were purchased from U.S. Biochemicals. Oligonucleotides were synthesized by the Monsanto Corp. Chemical Section. Other chemicals were of the highest purity obtainable. [γ-^{32}P]ATP (>6000 Ci/mmol; 1 Ci = 37 GBq), [α-^{32}P]-dCTP (3000 Ci/mmol), and deoxyadenosine 5′-[α-^{35}S]thiotriphosphate (dATP[^{35}S]) (1200 Ci/mmol) were obtained from Amersham or New England Nuclear.

Isolation of FMRFamide-Encoding Genomic and cDNA Clones. A *Drosophila* genomic library from the Canton S strain (18) was screened with a 1.2-kilobase (kb) fragment of the *Aplysia* cDNA for FMRFamide (19) that was nick-translated to a specific activity of ≈10^8 cpm/μg and used as a probe at 10^6 cpm/ml in a solution containing 30% (vol/vol) formamide, 4× SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7.0), 0.1% $NaDodSO_4$, salmon sperm DNA at 100 μg/ml, and 5× Denhardt's solution (1× Denhardt's solution = 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.02% bovine serum albumin) at 37°C. Filters were washed at 45°C in 2× SSC/0.1% $NaDodSO_4$, dried, and autoradiographed. A ≈400-base-pair (bp) fragment of one genomic clone (Mt-8) was hexamer-labeled to a specific activity of 10^9 cpm/μg and used to probe ≈10^6 clones of an adult head cDNA library (20). Hybridization was performed as above but without formamide and at 65°C; filters were washed at 65°C in 0.1× SSC/0.1% $NaDodSO_4$.

Nucleotide Sequence Analysis. Phage inserts or fragments thereof were subcloned into Bluescribe and M13-based vectors for sequence analysis in both orientations by using deoxyadenosine 5′-[α-^{35}S]thiotriphosphate, buffer gradient acrylamide gels, and either single-stranded or double-stranded (21) methods. DNA sequencing was performed as described by Sanger *et al.* (22) or by using a commercially available kit (Sequenase, U.S. Biochemicals).

DNA Blot Hybridization. Cloned genomic DNA was excised from Charon 4A phage arms, electrophoresed in agarose, and transferred to nitrocellulose (23). Oligonucleotide probes were end-labeled to a specific activity of 5 × 10^8 cpm/μg and used at a concentration of 3 × 10^6 cpm/ml in a solution of 6× SSC, 0.1% Sarkosyl, 10 mM EDTA, 1× Denhardt's solution, and salmon sperm DNA at 100 μg/ml. Hybridization was carried out at 30°C, and washing was at 35°C in 6× SSC/0.1% Sarkosyl. Genomic DNA was hybrid-

Abbreviations: FMRFamide, Phe-Met-Arg-Phe-NH_2; CRF, corticotropin-releasing factor.

*To whom reprint requests should be addressed.

†This sequence is being deposited in the EMBL/GenBank data base (Bolt, Beranek, and Newman Laboratories, Cambridge, MA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03232).

ized after transfer to nylon membranes (Nytran, Schleicher & Scheull) following the manufacturer's recommendations. Hexamer-labeled (24) cDNA probes were used at specific activities of 5×10^8 cpm/μg and at concentrations of 10^6 cpm/ml in a solution of 50% (vol/vol) formamide, 6× SSC, 1× Denhardt's solution, 1.4% $NaDodSO_4$, and salmon sperm DNA at 100 μg/ml at 42°C; washes were performed at 65°C in 0.1× SSC/1.4% $NaDodSO_4$.

RNA Blot Hybridization. Total RNA from different developmental stages was purified (22), poly(A)$^+$-enriched, size-fractionated in formaldehyde gels, then transferred to Nytran membranes. These blots were probed with cDNA that was hexamer-labeled to specific activities of 5×10^8 cpm/μg and used at concentrations of 10^6 cpm/ml as hybridization probes at 42°C in a solution of 50% (vol/vol) formamide, 5× SSC, 2× Denhardt's solution, salmon sperm DNA at 100 μg/ml, and 1.4% $NaDodSO_4$; blots were washed in 0.1× SSC/1.4% $NaDodSO_4$ at 65°C.

Chromosomal Localization. The *Drosophila* genomic clone Mt-8 was labeled in its entirety with biotin-UTP by hexamer-labeling, and the polytene chromosomes were prepared and probed as described (25) with only minor modifications. Hybridization signals were visualized with an alkaline phosphatase-conjugated avidin/biotin detection system (Vector Laboratories).

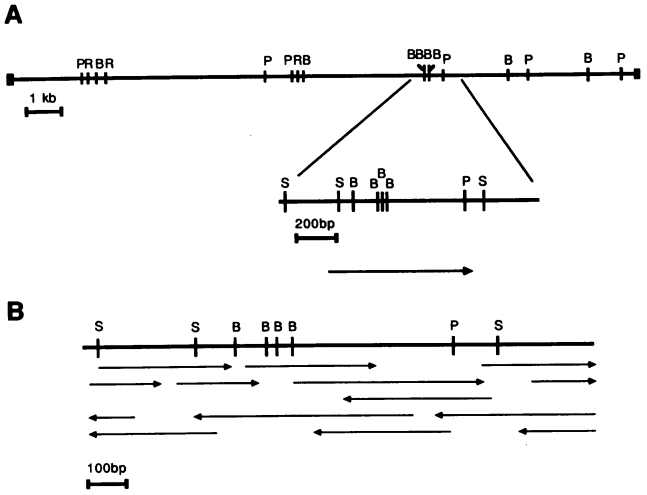

FIG. 1. Restriction map and DNA sequencing strategy for the cDNA corresponding to *Drosophila* FMRFamide mRNA. (*A*) The first line represents the genomic clone Mt-8; the second line represents the cDNA. Selected restriction sites are abbreviated as follows: R, *Eco*RI; S, *Rsa* I; B, *Bam*HI; P, *Pst* I. The arrow indicates the direction of transcription. (*B*) Sequencing strategy employed for the 1352-bp cDNA.

RESULTS

Isolation of the *Drosophila* FMRFamide-like Gene and cDNA. A 1195-bp *Taq* I fragment that includes most of the coding region of an *Aplysia* cDNA (19, 26) was used to probe a *Drosophila* genomic library at low stringency. Twelve independent clones were isolated and rescreened with a mixture of synthetic oligonucleotides whose sequence was based on the C-terminal amino acids of the *Aplysia* FMRFamide within its precursor: -Met-Arg-Phe-Gly-Lys. This rescreening was to ascertain that the DNA homology was due to correspondence in the peptide-coding regions of the probe. From the original 15 genomic clones, 5 gave strong hybridization signals in this secondary screening, and 1 clone (Mt-8) was further pursued, as it demonstrated multiple regions of hybridization within ≈2 kb of DNA. The multiple hybridization signals possibly corresponded to the multiple copies of sequences encoding FMRFamide in the *Aplysia* gene (26).

Partial DNA sequence analysis (data not shown) revealed a ≈400-bp region that could potentially encode seven short neuropeptides that terminated in the sequence Phe-Met-Arg-Phe-Gly. This region was used to screen ≈10^6 clones from a cDNA library made from adult *Drosophila* heads (20), and 2 positive clones were recovered. The clones appeared identical by size and restriction enzyme analysis. A restriction map of the genomic clone Mt-8 and of one of these cDNAs is presented in Fig. 1*A*; the strategy used to sequence the cDNA is illustrated in Fig. 1*B*.

Primary Structure of *Drosophila* FMRFamide mRNA and Precursor Protein. The nucleotide sequence of the 1352-bp cDNA is presented in Fig. 2, along with its deduced translation product. The coding strand is defined by the presence of a 29-residue poly(A)$^+$ tail at the 3′ end of one strand. The methionine residue that initiates translation is considered to be the AUG triplet at nucleotide position 19 because (*i*) it precedes the longest open reading frame present within the protein, (*ii*) it is in frame with the peptide sequences, and (*iii*) it precedes a deduced hydrophobic amino acid sequence (14/24 residues) indicative of a signal peptide. Given that this triplet represents the second AUG triplet in the cDNA, it is possible that translation may sometimes initiate at downstream AUG triplets (27). A potential signal cleavage site that has homology to a consensus sequence (28) occurs in the preproFMRFamide after amino acid residue 24 (alanine). There are 18 bp in the 5′-untranslated sequence of the cDNA; the correspondent mRNA may be 100–200 bp longer (see below). The assigned open reading frame predicts a precursor protein of 342 amino acids that has a molecular mass of ≈39 kDa. Three sequences that match the consensus ANUAAA site for poly(A)$^+$ addition are found in the 3′-untranslated region, including one that is within 28 bp of the poly(A)$^+$ tail (where, N = A, C, G, or T).

The deduced precursor (Fig. 3) contains a complex set of at least 15 distinct neuropeptides in nearly tandem array. Each of these peptide sequences is flanked by potential endoproteolytic cleavage sites (single, di-, or tri-basic residues), and each ends with a potential C-terminal amidation site (30) in a glycine residue. Ten peptides have the sequence -Phe-Met-Arg-Phe-Gly at their C termini, and these are each extended 3–5 residues toward the N terminus by one of five separate sequences. Five peptides have the extended sequence Asp-Pro-Lys-Gln-Asp-, two have Thr-Pro-Ala-Glu-Asp-, and the remaining three have Ser-Asp-Asn-, Ser-Pro-Lys-Gln-Asp-, and Pro-Asp-Asn-. Three other peptides have sequences resembling -Phe-Met-Arg-Phe-Gly: -Phe-Met-His-Phe-Gly, -Phe-Val-Arg-Ser-Gly, and -Phe-Ile-Arg-Phe-Gly. Each is preceded by its own distinct tetra- or pentapeptide sequence. The -Phe-Ile-Arg-Phe-NH_2 moiety matches the carboxyl region of a peptide purified from *Helix* (31).

The *Drosophila* FMRFamide precursor may harbor a variety of additional neuropeptides. For example, a peptide sequence (residues 113–142) between the -Phe-Met-His-Phe-Gly sequence and the first -Phe-Met-Arg-Phe-Gly sequence displays in-register homology to corticotropin-releasing factor (CRF, or related peptides) at 9 of 26 residues (32). These homologies include glutamine (*Drosophila* residue 115), Pro-Pro (residues 117 and 118) seen in ovine CRF and frog sauvagine, and leucine (residue 128) seen in ovine CRF, frog sauvagine, carp urotensin, and *Aplysia* CRF. This [and other homologies noted for the case of the *Aplysia* FMRFamide gene (26)] strengthens the argument that the invertebrate FMRFamide gene is ancestrally related to certain vertebrate neuropeptide genes. We note the presence of other potential cleavage sites at residue 57 [Leu-Leu (29)], at residue 87 (an

```
                    Met Phe Leu Leu Ala Leu Tyr Gln Met Gln Ser Ala Ile His Ser Glu Ile Ile Asp Thr Pro Asn Tyr Ala Gly Asn Ser Leu Gln   29
1   GAGATGGGCATTGCCTTG ATG TTC CTG CTG GCC CTG TAC CAG ATG CAG TCG GCC ATC CAC AGC GAG ATC ATC GAT ACG CCC AAC TAT GCG GGC AAC TCG TTG CAG

    Asp Thr Asp Ser Glu Val Ser Pro Ser Gln Asp Asn Asp Leu Val Asp Ala Leu Leu Gly Asn Asp Gln Thr Glu Arg Ala Glu Leu Glu Phe Arg His Pro   63
106 GAC ACT GAC TCC GAG GTG AGT CCA TCG CAG GAC AAT GAC CTG GTA GAT GCA CTG CTC GGC AAC GAT CAG ACC GAG AGG GCG GAG CTG GAG TTC CGG CAC CCC

    Ile Ser Val Ile Gly Ile Asp Tyr Ser Lys Asn Ala Val Val Leu His Phe Gln Lys His Gly Arg Lys Pro Arg Tyr Lys Tyr Asp Pro Glu Leu Glu Ala   97
208 ATC TCT GTG ATT GGC ATC GAC TAC TCG AAG AAC GCC GTG GTG CTG CAC TTC CAG AAA CAC GGC CGG AAA CCG CGC TAC AAG TAC GAT CCC GAG CTG GAG GCC

    Lys Arg Arg Ser Val Gln Asp Asn Phe Met His Phe Gly Lys Arg Gln Ala Glu Gln Leu Pro Pro Glu Gly Ser Tyr Ala Gly Ser Asp Glu Leu Glu Gly   131
310 AAG CGA AGG TCG GTG CAG GAC AAC TTC ATG CAC TTC GGC AAG AGG CAG GCG GAG CAG CTG CCA CCG GAG GGC AGC TAT GCT GGA TCC GAT GAA CTG GAG GGC

    Met Ala Lys Arg Ala Ala Met Asp Arg Tyr Gly Arg Asp Pro Lys Gln Asp Phe Met Arg Phe Gly Arg Asp Pro Lys Gln Asp Phe Met Arg Phe Gly Arg   165
412 ATG GCC AAG CGA GCA GCT ATG GAT CGG TAT GGC AGA GAT CCC AAG CAG GAC TTC ATG CGG TTT GGT CGG GAT CCG AAA CAG GAC TTC ATG AGG TTT GGC AGG

    Asp Pro Lys Gln Asp Phe Met Arg Phe Gly Arg Asp Pro Lys Gln Asp Phe Met Arg Phe Gly Arg Asp Pro Lys Gln Asp Phe Met Arg Phe Gly Arg Thr   199
514 GAT CCA AAG CAG GAC TTC ATG AGA TTC GGT CGG GAT CCC AAG CAG GAT TTC ATG AGA TTC GGT CGA GAT CCC AAG CAG GAT TTC ATG AGG TTT GGA CGC ACT

    Pro Ala Glu Asp Phe Met Arg Phe Gly Arg Thr Pro Ala Glu Asp Phe Met Arg Phe Gly Arg Ser Asp Asn Phe Met Arg Phe Gly Arg Ser Pro His Glu   233
616 CCG GCT GAG GAT TTC ATG AGG TTC GGA CGC ACT CCG GCG GAG GAC TTC ATG AGG TTC GGA CGC TCC GAC AAT TTC ATG CGC TTC GGA CGC AGT CCC CAC GAG

    Glu Leu Arg Ser Pro Lys Gln Asp Phe Met Arg Phe Gly Arg Pro Asp Asn Phe Met Arg Phe Gly Arg Ser Ala Pro Gln Asp Phe Val Arg Ser Gly Lys   267
718 GAG CTT CGC AGT CCC AAA CAG GAT TTC ATG CGA TTC GGT CGC CCG GAC AAC TTC ATG CGC TTC GGG CGT TCC GCT CCG CAG GAT TTT GTG CGC TCC GGG AAG

    Met Asp Ser Asn Phe Ile Arg Phe Gly Lys Ser Leu Lys Pro Ala Ala Pro Glu Ser Lys Pro Val Lys Ser Asn Gln Gly Asn Pro Gly Glu Arg Ser Pro   301
820 ATG GAC TCA AAC TTC ATT CGA TTC GGT AAG AGC TTA AAG CCG GCG GCT CCC GAG TCC AAG CCA GTC AAG TCC AAT CAA GGC AAC CCA GGC GAA CGC AGT CCA

    Val Asp Lys Ala Met Thr Glu Leu Phe Lys Lys Gln Glu Leu Gln Asp Gln Gln Val Lys Asn Gly Ala Gln Ala Thr Thr Thr Gln Asp Gly Ser Val Glu   335
922 GTG GAC AAG GCC ATG ACG GAG CTG TTC AAG AAA CAG GAG CTG CAG GAT CAG CAG GTG AAG AAC GGC GCA CAG GCG ACC ACC ACG CAG GAT GGG AGT GTG GAA

    Gln Asp Gln Phe Phe Gly Gln STOP
1024 CAG GAC CAG TTC TTC GGC CAG TGA GGTAGTCCTGCGGGACGCCTCCTTGTAAATAGATATGGACAAATGTACGCAAGGATCTAAATTGATATACGTATATAACCCACTCCTCACACGAACTCCTGAC

1168 TTATGCCTGAACTATGAATTTTTAATGAATGGGCTGGATTAAAAATTCACCGTGCTTTGAAGTTCTTATCTATAAATATATCTAGTGTAATATTGAAGAAATTGAAATTGGCGTGAATAAAATCCTGTGGCAACA

1275 TTTTAAATAAAGATTGCTTTACTGTAAATTATGGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIG. 2. Nucleotide sequence of the *Drosophila* FMRFamide cDNA (1352 bp, numbered at the left) and the deduced amino acid sequence of the polyprotein precursor (342 amino acids, numbered at the right). Three polyadenylylation consensus sequences (underlined) and a poly(A)$^+$ tail of 29 bases are present at the 3′ end. The initiator ATG precedes a predicted ≈39-kDa precursor protein that includes a number of Phe-Met-Arg-Phe-Gly sequences (overbar) or related sequences; single or dibasic residues flank each of these putative neuropeptides.

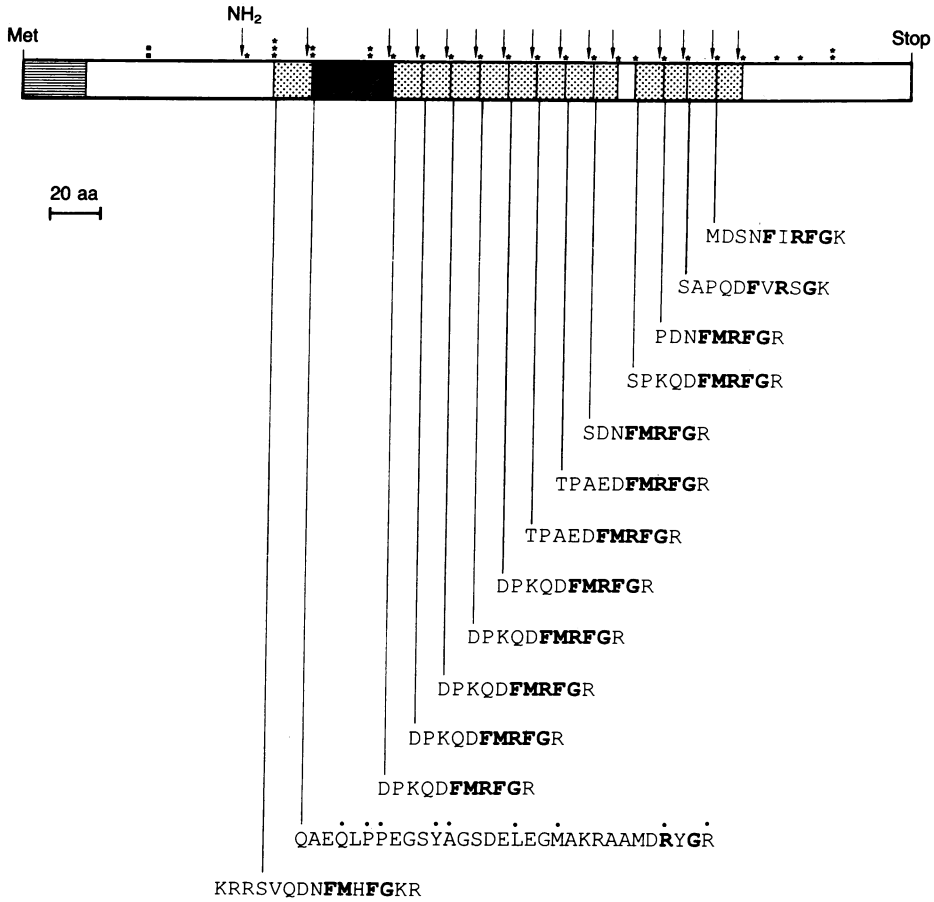

FIG. 3. Schematic representation of the FMRFamide polyprotein precursor in *Drosophila*. The protein has a 25-residue hydrophobic leader sequence (horizontal lines). Deduced peptide cleavage products are represented as boxes within the precursor. Phe-Met-Arg-Phe-Gly-like peptides (stippled) and a putative CRF-like peptide (hatched) (c.f., ref. 26) are included. Arrows indicate positions of glycine residues that precede basic residues (stars indicate number of basic residues). Glycine residues donate the amide moiety for peptide amidation (NH_2), and basic residues may be used as sites for endoproteolytic cleavage. The order of the deduced peptides are listed below the precursor with "FMRFG" or similar sequences listed in boldface letters. Solid squares above residues 47 and 48 denote a Leu-Leu that may also serve as a cleavage site (29). Dots above residues in the longest sequence (second from bottom) indicate those residues that are homologous to correspondent positions within CRF or related peptides. The single-letter amino acid code is used.

arginine, preceded by a glycine), at residue 290 (a lysine followed by serine, many of the FMRFamide peptides upstream begin with a serine residue), at residue 299 (arginine, again followed by serine), and at residue 311 (Lys-Lys).

Hybridization Analysis of the *Drosophila* FMRFamide Gene and mRNA. Southern blot analysis indicates that the *Drosophila* FMRFamide gene is present as a single copy in the haploid genome (Fig. 4*A*). We cannot exclude the possibility that other, unrelated *Drosophila* genes may exist that contain small regions encoding peptides with sequences related to FMRFamide. For example, we isolated a distinct and unrelated genomic clone (Mt-10) in the original screen of the *Drosophila* genome by homology to the *Aplysia* cDNA. Like clone Mt-8, it displays numerous regions of hybridization to the synthetic oligonucleotide mixture that is based on the *Aplysia* sequence -Met-Arg-Phe-Gly-Lys (26). In addition, a gut peptide with the C-terminal sequence -Met-Arg-Phe-NH_2 was purified from cockroach (33); if such a peptide exists in *Drosophila* it would likely be recognized by anti-FMRFamide antiserum, but it does not appear to be encoded by the gene herein described.

By using the entire 1352-bp *Drosophila* cDNA as a probe, RNA gel blot analysis of poly(A)$^+$ RNA produced a relatively constant transcriptional pattern throughout several developmental stages (Fig. 4*B*). The following two major bands are evident: an mRNA of ≈1.7 kb that presumably represents the class of transcript giving rise to the cDNA we analyzed and a second smaller, but more abundant, band(s) at ≈0.7 kb. By comparison with the ≈1.7-kb transcript, the 1352-bp cDNA clone we have analyzed is at most 350 bp less than full length. By comparison with genomic sequence (data not shown), amino acid residues 1–150 appear to be encoded by a single exon; we do not yet know if introns interrupt any of the remaining portions of the cDNA we have analyzed.

Chromosomal Localization of the FMRFamide Gene. By using *in situ* hybridization to salivary gland polytene chromosomes, we have localized the *Drosophila* FMRFamide gene to a position at or near band 46C on the right arm of the second chromosome (Fig. 5). This region of the genome may be covered by the deficiency Df(2R)*eve*$^{1.27}$ that affects the nearby gene even-skipped (34).

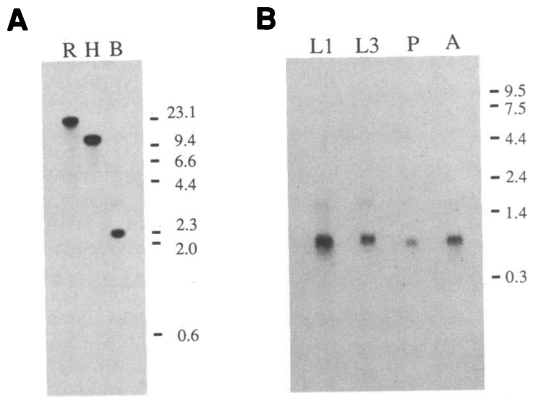

FIG. 4. Hybridization analysis of the *Drosophila* FMRFamide gene and mRNAs. (*A*) Genomic DNA (20 μg per lane) was digested with various enzymes and probed with the entire 1352-bp *Drosophila* FMRFamide cDNA that had been radiolabeled. Lanes: R, *Eco*RI; H, *Hin*dIII; B, *Bam*HI. Molecular size standards (× 10^{-3} base pairs) (restriction digest of phage λ DNA) are indicated to the right. (*B*) Poly(A)$^+$ RNA (20 μg per lane, except lane P with 15 μg) from four developmental stages was probed with the entire *Drosophila* FMRFamide cDNA that had been radiolabeled. Lanes: L1, first larval instar; L3, third larval instar; P, pupal stage; A, adult. Molecular size standards (× 10^{-3} base pairs) (Bethesda Research Laboratories) are indicated to the right.

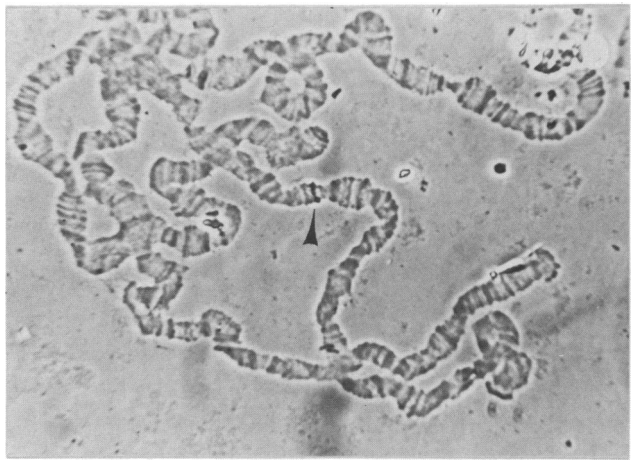

FIG. 5. Localization of the FMRFamide gene to a locus at or near band 46C (arrowhead) of salivary gland polytene chromosomes.

DISCUSSION

Whereas insects have proven to be excellent systems for the analysis of neuropeptide expression and function (e.g., ref. 35), little information exists concerning their biosynthesis or the genes that encode them. These experiments describe the isolation in an insect of a neuropeptide-encoding gene. More significantly, isolation of this gene in *Drosophila* provides the basis for a genetic analysis of neuropeptide regulation and function throughout development. Such an analysis can make use of the molecular techniques that are available to transform these animals and the wealth of genetic information and resources that are at hand to manipulate gene dosage.

Nambu *et al.* (36) used a serum antibody to synthetic FMRFamide to isolate an endogenous *Drosophila* peptide with an amino acid sequence that matches that of the most prevalent peptide predicted from the deduced precursor: Asp-Pro-Lys-Gln-Asp-Phe-Met-Arg-Phe-NH_2. This is consistent with the hypothesis that the mRNA we have described is translated and that the precursor is posttranslationally cleaved and amidated as suggested by its structural analogies to other neuropeptide precursors. Whether other portions of the precursor are similarly processed to their suspected final forms could be determined by the use of antisera specific to processed forms of those other putative peptides.

The *Drosophila* FMRFamide gene, like the *Aplysia* gene with which it was detected, encodes multiple copies of related neuropeptides within a single precursor protein. Inspection of the sequence reveals the inclusion of 13 diverse FMRFamide-like peptides in the *Drosophila* precursor that differ from each other mainly in the N terminus. This pattern of variable structure suggests either of the two following inferences. (*i*) The N-terminal region of the molecules is not physiologically significant for the animal and hence has drifted during evolution, or (*ii*) this variable region may confer differing stabilities or biological actions onto the various peptides and hence represents the structural basis for complex physiological actions. Direct examination of the activities of synthetic peptides based on these deduced sequences will help resolve these issues.

Immunocytochemistry with an antiserum to FMRFamide indicates the presence of immunoreactive material in ≈35 neurons in the central and enteric nervous systems (17). It is likely that the gene described in this report is responsible for most and perhaps all of these immunoreactive forms. However, other unrelated genes may exist that encode small

peptides having limited homology to FMRFamide. For example, in *Aplysia*, a second neuropeptide gene that is expressed in the neuron L5 encodes a peptide that ends with the sequence -Arg-Phe-NH_2 (37). Additionally, an insect neuropeptide with a -Met-Arg-Phe-NH_2 C terminus has been purified (33). These indications of diverse -Arg-Phe-NH_2 peptides and genes underscore the importance of using specific antisera and nucleic acid probes to determine the exact spatial pattern of expression of the *Drosophila* FMRFamide gene.

RNA gel blot analysis of FMRFamide mRNA revealed the expression of at least two size classes; the larger class probably corresponds to the cDNA we have analyzed. We have not yet isolated a cDNA corresponding to the smaller class; either (*i*) it is present in the head at much lower abundance than the larger class, or (*ii*) it is homologous to the larger class only in untranslated regions of the mRNA. While more rigorous analysis of the transcriptional activity of this gene is needed, these data indicate its potential diversity and, therefore, potential cell-specific regulation in the biosynthesis or posttranslational processing of the putative polyproteins.

The identification of the gene locus at band 46C provides an entry point to begin manipulating structure and dosage *in vivo*. This region of the 2nd chromosome is not well characterized but does contain certain deficiencies and translocations that should allow the establishment of mutant *Drosophila* stocks that lack FMRFamide-related neuropeptide expression. These stocks will permit direct *in vivo* examination of the physiological and developmental consequences of the absence of this set of biologically active molecules.

Against this mutant background, structurally altered FMRFamide genes can be stably integrated into the germ line so as to precisely test hypotheses concerning the functions, copy number, and linear order of multiple neuropeptides within a single polyprotein precursor. Similar methods would also be useful to examine the rules by which precise cell-specific expression of a neuropeptide gene is elaborated during development.

We are grateful to Richard Scheller and colleagues for sharing the *Aplysia* FMRFamide cDNA and to Paul Salvaterra and Gerry Rubin for sharing their *Drosophila* cDNA libraries. We thank Kathy Naylor, Philip Copenhaver, Jim Krause, Karen O'Malley, Aguan Wei, and Larry Salkoff for assistance and advice. We thank Ian Duncan for help in interpreting polytene chromosome bands; Sally Elgin for providing *Drosophila*; and Karen O'Malley, Larry Salkoff, David Harris, and John Merlie for a critical reading of the manuscript. L.E.S. was supported by a National Institutes of Health training grant; the work was supported by grants from the Monsanto Corporation, the National Institutes of Health (NS 21749), and the Washington University Center for Cellular and Molecular Neurobiology to P.H.T., who is a Sloan Fellow.

1. Kakidini, H., Furutani, Y., Takahishi, H., Noda, M., Morimoto, M., Hirose, T., Asai, M., Inayama, S., Nakanishi, S. & Numa, S. (1982) *Nature (London)* **298**, 245–249.
2. Land, H., Schutz, G., Schmak, H. & Richter, D. (1982) *Nature (London)* **295**, 299–303.
3. Amara, S. G., Jonas, V., Rosenfeld, M. G., Ong, E. S. & Evans, R. M. (1982) *Nature (London)* **298**, 240–244.
4. Scheller, R. A., Jackson, J. F., McAllister, L. B., Rothman, B. S., Mayeri, E. & Axel, R. (1983) *Cell* **32**, 7–22.
5. Nelson, C., Crenshaw, E. B., Franco, R., Lira, S. A., Albert, V. A., Evans, R. M. & Rosenfeld, M. G. (1986) *Nature (London)* **322**, 557–562.
6. Crenshaw, E. B., Russo, A. F., Swanson, L. W. & Rosenfeld, M. G. (1987) *Cell* **49**, 389–398.
7. Spradling, A. C. & Rubin, G. M. (1982) *Science* **218**, 341–347.
8. Scholnick, S. B., Ray, S. J., Morgan, B. A., McCormick, C. A. & Hirsh, J. (1986) *Science* **234**, 998–1002.
9. Price, D. A. & Greenberg, M. J. (1977) *Science* **197**, 670–671.
10. Dockray, G. J. C., Reeve, J. R., Jr., Shively, J., Gayton, R. J. & Barnard, C. S. (1983) *Nature (London)* **305**, 328–330.
11. Cottrell, G. A. (1982) *Nature (London)* **296**, 87–89.
12. Gayton, R. J. (1982) *Nature (London)* **298**, 275–276.
13. Doble, K. E. & Greenberg, M. J. (1982) *Neuropeptides* **2**, 157–167.
14. Walther, C. & Scheibe, M. (1987) *Neurosci. Lett.* **77**, 209–214.
15. Evans, P. & Myers, C. M. (1986) *J. Exp. Biol.* **126**, 403–422.
16. Carroll, L. S., Carrow, G. M. & Calabrese, R. L. (1986) *J. Exp. Biol.* **126**, 1–14.
17. White, K., Hurteau, T. & Punsal, P. (1986) *J. Comp. Neurol.* **247**, 430–438.
18. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, D. K. & Efstratiadis, A. (1978) *Cell* **15**, 687–701.
19. Schaeffer, M., Piccioto, M. R., Kreiner, T., Kaldany, R. R., Taussig, R. & Scheller, R. H. (1985) *Cell* **41**, 457–467.
20. Itoh, N., Slemmon, J. R., Hawke, D. H., Williamson, R., Morita, E., Ikatura, K., Roberts, E., Shively, J. E., Crawford, G. D. & Salvaterra, P. M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4081–4085.
21. Chen, E. Y. & Seeburg, P. H. (1985) *DNA* **4**, 165–170.
22. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
23. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
24. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
25. Langer-Sofer, P. R., Levine, M. & Ward, D. C. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4381–4385.
26. Taussig, R. & Scheller, R. H. (1986) *DNA* **5**, 453–461.
27. Kozak, M. (1984) *Nucleic Acids Res.* **12**, 3873–3893.
28. von Heinje, G. (1984) *J. Mol. Biol.* **173**, 243–251.
29. Smyth, D. G. & Massey, D. E. (1979) *Biochem. Biophys. Res. Commun.* **87**, 1006–1010.
30. Bradbury, A. F., Finnie, M. B. A. & Smyth, B. G. (1982) *Nature (London)* **298**, 686–688.
31. Price, D. A. (1982) *Comp. Biochem. Physiol. C* **72**, 325–328.
32. Vale, W. W., Rivier, C., Spiess, J. & Rivier, J. (1983) in *Brain Peptides*, eds. Kreiger, D., Brownstein, M. J. & Martin, J. B. (Wiley, New York), pp. 961–974.
33. Nachman, R. J., Holman, G. M., Haddon, W. F. & Ling, W. (1986) *Science* **234**, 71–73.
34. Nüsslein-Volhard, C., Weischaus, E. & Kluding, H. (1984) *Roux's Archiv.* **193**, 267–282.
35. Truman, J. & Taghert, P. H. (1983) in *Brain Peptides*, eds. Krieger, D., Brownstein, M. J. & Martin, J. B. (Wiley, New York), pp. 67–73.
36. Nambu, J., Andrews, P. C., Gottfried, J. & Scheller, R. H. (1987) *Soc. Neurosci. Abstr.* **13**, 1256.
37. Shyamala, M., Fisher, J. M. & Scheller, R. H. (1986) *DNA* **5**, 203–208.