MOLECULAR AND CELLULAR BIOLOGY, Aug. 1987, p. 2762–2771
0270-7306/87/082762-10$02.00/0

Vol. 7, No. 8

Aplysia californica Neurons Express Microinjected Neuropeptide Genes

LUC DESGROSEILLERS,[1] DAVID COWAN,[1] MICHAEL MILES,[2] ALEX SWEET,[1] AND RICHARD H. SCHELLER*

Departments of Biological Sciences,[1] and Neurology,[2] Stanford University, Stanford, California 94305

Received 18 February 1987/Accepted 20 April 1987

Neuropeptide genes are expressed in specific subsets of large polyploid neurons in *Aplysia californica*. We have defined the transcription initiation sites of three of these neuropeptide genes (the R14, L11, and ELH genes) and determined the nucleotide sequence of the promoter regions. The genes contain the usual eucaryotic promoter signals as well as other structures of potential regulatory impoortance, including inverted and direct repeats. The L11 and ELH genes, which are otherwise unrelated, have homology in the promoter regions, while the R14 promoter was distinct. When cloned plasmids were microinjected into *Aplysia* neurons in organ culture, transitions between supercoiled, relaxed circular, and linear DNAs occurred along with ligation into high-molecular-weight species. About 20% of the microinjected neurons expressed the genes. The promoter region of the R14 gene functioned in expression of the microinjected DNA in all cells studied. When both additional 5′ and 3′ sequences were included, the gene was specifically expressed only in R14, suggesting that the specificity of expression is generated by a multicomponent repression system. Finally, the R14 peptide could be expressed in L11, demonstrating that it is possible to alter the transmitter phenotype of these neurons by introduction of cloned genes.

Control of gene expression in the central nervous system (CNS) may require novel and complex regulatory networks to generate the vast and intricate communication network between cells. The specificity of this network is determined by the expression of specific sets of genes, including those which determine neurotransmitter and receptor phenotype. In addition, the CNS can be considered to develop continuously throughout the life of an animal as experience shapes the fine structure of this communication network. The complexity, which affords the nervous system with remarkable properties, makes cellular and molecular studies, including those of gene expression, a difficult task. Neurons are generally postmitotic, and the vast number of unique cells makes the study of identified neurons with known functions very difficult in vertebrates.

Many invertebrate species, including mollusks, have provided partial solutions to these problems due to the numerical simplicity of their CNSs and the large size of the component neurons. The CNS of the gastropod mollusk *Aplysia californica* contains about 20,000 neurons, many of which are reproducibly identifiable, allowing physiological studies of the roles of these cells in governing simple behaviors. The 2,000 neurons of the *Aplysia* abdominal ganglion have been the subject of numerous investigations concerning the neuronal elements underlying simple forms of learning, while other studies have focused on the cellular mechanisms governing fixed action patterns (27, 28).

The neuronal communication network mediating animal behaviors is linked by a large variety of secreted substances, such as acetylcholine and glutamine, which act over short distances and short time periods within the constrained space of the synapse. Neuropeptides can function in a similar fashion, but in addition often serve to modulate the actions of classic transmitters. Neuropeptides can travel some distance from their site of release, and the duration of their actions may be quite prolonged (4). Many, perhaps even most, *Aplysia* neurons use neuropeptides as chemical messengers, and in some cells up to 10 to 25% of the mRNA encodes precursors for the biologically active peptide products (for a review, see reference 26). RNA blotting, in situ hybridization, and immunohistochemistry techniques all demonstrate that the various neuropeptides are expressed in cell-specific patterns ranging from as few as 12 cells in the case of the R14 peptide to as many as 100 or so cells in the case of the L11 peptide (Fig. 1) (for example, see references 30, 39, and 47).

Thus, in *A. californica*, it is possible to isolate genes encoding neuropeptide precursors by differential screening of cDNA and genomic libraries with probes made from the mRNA of individually identified neurons (for a review, see reference 46). Analysis of the cDNA clones allows definition of the primary sequence of the peptide precursor; therefore, precise studies of the proteolytic processing and packaging are possible. As in all animals from yeasts to humans, the *Aplysia* precursors are proteolytically cleaved at basic residues, and the resulting fragments are found localized in the same secretory granules (31, 39a). Physiological studies of these neuropeptide products are elucidating their roles in governing simple behaviors, as well as the cellular mechanisms whereby the responses are elicited. Thus, we are able to follow quite precisely the flow of information encoded in these neuropeptide genes through biosynthesis to a physiological response and resulting behavior.

While this system provides an excellent model for studies of gene expression in the nervous system, little is known about the molecular mechanisms that regulate cell-specific expression of the peptide genes. Even less is understood about the mechanisms whereby neuronal activity may regulate expression of the genes in different physiological states. To begin approaching these issues, we determined the transcription start sites and characterized the promoter regions of a peptide gene expressed in neurons R3 through R14 (referred to as the R14 gene) and a second gene expressed in L11, as well as the left-upper-quadrant (LUQ) cells (referred to as the L11 gene). We also studied the gene family expressed in the bag cells and the atrial gland, which govern

* Corresponding author.

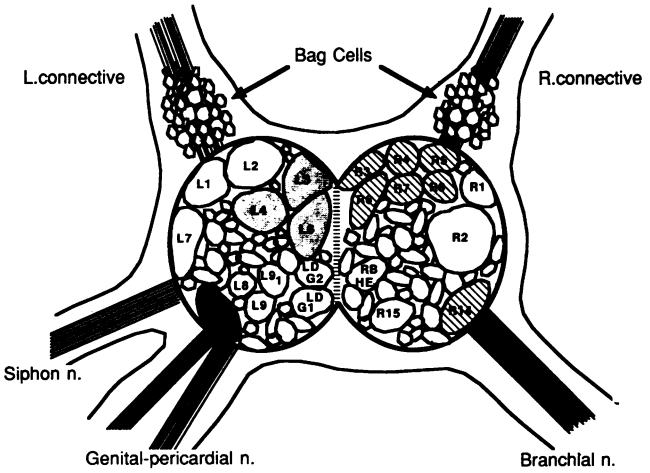

FIG. 1. Schematic representation of the *Aplysia* abdominal ganglion dorsal surface. Some well-identified neurons of the abdominal ganglion are indicated by their corresponding number in the right (R) or left (L) hemiganglion. Three sets of neurons, each of which expresses a unique neuropeptide gene, are indicated. The bag cells are localized on the rostral connective and govern egg-laying behavior via a set of neuropeptides processed from the ELH precursor. The neurons R3 through R14 (hatched areas) express a neuropeptide gene called the R14 gene after the cell used in its initial isolation. The L11 (dark stippling) neuron also expresses large amounts of a neuropeptide gene, and the LUQ cells L3, L4, and L6 (light stippling) express this gene at about a 10-fold lower level. The ganglion is about 2 to 4 mm in diameter and was dissected from the animal, pinned in a dish, and microinjected in this position.

egg-laying behavior (referred to as the ELH gene) (Fig. 1). Many *Aplysia* abdominal ganglion neurons are highly polyploid and quite large (up to 500 μm diameter), making it possible to reintroduce cloned genes into the cells by microinjection. Using this procedure, we have been able to study gene expression in identified neurons and thus have a model system for investigating the function of products introduced into nerve cells at the DNA level.

MATERIALS AND METHODS

Animals. *A. californica* weighing 75 to 150 g were purchased from Sea Life Supply (Sand City, Calif.) and maintained in an aerated recirculating seawater tank at 15°C for less than 2 weeks before use.

Microinjection. Abdominal ganglia from animals weighing 75 to 150 g were dissected and pinned in a dish containing artificial sea water (490 mM NaCl, 11 mM KCl, 19 mM $MgCl_2$, 30 mM $MgSO_4$, 11 mM $CaCl_2$, and 10 mM Tris, pH 7.6). Individual cells of the dorsal face were injected with a DNA solution suspended in 1 M KCl and 0.01% fast green. About 10^6 to 10^7 copies of the gene were injected in a volume of about 0.1 to 1 nl. The membrane potential of the cells was recorded during all procedures. After injection, the ganglion was kept for 48 h in L15 tissue culture medium at 16°C. The medium was then changed to cold propylene glycol-sea water (40:60), and individual cells were dissected.

RNA extraction. Dissected cells were lysed in 4 M guanidium thiocyanate–2% Sarkosyl–12.5 mM EDTA–50 mM Tris (pH 7.5)–0.14 M 2-mercaptoethanol at 60°C. The lysate was diluted with an equal volume of 0.1 M sodium acetate, pH 5.0 and extracted twice with phenol-chloroform at 60°C and once with chloroform. The aqueous phase was ethanol precipitated in the presence of 10 μg of yeast RNA as carrier. The RNA was then suspended in 10 mM Tris (pH 7.9)–10 mM NaCl–6 mM $MgCl_2$ and incubated with RQ1 DNase (Promega Biotech) (1 U/μg of DNA) for 15 min at 37°C. The RNA was extracted successively by phenol, phenol-chloroform, and chloroform and ethanol precipitated.

DNA extraction. Single dissected cells were lysed in 10 mM Tris (pH 7.5)–1 mM EDTA–200 mM NaCl–0.6% sodium dodecyl sulfate, extracted successively with phenol, phenol-chloroform, and chloroform, and ethanol precipitated with 10 μg of yeast RNA as carrier.

DNA sequencing. DNA fragments were sequenced by the method of Maxam and Gilbert (37) essentially as described before (39).

Analysis of DNA and RNA. Southern and Northern (RNA blot) electrophoresis techniques and nucleic acid transfer to nitrocellulose were done as described before (39). For the dot blots, DNase-treated RNA was blotted to nitrocellulose and hybridized to ^{32}P-labeled probes (1). Filters were washed at 85°C for 2 h in 0.05 M SSC–0.05% sodium dodecyl sulfate before rehybridization.

Construction of plasmids. The techniques used in the construction of hybrid molecules (restriction endonuclease digestion, isolation of DNA fragments from agarose or acrylamide electrophoresis gels, ligation, subcloning, colonies, screening, etc.) have all been described (36). The genomic R14 clone was first subcloned into pBR322 at the *Hin*dIII site. The pRL-2A plasmid was constructed in two steps. First, a 192-base-pair (bp) *Hae*III partial digest fragment (nucleotides 376 to 568 [48]) isolated from L11 cDNA was isolated and ligated at the *Alu*I site (nucleotide 277 [39]) of the subclone pMP65 (49), which contains the second exon of the R14 gene. Second, one clone which contained the L11 insert in the proper orientation was purified, and the *Eco*RI-*Sst*I fragment was substituted for the corresponding fragment in the R14 genomic clone.

pRL-2B cloned DNA was derived from the pRL-2A plasmid. pRL-2A DNA was digested to completion with *Bam*HI and recircularized. This eliminated all the 5'-flanking region except for 420 nucleotides. To construct pR14CAT-E1 and pR14-CAT plasmids, we first isolated from pR14E1 (a subclone that contains the *Bam*HI-*Sal*I fragment of R14 genomic clone; unpublished data) the *Bam*HI-*Fok*I and *Bam*HI-*Taq*I fragments, which contain the R14 promoter with and without the first exon, respectively. To do so, we first digested pR14E1 to completion with *Taq*I or *Fok*I endonuclease, filled with protruding ends with the Klenow fragment of DNA polymerase I, and then digested to completion with *Bam*HI nuclease. Both fragments were then ligated, along with the *Hin*dIII-*Bam*HI fragment of the pR14 clone, into the *Hin*dIII-*Bam*HI-digested pCAT-B' (an enhancerless and promoterless chloramphenicol acetyltransferase [CAT] plasmid) in which the *Bam*HI site had been filled in.

pR14CAT-EB and pR14CAT-B were derived from pR14CAT-E1 and pR14CAT, respectively, by digestion with *Hin*dIII and *Bam*HI, filling the sticky ends with the Klenow enzyme, and recircularization.

pLR-1 was constructed by ligation of the *Hin*dIII-*Sal*I fragments isolated from the 5' end of the L11 genomic clone and 3' end of the R14 genomic clone, respectively.

S1 mapping and primer extension. The 5' ends of the transcripts were analyzed with S1 nuclease as described (50). Hybridiztion was performed at 50°C in 80% formamide. The hybrids were digested with 20 or 50 U of S1 nuclease at 37°C for 1 h, and the protected fragments were analyzed by

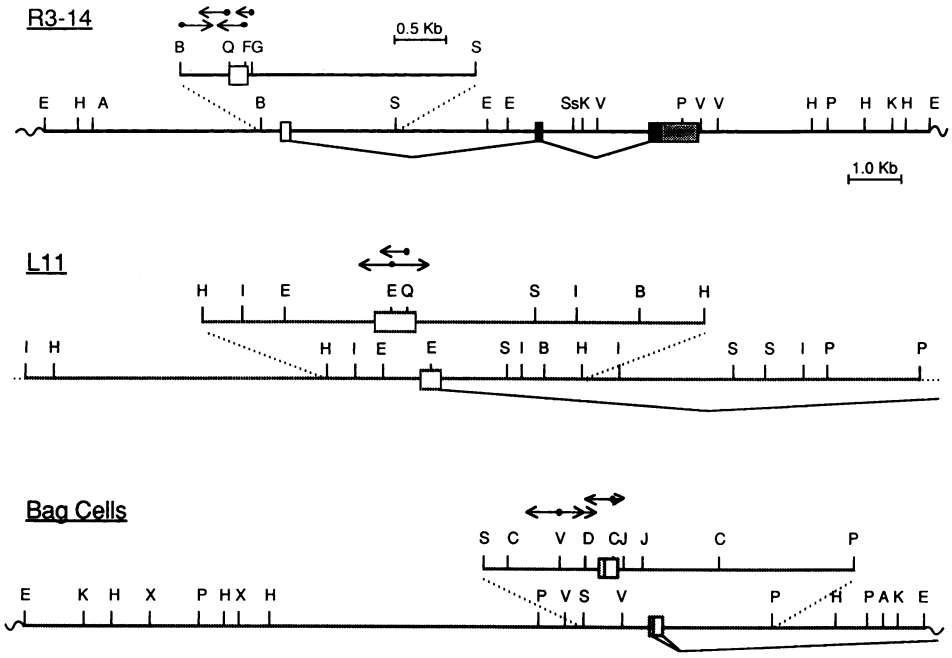

FIG. 2. R14, L11, and bag cell genomic clones containing promoter regions. The positions of restriction enzyme sites are shown along with the 5′ untranslated (open boxes), coding (solid boxes), and 3′ untranslated (stippled boxes) portions of the mRNAs. The R14 peptide clone contains the entire gene; the L11 and bag cell (or ELH gene) clones contain only the first exons. The lines below the clones indicate the splicing pattern. The ELH gene has an alternative splicing pattern in the 5′ untranslated region, as indicated. The arrows indicate the strategy used to sequence the promoter regions as shown in Fig. 4. Restriction enzymes: A, *Sma*I; B, *Bam*HI; C, *Hin*cII; D, *Hae*III; E, *Eco*RI; F, *Fok*I; G, *Bgl*II; H, *Hin*dIII; I, *Xba*I; J, *Hha*I; K, *Kpn*I; P, *Pst*I; Q, *Taq*I; S, *Sal*I; Ss, *Sst*I; V, *Pvu*II; X, *Xho*I.

electrophoresis through 6% polyacrylamide–urea sequencing gels.

The primer extension experiments were done as described (38). The extended fragments were analyzed by electrophoresis through 6% polyacrylamide–urea sequencing gels.

RNase mapping. The RNase mapping was done as described previously (51). The RNA isolated from single cells was dissolved in 80% formamide, hybridized overnight at 47°C to ^{32}P-labeled antisense riboprobe, digested with RNases A and T_1 at 35°C for 30 min, and migrated on a 5% polyacrylamide nondenaturing gel. Synthesis of the SP6 probes was done as described (51).

CAT assay. Approximately 48 h after injection, single cells were dissected, suspended in water, and frozen and thawed twice. The CAT assay and separation of the reaction products were performed as described (21).

Immunohistochemistry. The R3-14 peptide antibodies have been described previously (30). The techniques used for immunohistochemistry were essentially as described (30).

RESULTS

***Aplysia* neuropeptide gene promoters.** Genomic and cDNA clones encoding neuropeptide genes expressed in R3-14, L11, and the bag cell neurons had previously been isolated and characterized (39, 45, 47, 48). The transcribed regions spanned a minimum of 7 kilobases (kb) in the case of the R14 peptide gene and over 30 kb in the case of the L11 peptide gene. The cDNA clones contained substantial 5′ untranslated sequences and were approximately the same size as the mRNA determined by RNA blotting techniques. As a first step in characterizing the promoters, we generated more detailed restriction enzyme maps in the region of the 5′ end of the mRNA (Fig. 2).

This structural information was used in conjunction with primer extension and S1 mapping experiments to precisely define the sites of transcription initiation of the three genes (Fig. 3 and 4). Extension of bag cell or atrial gland RNA with a 98-nucleotide fragment from the 5′ end of an ELH cDNA revealed two major bands of approximately equal intensity in both tissues, indicating that the same relative start sites are used by the members of this small gene family (Fig. 3A). Thus, while the nucleotide sequences of the promoter and adjacent regions are similar (35), these genes are very abundantly expressed in distinct tisues. This suggests that elements controlling tissue-specific gene expression are located outside the promoter region itself.

Whole abdominal ganglion RNA was used with both S1 mapping and primer extension techniques to map a single transcription initiation site of the R14 gene (Fig. 3B).The extension of a 58-bp probe located at the 5′ end of the R14 cDNA clone gave rise to bands of 128 to 129 nucleotides (Fig. 3B, lane 1). Similarly, a 570-nucleotide probe from the genomic R14 clone protected a nest of fragments centered at 150 nucleotides following digestion with 20 or 50 U of enzyme (Fig. 3B, lanes 2 and 4 and lanes 3 and 5, respectively). Both techniques located the start site at the same position (±2 nucleotides), 36 nucleotides downstream from the TATA box (Fig. 4).

In contrast S1 mapping of the L11 gene resulted in a complex set of bands which we do not yet completely understand (Fig. 3C). The intense series of bands located at 288 to 292 nucleotides was only found when *Aplysia* RNA was added to the reaction mixture, while the set of bands at 326 to 330 nucleotides was also observed in control experiments when tRNA was used in the hybridization and S1 digestion. The 326- to 330-nucleotide bands may be a result of an inverted repeat sequence at this position. The remain-

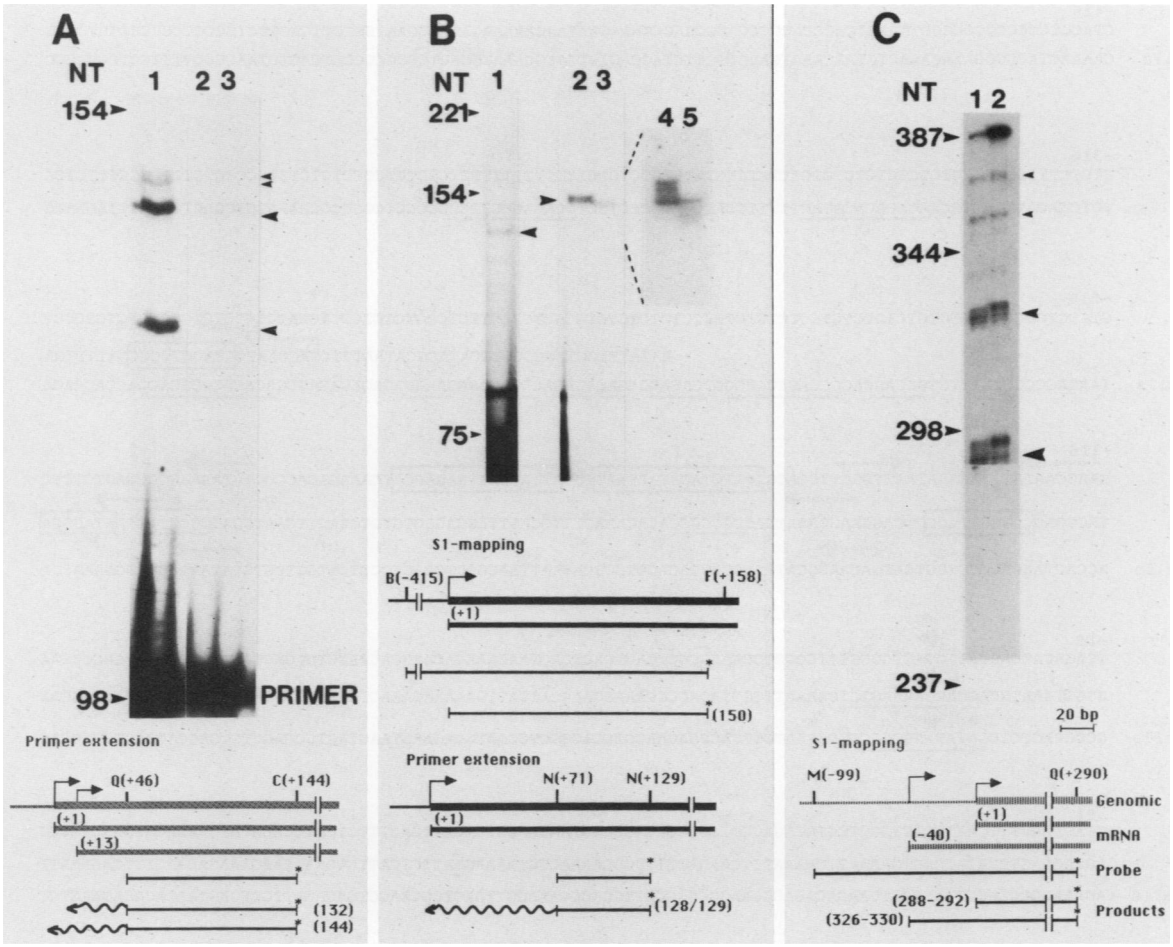

FIG. 3. Analysis of transcription initiation sites. ^{32}P-labeled primer fragments isolated from the cDNA clones were hybridized to RNA in 80% formamide and extended by using reverse transcriptase. ^{32}P-labeled genomic fragments were hybridized to RNA, followed by treatment with the single-strand-specific nuclease S1. The exact definition of the genomic and cDNA products is illustrated below the gels. Reaction mixtures were fractionated on 6% polyacrylamide sequencing gels. Size markers are indicated on the left (in nucleotides), and the bands of interest are demarcated by arrowheads on the right. Smaller arrowheads indicate minor sites. The exact sizes of the products were determined by comparing with sequencing ladders (not shown). (Bottom) The asterisks represent the position of label on the probes, and the wavy lines indicate the extended products. The arrows above the genomic DNA fragments indicate transcription start sites. Enzyme abbreviations: N, *Hin*fI; M, *Dde*I; Q, *Taq*I; C, *Hin*cII; B, *Bam*HI; F, *Fok*I. (A) Primer extension analysis of the ELH gene family with atrial gland polyadenylated RNA (lane 1) or bag cell total RNA (lane 2 and 3). A 98-nucleotide fragment from an ELH cDNA clone was used as the primer. (B) Primer extension (lane 1) and S1 mapping (lanes 2 to 5) of the R14 promoter. A 58-nucleotide fragment from an R14 cDNA clone was used for priming the extension reactions and a 570-nucleotide genomic fragment was used for the S1 analysis. Lanes 2 and 4 are 20 U of S1 nuclease and lanes 3 and 5 are 50 U of S1. Lanes 4 and 5 are longer electrophoretic separations of the S1 products. (C) S1 mapping of the L11 gene with 50 (lane 1) or 20 (lane 2) U of S1 nuclease. A 398-nucleotide genomic fragment was used as the probe.

ing higher-molecular-weight bands may arise from protection by authentic transcripts; however, they were not consistent with our definition of the promoter as described below.

The first exons of all three genes are untranslated leaders separated from the remainder of the coding regions by introns of at least 3 kb. The donor splice junctions all had GT nucleotides at the beginning of the intron, conforming to the splicing consensus rules (6) (Fig. 4). The major site of transcription initiation is indicated as +1 by the large arrow. Between 21 and 36 residues upstream we found sequences which were exact or very close matches to the TATAAAA box observed at this position in numerous eucaryotic genes (M. Goldberg, Ph.D. thesis, Stanford University, Stanford, Calif., 1979). The region spanning −86 to −142 nucleotides had several sequences homologous with the CAAT box, a second sequence often found in the promoters of eucaryotic genes (6). In fact, the R14 promoter contained two perfect CAAT consensus sequences in an inverted orientation at positions −90 to −116. Based on these observations we conclude that these *Aplysia* neuropeptide genes have the common recognizable sequence elements found in eucaryotic promoters.

We noted several other interesting structural characteristics in these promoters (Fig. 4). A number of inverted and direct repeats were found, the most striking of which was a 13 of 16 nucleotides inverted repeat located between the R14 TATA and CAAT homologies, making this region rich in potentially important regulatory sequences. Just upstream of the distal CAAT homology in the R14 promoter, the sequence degenerated to a CT and homopolymer repeat for several hundred bases. A number of possible homologies to other eucaryotic promoter regions are labeled I to V and described in the legend to Fig. 4.

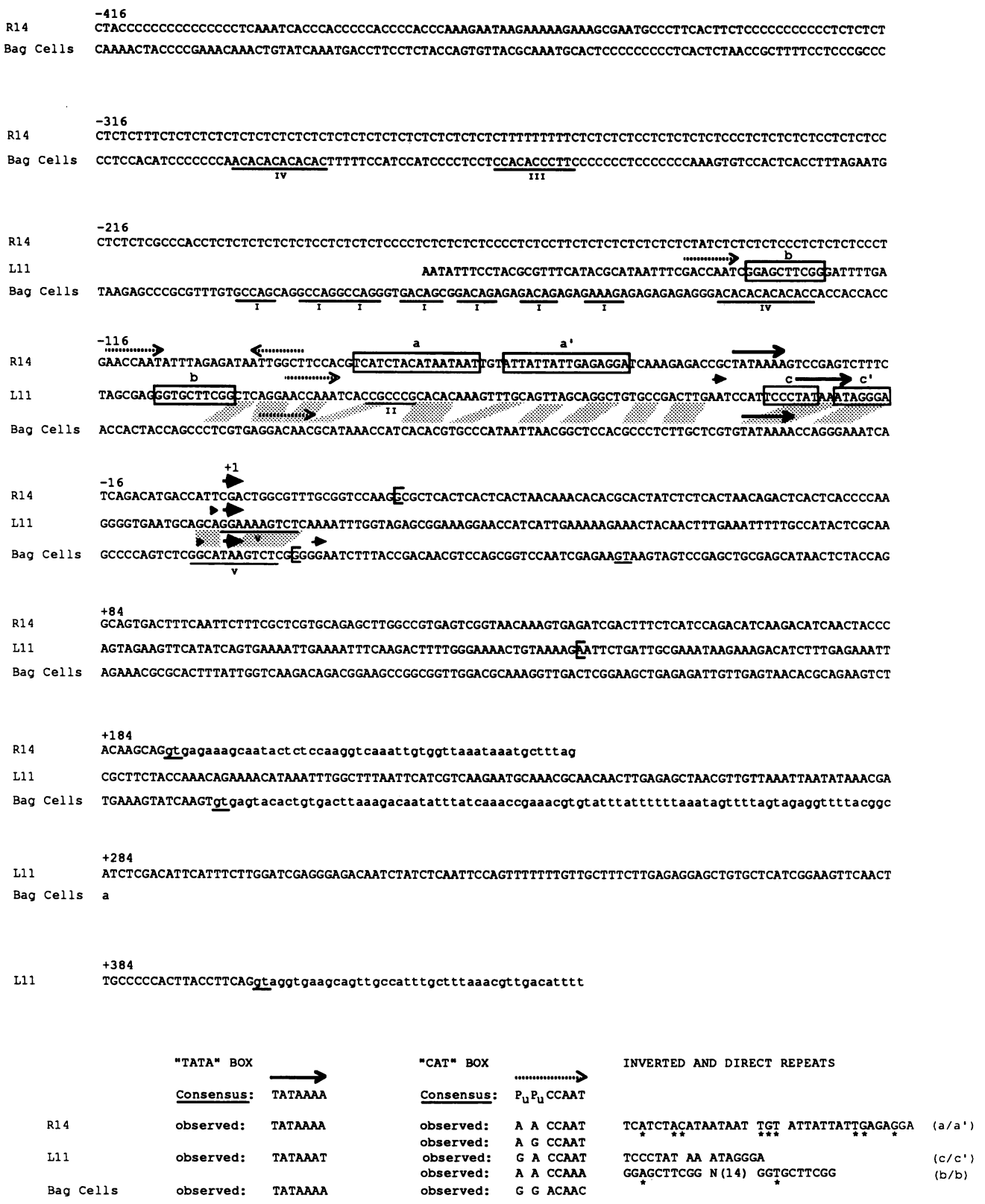

FIG. 4. Nucleotide sequences of R14, L11, and ELH gene promoters. The sequences were determined by the Maxam and Gilbert technique (37) and are aligned by the prominent transcription start sites indicated by the large arrows. Minor start sites are indicated by the smaller arrows. The TATA (→) and CAAT (----→) sequences are indicated. Direct and inverted repeats are labeled a and a′, b and b′, and c and c′, respectively. Regions of homology between the L11 and ELH promoters are shaded. The beginning of the corresponding cDNA clone is indicated by brackets, and the dinucleotide GT spice donor sequence is underlined. The 5′ and transcribed sequences are in capital letters,

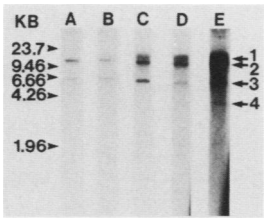

FIG. 5. Blot analysis of DNA microinjected into neurons. Individual abdominal ganglion neurons were injected with 10^8 copies of supercoiled (lanes A and B) or linear (lanes C to E) plasmid pR14CAT-EB (see map in Fig. 7). Two days postinjection, total DNA from individually dissected neurons was isolated, subjected to electrophoresis on a 1% agarose gel, transferred to nitrocellulose, and hybridized to a nick-translated *cat* probe. Size markers are indicated on the left. The numbers on the right correspond to positions of: 1, high-molecular-weight genomic DNA; 2, open circular DNA; 3, linear DNA; and 4, supercoiled DNA. Lane E is a sevenfold-longer exposure of lane D.

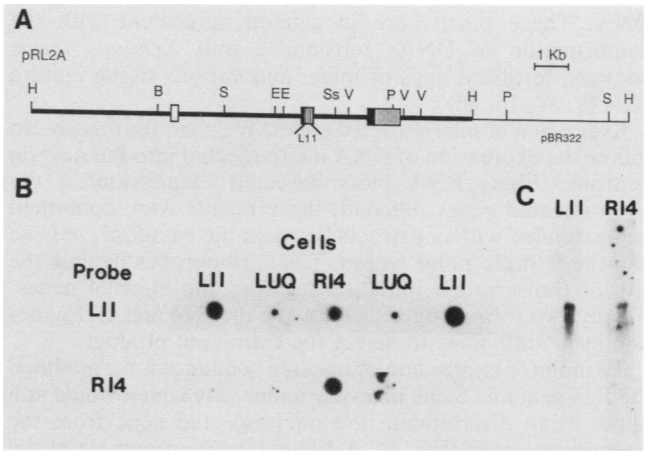

FIG. 6. Specific expression of an R14-L11 hybrid gene. (A) Schematic representation of pRL-2A plasmid DNA. The clone was constructed by inserting 192 nucleotides of L11 coding region into the second exon of the R15 gene. The symbols are the same as in Fig. 2. The cells were injected with about 10^7 copies of the construct; 2 days postinjection, the neurons were dissected and the RNA was purified. (B) Half of the RNA was bound to nitrocellulose for a dot blot assay. The blot was hybridized to the L11 probe, exposed, stripped of the bound radioactivity, and rehybridized to the R14 probe. As a control the probe was washed off and rehybridized with pBR322; no hybridization was detected. However, in another series of injections, the second half of the purified RNA was not DNase treated. In this case every cell hybridized to the pBR322 probe, showing that all neurons received an approximately equal amount of injected DNA (not shown). (C) Half of the RNA was fractionated on formaldehyde-agarose gels, blotted, and hybridized. Here we show hybridization to RNA from an injected L11 neuron and an injected R14 neuron probed with L11 sequences. The transcripts are consistent with previous studies of the expression of these genes.

When the L11 and ELH gene promoters were aligned at their transcription initiation sites, many short regions of homology, ranging from 3 to 7 nucleotides, were observed (Fig. 4). We were not able to find similar stretches of homology to the R14 promote or random sequences. These two genes do not have any other similarities at the DNA or protein level. Therefore, these promoter regions may have converged on this homology, which most likely represents sites for protein-DNA interactions.

Fate of DNA microinjected into *Aplysia* neurons. Clearly, one needs functional assays to delineate the roles of the promoter regions described above, as well as to define other potential regulatory sequences. Below we outline a series of experiments which addressed the fate of DNA sequences microinjected into *Aplysia* neurons. We began this analysis by investigating the state of the injected DNA and then analyzed its expression.

Abdominal ganglia were dissected and pinned in a Silgard dish containing artificial sea water. Microelectrodes were filled with DNA at a concentration of 200 μg/ml, suspended in 1 M KCl and 0.01% fast green. In the expression studies a volume of about 0.1 to 1 nl was injected, corresponding to 10^6 to 10^7 copies of the DNA sequence. We recorded the membrane potential during the injection procedure and monitored retention of the dye during the incubation period. In the experiments reported below the ganglia were maintained for 48 h at 16°C in L15 medium.

Figure 5 shows an analysis of supercoiled and linear DNA injected into *Aplysia* neurons. For these experiments it was necessary to inject about 10^8 copies of the constructs to allow subsequent detection in Southern blots. Lanes A and B show injected supercoiled DNAs which migrated at the position of band 4. After the 48-h incubation period, the supercoiled plasmids were recovered as molecules which migrated precisely at the positions of relaxed circles (band 2) and linear plasmid (band 3). Even after longer exposure, no other bands were observed in these lanes. An analysis of injected linear DNA is shown in Fig. 5, lanes C and D. After incubation the injected linear DNA was found in a variety of other forms, including relaxed circular DNA, and after longer exposure a small amount of material was occasionally detected at the position of supercoiled DNA (Fig. 5, lane E, band 4). In addition, some material was found to migrate as a smear at higher molecular weights (band 1). When the gels were stained with ethidium bromide prior to blotting, the bulk of the chromosomal DNA from the neurons was found to migrate at this position. This high-molecular-weight DNA most likely represents long concatenates of the plasmid, but we do not know whether this material was integrated into the chromosome. The other possibility is that this DNA represents many independent integrations into the chromosomal

and the introns are in lowercase letters. The numbers are relative to the transcription initiation start site. The lower region shows a comparison with consensus TATA and CAAT sequences, as well as the repeated sequences. The repeat in the R14 (a/a') gene is homologous to an inverted repeat sequence seen in the human beta interferon gene (TCCTCTGAATAGAGAGAGGA; the underlined region indicates the homology) (19). Sequences which may correspond to known eucaryotic promoter enhancer signals are underlined and labeled I to V. I: GC/ACAG. This sequence is repeated seven times in the 5' end of the ELH gene and is similar to the core sequence found in a number of enhancers (PuPuGA/CCAGPuPu) (2, 10, 17). II: CCGCCC. This sequence in the L11 gene is similar to the simian virus 40 21-bp repeat consensus (CCGCCC) (15). III: CCACACCCTT. A bag cell sequence similar to the globin upstream sequence (CCXCACCCTG) (13). IV: ACACACACACAC. Alternating stretches of purines and pyrimidines which may adopt a ZDNA conformation and possess enhancer activity (23). V: GGA/CAT/AAAGTCTC. A shared sequence near the L11 and bag cell transcription initiation sites that is similar to an insect non-heat shock gene sequence described previously (ATCAG/TTC/T) (24).

DNA. These results are in general agreement with the conformation of DNAs introduced into *Xenopus laevis* oocytes, fertilized eggs of mice, and various tissue culture cells (8, 14, 16, 40).

Expression of microinjected genes. We used four assays to follow the expression of DNA microinjected into the *Aplysia* neurons. First, RNA blots detected expression of the microinjected genes. Second, these results were confirmed and extended with constructs between the bacterial *cat* gene and the R14 promoter region. Third, riboprobes defined the site of transcription initiation used by the injected genes. Finally, we formed hybrids between the R14 and L11 genes and used antibodies to detect the expressed product.

To monitor expression of injected sequences, we modified the R14 genomic clone in a very minor way which would still allow us to discriminate the microinjected gene from the endogenous gene (Fig. 6). A 192-nucleotide fragment of the L11 coding region was inserted, in frame, into the middle exon of the R14 gene. In this way expression of the construct in R3-14 neurons could be detected by probing with L11, while expression in other neurons could be detected by probing with R14 gene fragments.

Many identified neurons have been injected with the R14-L11 hybrid construct and analyzed by RNA blot assays. Five injected cells are shown in Fig. 6B. When the L11 probe was hybridized to the dot blot RNA, all of the neurons showed hybridization above background. When this blot was washed and reprobed with R14 gene sequences, expression was detected only in R14, not in the L11 and LUQ cells. Therefore, these experiments suggest that the microinjected gene is expressed specifically in R14 and not the other cells investigated. These results also show that the microinjected gene is expressed at roughly the same level as the endogenous R14 gene (compare hybridization in R14 cells with the L11 and R14 probes). DNase treatment and control experiments established that the hybridization was to RNA and not to the microinjected DNA (Fig. 6). To further establish this point, we analyzed the other half of the RNA from these cells in Northern assays. Figure 6C shows hybridization of the L11 gene probe to the RNA from a microinjected L11 and a microinjected R14 neuron. The endogenous L11 transcript of 1.2 kb was detected in L11 cells, while R14 probes did not hybridize. In the RNA from the R14 neuron, hybridization with the L11 probe revealed a band consistent with transcription of the microinjected gene initiating at the correct site. We injected 38 cells with this construct. Five of the 11 R14 neurons expressed the microinjected gene, while we did not detect expression in 10 R3-13 neurons. None of the 17 injected neurons in the left hemiganglion, including L11 and LUQ cells, expressed the gene. These data demonstrate that the pRL-2A construct was specifically expressed in R14 and not other neurons, with a confidence level of 0.025 (5 of 11 versus 0 of 17; χ^2 test).

To further analyze this promoter region, a series of four constructs between the bacterial *cat* gene and the R14 promoter were injected into *Aplysia* neurons (Fig. 7). The locations of the R14 DNA used in these constructs relative to the transcribed portion of the gene are shown in Fig. 2. The constructs included either 4 or 0.4 kb of 5′-flanking sequence and a portion of the 5′ exon. These DNAs were injected into a variety of cells in organ culture (over 150 cells injected), incubated for 48 h, dissected from the ganglion, and assayed for CAT activity (Fig. 7). Approximately 20% of the microinjected neurons expressed these genes. No differences in expression of these four constructs were observed, and the amount of activity varied from barely detectable to

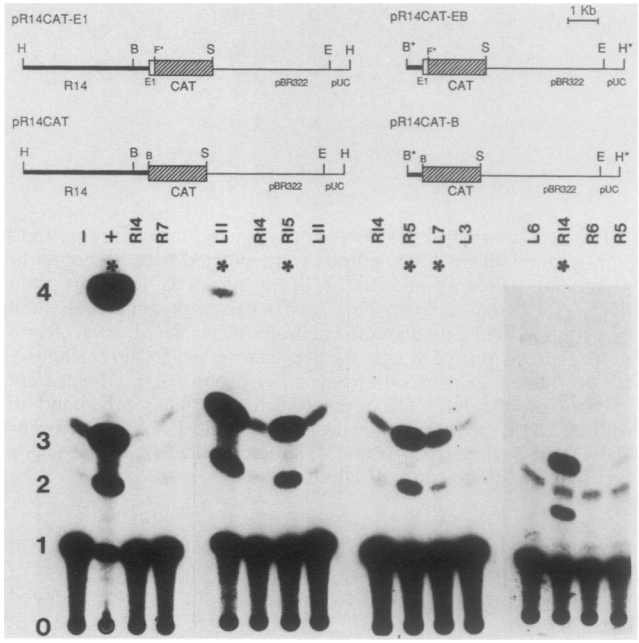

FIG. 7. Analysis of the expression of R14-*cat* hybrid genes. (Top) Schematic representation of four recombinants between the 5′ region of the R14 gene and the *cat* gene. The recombinants contain either 4 kb (pR14CAT-E1 and pR14CAT) or 420 nucleotides (pR14CAT-EB and pR14CAT-B) of 5′-flanking DNA and the first untranslated exon (pR14CAT-E1 and pR14 CAT-EB). The symbols for the restriction enzymes are as in Fig. 2, and the striped boxes represent the *cat* gene. E1 represents the R14 first exon, and asterisks represent loss of a site after filling in and ligation. The small B represents a new *Bam*HI site generated from filling in and blunt-end ligation. (Bottom) Thin-layer chromatography analysis of CAT activity. The constructs shown above were expressed in all cells investigated. Four representative autoradiographs of chromatography plates are shown. Assays with either no enzyme or purified bacterial CAT are shown in the first two lanes, respectively. All other lanes contained individually microinjected neurons. The numbers to the left represent the positions of the origin (0), ^{14}C-labeled chloramphenicol (position 1) and the mono- (positions 2 and 3) and di- (position 4) acetylated forms. The stars represent cells with activity above background. Reading from left to right, these cells were injected with the following constructs: R14, R7 (pR14CAT-E1), L11 (pR14CAT-EB), R14, R15, L11 (pR14CAT), R14, R5, L7, L3, L6, R14, R6, and R5 (pR14CAT-EB).

greater than 10-fold the background level. In addition, expression was not restricted to neurons which normally express the R14 gene, suggesting that the 420-nucleotide *Bam*HI-*Taq*I fragment serves as a constitutive promoter element.

It was important to determine whether the promoter, as defined earlier in this report, was active in expressing the microinjected genes. This was done by injecting the construct shown in Fig. 8 and using SP6 riboprobes to determine the region protected by transcripts. This construct was a truncated version of pRL-2A, which contains the complete mRNA coding region, 0.4 kb of 5′ and 2.3 kb of 3′ DNA. Figure 8, lane A, shows an uninjected R14 neuron which expressed the endogenous peptide gene. The 153-nucleotide band was approximately the expected size for initiation at the promoter site as described earlier. Lanes B, C, and D represent RNase protection of transcripts expressed in injected cells which do not normally express the R14 gene. The size of this transcript again corresponded to that expected

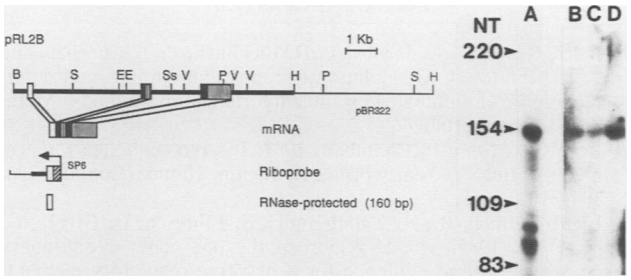

FIG. 8. RNase mapping of a microinjected R14 gene. (Left) Schematic representation of the microinjected gene and the RNase mapping strategy. The R14 gene was truncated at the *Bam*HI site. The SP6 probe contained 265 nucleotides of the 5′ end of the *cat* gene, the R14 first exon, and 420 nucleotides of the R14 promoter region. Expression of the microinjected gene will result in protection of a 160-nucleotide fragment. Symbols and abbreviations are as in Fig. 2. (Right) RNase mapping of microinjected cells. RNA was hybridized overnight in 80% formamide at 47°C, digested with RNases A and T_1 at 35°C for 30 min, and fractionated on nondenaturing 5% polyacrylamide gels. DNA size markers (in nucleotides) are shown to the left and may not exactly predict the size of the RNA bands. Lanes: A, uninjected R14; B, injected L7; C, injected L11; D, injected LUQ (L2).

for transcription initiation at the promoter site displayed in Fig. 4. This construct was expressed in microinjected neurons regardless of their identity.

Therefore, while we have not precisely defined the sequences necessary for R14-specific expression, the above data suggest an interaction between the promoter region and sequences both 5′ and 3′, forming a multicomponent repressor system.

DNA-mediated switching of transmitter phenotype. At this point we had established that large *Aplysia* neurons can express microinjected genes, suggesting that the system will be useful for investigations of neuron cell biology and function. For instance, can neuropeptides normally expressed in one cell be expressed in another cell, and if so, are they processed and packaged in a similar fashion? To begin an analysis of this and other questions, we joined the 5′ region of the L11 gene to the coding portion of the R14 gene at a unique *Sal*I site in the first intervening sequence of both genes (Fig. 9). This gene was injected into various cells, and expression of the peptide product was assayed by using antibodies which recognize the R14 neuropeptide precursor. These antibodies were generated against synthetic peptides defined by the sequence of a cDNA clone and have been extensively characterized elsewhere (30). Injection of this construct into L11 neurons resulted in the synthesis of R14-immunoreactive material; however, we do not yet know whether the material was processed and localized in dense core vesicles as expected. In control experiments, microinjection of numerous other DNAs lacking the R14 sequences did not generate any immunoreactivity. We have not analyzed enough cells to have statistically significant evidence relevant to the specificity of the expression, yet these data demonstrate the utility of this system in expressing neuronal protein products in identified neurons.

DISCUSSION

Our analysis of *Aplysia* promoters demonstrates that these sequences conform to the general consensus derived from the analysis of eucaryotic genes. TATA and CAAT sequences were found in the usual positions, emphasizing their important roles in governing expression of these genes (6; Goldberg, Ph.D. thesis). The presence of two CAAT sequences in the L11 and R14 promoters is unusual. The second CAAT sequence in the L11 gene is quite distal to the TATA box and is not a perfect match to the consensus sequence. Nevertheless, this sequence could be involved in the potential secondary initiation sites observed in the promoter mapping experiments. The R14 gene has two perfect matches to the consensus sequence, in inverted orientation, separated by 11 nucleotides, or about one turn of B DNA. Graves et al. (22) have shown that an inverted copy of the CAAT sequence can bind a nuclear transcription factor and function to modulate transcription of the herpesvirus thymidine kinase (*tk*) gene. It is therefore likely that both of these sequences can function in vivo; however, it is not clear whether they act in a coordinate fashion or independently.

The unexpected finding of many short homologies between the L11 and bag cell promoters suggests that similar or identical molecules may interact with these regions of DNA. These otherwise unrelated genes may be coordinately regulated by common molecules, for example during egg laying, and independently regulated by factors which recognize unique sequences outside the promoter regions.

Of additional interest is the homology (sequence V in Fig. 4) located at the transcription initiation site of the L11 and bag cell genes. A similar sequence is found at the initiation sites of many insect (24), sea urchin histone (9), and other genes (29). The sequence was not found in the R14 gene, nor has it been identified in vertebrate genes. Its role in gene expression remains unknown; however, its presence in an additional phylum suggests that its function is quite important.

Cloned *Aplysia* genes were expressed when microinjected into neurons maintained in organ culture. As in other systems, the DNA underwent a number of transitions after microinjection, including linearization, cyclization, ligation, and supercoiling. The results demonstrate that about 20% of the microinjected neurons expressed the exogenous genes; however, it was difficult to determine whether the DNA was

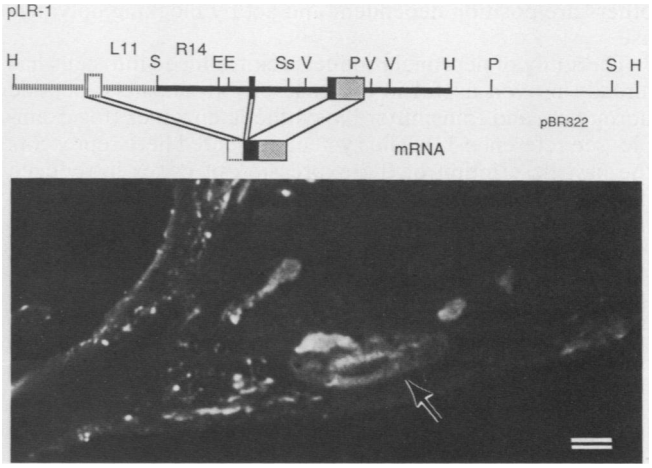

FIG. 9. Injection of DNA can alter transmitter phenotype. (Top) Hybrid L11-R14 gene constructed by joining the first exon of L11 to the second and third exons of R14 at a unique *Sal*I site within the first intron. Abbreviations and symbols are as in Fig. 2. (Bottom) Microinjected L11 neuron stained with antibodies which recognize the R14 neuropeptide precursor and peptides (arrow). Uninjected neurons or neurons injected with other constructs did not stain with this antibody. Bar, 50 μm.

being directly introduced into the nucleus. *Aplysia* neurons are highly polyploid, containing up to 2×10^5 copies of the genome (12, 34), while the injections presented in this report were usually of about 10^6 to 10^7 copies of DNA. Injection of an order of magnitude less DNA did not result in detectable expression. Thus, while we were able to work with single neurons, copies of important regulatory factors may have been diluted in these experiments. Finally, we have not yet systematically investigated potential physiological factors which may influence expression of the genes and lead to uncontrolled variation between individual animals.

Many different constructs are expressed when driven by *Aplysia* promoters; however, viral promoters were poorly effective in driving the *cat* gene (the constructs investigated were MSV-CAT [33], SV2-CAT [21], and RSV-CAT [20]). The 420-nucleotide *Bam*HI-*Taq*I fragment of the R14 promoter was sufficient to direct accurate transcription of the *cat* gene, but this expression was not restricted to a particular cell type. We also demonstrated nonspecific expression of R14 constructs, which included 420 nucleotides of 5′ DNA, the whole transcribed sequence, and 2.3 kb of 3′ untranslated sequence. Only when a large genomic fragment was introduced into cells including 4 kb of 5′ DNA, all introns and exons, as well as 2.3 kb of 3′ DNA was the gene expressed in a cell-specific fashion. This specific expression was limited to R14 and did not include the anatomically distinct neurons R3 to R13, which express the endogenous gene.

While these issues are in need of more detailed investigation, there appear to be multiple *cis*-acting sequences governing expression of the R14 gene. The constitutive promoter region was found directly in front of the gene. The fact that additional 5′ and 3′ sequences were required for cell-specific expression suggests that a complex repression mechanism may turn off the gene in non-R14 neurons. Negative regulation is well established in procaryotes as well as yeast (5, 7, 43), and examples in which deletion of a sequence leads to increased expression of a gene have been reported in eucaryotic systems as well (3, 11, 18, 19, 25, 41, 42, 44, 51). Some of these repressor functions act at a distance and in an orientation-independent manner, while others are position dependent and act by blocking polymerase.

The study of neuronal products reintroduced into cells has already proven useful in the case of ion channels, peptide hormones, and transmitter biosynthetic enzymes (for example, see reference 32). The system presented here represents the first description of the expression of genes introduced directly into neurons. This affords the possibility for studying neuronal products in a well-defined environment. In particular, the ability to study gene expression in different physiological states of the neuron may help lead to an understanding of the relationship between neuronal electrical activity and gene expression.

ACKNOWLEDGMENTS

We thank Corey Goodman and Wayne Sossin for the critical reading of this manuscript. We also thank Jack Dixon of Purdue University for help with the initial experiments and useful discussions.

L.D. was supported by a centennial grant from the Medical Research Council of Canada. This work was supported by a Public Health Service grant to R.H.S. from the National Institutes of Health.

LITERATURE CITED

1. **Beltz, G. A., K. A. Jacobs, T. H. Eickbush, P. T. Cherbas, and F. C. Kafatos.** 1983. Isolation of multigene families and determination of homologies by filter hybridization methods. Methods Enzymol. **100**:266–285.
2. **Benoist, C., and P. Chambon.** 1981. In vivo sequence requirements of the SV40 early promoter region. Nature (London) **290**: 304–310.
3. **Bienz-Tadmor, B., R. Zakut-Houri, S. Libresco, D. Givol, and M. Oren.** 1985. The 5′ region of the p53 gene: evolutionary conservation and evidence for a negative regulatory element. EMBO J. **4**:3209–3213.
4. **Bousfield, D.** 1985. Neurotransmitters in action. Elsevier Biomedical Press, Amsterdam.
5. **Brand, A. H., L. Breeden, J. Abraham, R. Sternglanz, and K. Nasmyth.** 1985. Characterization of a "silencer" in yeast: a DNA sequence with properties opposite to those of a transcriptional enhancer. Cell **41**:41–48.
6. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
7. **Brent, R.** 1985. Repression of transcription in yeast. Cell **42**:3–4.
8. **Brinster, R. L., H. Y. Chen, M. E. Trumbauer, M. K. Yagle, and R. D. Palmiter.** 1985. Factors affecting the efficiency of introducing foreign DNA into mice by microinjecting eggs. Proc. Natl. Acad. Sci. USA **82**:4438–4442.
9. **Busslinger, M., R. Portmann, J. C. Irminger, and M. L. Birnstiel.** 1980. Ubiquitous and gene-specific regulatory 5′ sequences in a sea urchin histone DNA clone coding for histone protein variants. Nucleic Acids Res. **8**:957–977.
10. **Carthew, R. W., L. A. Chodosh, and P. A. Sharp.** 1985. An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell **43**:439–448.
11. **Chambon, P., M. P. Gaud, J. P. Lepennec, A. Dierich, and D. Astinotti.** 1984. Steroid hormones relieve repression of the ovalbumine gene promoter in chick oviduct tubular gland cells. *In* F. Labrie and L.Proux (ed.), Endocrinology. Excerpta Medica, Amsterdam.
12. **Coggeshall, R. E., B. A. Yaksta, and F. J. Swartz.** 1971. A cytophotometric analysis of the DNA in the nucleus of the giant cell, R2, in *Aplysia*. Chromosoma **32**:205–212.
13. **Dierks, P., A. Van Ooyen, M. D. Cochran, C. Dobkin, J. Reiser, and C. Weissmann.** 1983. Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit beta-globin gene in mouse 3T6 cells. Cell **32**:695–706.
14. **Etkin, L. D., B. Pearman, M. Roberts, and S. L. Bektesh.** 1984. Replication, integration and expression of exogenous DNA injected into fertilized eggs of Xenopus laevis. Differentiation **26**:194–202.
15. **Everett, R. D., D. Baty, and P. Chambon.** 1983. The repeated GC-rich motifs upstream from the TATA box are imortant elements of the SV40 early promoter. Nucleic Acids Res. **11**: 2447–2464.
16. **Folger, K. R., E. A. Wong, G. Wahl, and M. R. Capecchi.** 1982. Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. Mol. Cell. Biol. **2**: 1372–1387.
17. **Fujita, T., S. Ohno, H. Yasumitsu, and T. Tanigushi.** 1985. Delimitation and properties of DNA sequences required for the regulated expression of human interferon-beta gene. Cell **41**: 489–496.
18. **Gonzalez, F. J., and D. W. Nebert.** 1985. Autoregulation plus upstream positive and negative control regions associated with transcriptional activation of the mouse P1-450 gene. Nucleic Acids Res. **13**:7269–7288.
19. **Goodbourn, S., H. Burstein, and T. Maniatis.** 1986. The human B-interferon gene enhancer is under negative control. Cell **45**: 601–610.
20. **Gorman, C. M., G. T. Merlino, M. C. Willingham, I. Pastan, and B. H. Howard.** 1982. The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of

eucaryotic cells by DNA-mediated transfection. Proc. Natl. Acad. Sci. USA **79**:6777–6781.
21. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
22. **Graves, B. J., P. F. Johnson, and S. L. McKnight.** 1986. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. Cell **44**:565–576.
23. **Hamada, H., M. Seidman, B. H. Howard, and C. M. Gorman.** 1984. Enhanced gene expression by the poly(dT-dG)-poly(dc-dA) sequence. Mol. Cell. Biol. **4**:2622–2630.
24. **Hultmark, D., R. Klemenz, and W. J. Gehring.** 1986. Translational and transcriptinal control elements in the untranslated leader of the heat-shock gene hsp22. Cell **44**:429–438.
25. **Jones, P. B. C., D. R. Galeazzi, J. M. Fisher, and J. P. Whitlock, Jr.** 1985. Control of cytochrome P1-450 gene expression by dioxin. Science **227**:1499–1502.
26. **Kaldany, R. R., J. R. Nambu, and R. H. Scheller.** 1985. Neuropeptides in identified *Aplysia* neurons. Annu. Rev. Neurosci. **8**:431–455.
27. **Kandel, E. R.** 1976. Cellular basis of behavior. W. H. Freeman and Co., San Francisco.
28. **Kandel, E. R.** 1979. Behavioral biology of *Aplysia*. W. H. Freeman and Co., San Francisco.
29. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eucaryotic mRNA's. Nucleic Acids Res. **12**:857–872.
30. **Kreiner, T., J. B. Rothbard, G. K. Schoolnik, and R. H. Scheller.** 1984. Antibodies to synthetic peptides defined by cDNA cloning reveal a network of peptidergic neurons in *Aplysia*. J. Neurosci. **4**:2581–2589.
31. **Kreiner, T., W. Sossin, and R. H. Scheller.** 1986. Localization of *Aplysia* neurosecretory peptides to multiple populations of dense core vesicles. J. Cell Biol. **102**:769–782.
32. **Kubo, T., K. Fukuda, A. Mikami, A. Maeda, H. Takahashi, M. Mishina, T. Haga, K. Haga, A. Ichiyama, K. Kangawa, M. Kojima, H. Matsuo, T. Hirose, and S. Numa.** 1986. Cloning, sequencing and expression of complementary DNA encoding the muscarinic acetylcholine receptor. Nature (London) **323**: 411–416.
33. **Laimins, L. A., G. Khoury, C. M. Gorman, B. Howard, and P. Gruss.** 1982. Host-specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. USA **79**:6453–6457.
34. **Lazek, R. J., and W. J. Dower.** 1971. *Aplysia californica*: analysis of nuclear DNA in individual nuclei of giant neurons. Science **172**:278–280.
35. **Mahon, A. C., J. R. Nambu, R. Taussig, M. Shyamala, A. Roach, and R. H. Scheller.** 1985. Structure and expression of the egg laying hormone gene family in *Aplysia*. J. Neurosci. **5**: 1872–1880.
36. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
37. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
38. **McKnight, S. L., and R. Kingsbury.** 1982. Transcriptional control signals of a eucaryotic protein-coding gene. Science **217**:316–324.
39. **Nambu, J. R., R. Taussig, A. C. Mahon, and R. H. Scheller.** 1983. Gene isolation with cDNA probes from identified *Aplysia* neurons: neuropeptide modulators of cardiovascular physiology. Cell **35**:47–56.
39a. **Newcomb, R., and R. Scheller.** 1987. Proteolytic processing of the *Aplysia* egg-laying hormone and R3-14 neuropeptide precursors. J. Neurosci. **7**:854–863.
40. **Perucho, M., D. Hanahan, and M. Wigler.** 1980. Genetic and physical linkage of exogenous sequences in transformed cells. Cell **22**:309–317.
41. **Ragg, H., and C. Weissmann.** 1983. Not more than 117 base pairs of 5′ flanking sequence are required for inducible expression of a human IFN-A gene. Nature (London) **303**:439–442.
42. **Remmers, E. F., J. Q. Yang, and K. B. Marcu.** 1986. A negative transcriptional control element located upstream of a murine c-myc gene. EMBO J. **5**:899–904.
43. **Reznikoff, W. S., D. A. Siegele, D.W. Cowing, and C. A. Gross.** 1985. The regulation of transcription initiation in bacteria. Annu. Rev. Genet. **19**:355–387.
44. **Rosen, C. A., J. G. Sodroski, and W. A. Haseltine.** 1985. The location of cis-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. Cell **41**:813–823.
45. **Scheller, R. H., J. F. Jackson, L. B. McAllister, B. S. Rothman, E. Mayeri, and R. Axel.** 1983. A single gene encodes multiple neuropeptides mediating a stereotyped behavior. Cell **32**:7–22.
46. **Scheller, R. H., and M. D. Kirk.** 1987. *Aplysia* neuropeptides: a molecular genetic, biochemical and physiological analysis. Trends Neurosci. **10**:46–52.
47. **Taussig, R., R. R. Kaldany, J. B. Rothbard, G. Schoolnik, and R. H. Scheller.** 1985. Expression of the L11 neuropeptide gene in the Aplysia central nervous system. J. Comp. Neurol. **238**: 53–64.
48. **Taussig, R., R. R. Kaldany, and R. H. Scheller.** 1984. A cDNA clone encoding neuropeptides isolated from *Aplysia* neuron L11. Proc. Natl. Acad. Sci. USA **81**:4988–4992.
49. **Taussig, R., M. R. Picciotto, and R. H. Scheller.** 1984. Two introns define functional domains of a neuropeptide precursor in Aplysia, p. 551–660. *In* E. H. Davidson and R. A. Firtel (ed.), Molecular biology of development, vol. 19. Alan R. Liss, Inc., New York.
50. **Weaver, R. F., and C. Weissmann.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15S beta-globin mRNA precursor and mature 10S beta-globin mRNA have identical map coordinates. Nucleic Acids Res. **7**: 1175–1193.
51. **Zinn, K., D. DiMalo, and T. Maniatis.** 1983. Identification of two distinct regulatory regions adjacent to the human B-interferon gene. Cell **34**:865–879.