Isolation, Characterization, and Evolutionary Aspects of a cDNA Clone Encoding Multiple Neuropeptides Involved in the Stereotyped Egg-Laying Behavior of the Freshwater Snail Lymnaea stagnalis

E. Vreugdenhil,[1] J. F. Jackson,[3] T. Bouwmeester,[1] A. B. Smit,[1] J. Van Minnen,[1] H. Van Heerikhuizen,[2] J. Klootwijk,[2] and J. Joosse[1]

[1]Biological Laboratory, and [2]Biochemisch Laboratorium, Vrije Universiteit, 1081 HV Amsterdam, The Netherlands, and [3]Memorial Sloan-Kettering Cancer Center, New York, New York 10021

The cerebral neurosecretory caudodorsal cells (CDCs) of the freshwater pulmonate snail Lymnaea stagnalis control egg laying, an event that involves a pattern of stereotyped behaviors. The CDCs synthesize and release multiple peptides, among which is the ovulation hormone (CDCH). It is thought that each peptide controls a specific aspect of the processes involved in egg laying. We isolated and characterized a CDC-specific cDNA clone that encodes the ovulation hormone (CDCH). RNA blot analysis and in situ hybridization experiments demonstrated that the CDCs are the major cell groups in the cerebral ganglia that transcribe the CDCH gene. In addition to CDCH, the 259-amino acid-long CDCH prohormone contains 11 other predicted peptides. The overall homology of the CDCH preprohormone with the egg-laying hormone (ELH) preprohormones of the marine opisthobranch snails Aplysia californica and A. parvula is very low (29 and 26%, respectively). However, a more detailed comparison revealed a highly differential pattern of conservation of peptide regions. Significant homology was found between the regions containing (1) CDCH and ELH, (2) repeated pentapeptides, (3) alpha-caudodorsal cell peptide and alpha-bag cell peptide, and (4) 2 regions representing as yet unidentified peptides. Insignificant homology was found when comparing regions containing the other predicted peptides. The conserved peptides probably control similar aspects of the egg-laying fixed action patterns in these distantly related gastropod species. The pentapeptide region exhibits the highest level of homology (75%); in addition, an extra pentapeptide has been generated on the CDCH precursor. This indicates a vital function of these peptides in Aplysia, as well as in Lymnaea species.

Lower animals provide simple behavioral systems that are attractive for the analysis of the role of the CNS in the control of behavior. The gastropod molluscs Lymnaea stagnalis and Aplysia californica are particularly advantageous models for such studies; their CNS consist of only 10,000–20,000 neurons, many of which are unusually large and uniquely identifiable. Egg laying in these animals is an extensively studied and well-defined fixed action pattern involving a series of internal processes (ovulation, egg and egg mass or egg string formation) that closely correspond to a pattern of overt behaviors such as alteration of locomotion and feeding, specific postures, and ovipositions (Cobbs and Pinsker, 1982; Rothman et al., 1985; Geraerts et al., 1987a). In the freshwater pulmonate L. stagnalis, these events are controlled by the neurosecretory caudodorsal cells (CDCs), large (diameter ca. 70 μm) polyploid cells that are located in 2 homogeneous clusters (total of 100 neurons) in the cerebral ganglia near the origin of the intercerebral commissure (COM), which serves as their neurohemal area (Roubos, 1984). The control event in egg laying is a ca. 60 min, pacemaker-driven, electrical discharge of all CDCs that can be elicited by the appropriate environmental stimuli (Kits, 1980; Ter Maat et al., 1983). During the discharge, the CDCs release the ovulation hormone (CDCH) from the contents of secretory granules in CDC axon terminals located in the COM (Geraerts et al., 1983, 1984). CDCH, a 36-amino acid-long peptide (Ebberink et al., 1985), induces ovulation and egg-mass production; furthermore, it probably also stimulates synthesis of secretory products in the female accessory sex glands and affects neurons in the neuronal circuits controlling locomotion and feeding (Goldschmeding et al., 1983; Wijdenes et al., 1983; Jansen and Ter Maat, 1985). During the discharge, 2 other identified peptides are released by the CDCs: CDCA (Mr < 1500 Da), which acts as an autotransmitter (Ter Maat et al., 1988a), and calfluxin, which affects calcium movements in the albumen gland cells, a female accessory sex gland (Dictus et al., 1987). In addition, the CDCs release at least 7 other peptides with unknown functions (Geraerts et al., 1987b). The combined actions of this set of released peptides are thought to be responsible for the initiation and coordination of the egg-laying fixed action pattern (Geraerts et al., 1987a).

Egg laying in the distantly related marine opisthobranch A. californica is controlled by the neurosecretory bag cells (BC), which have similar morphological and electrophysiological characteristics as the CDCs (Geraerts et al., 1987a). Several biologically active peptides have been purified from BC extracts and sequenced, including the egg-laying hormone (ELH) and alpha-bag cell peptide (α-BCP) (Chiu et al., 1979; Rothman et al., 1983). ELH performs similar functions as CDCH. This sim-

ilarity in biological action is reflected in the primary structure of both peptides; these peptides are 36 amino acids long, have a strongly basic character, and share 16 amino acids in the same position (Ebberink et al., 1985).

Using molecular biological techniques it was demonstrated that ELH and α-BCP are synthesized as part of a common preprohormone (Scheller et al., 1983; Jackson et al., 1986). Pulse-chase experiments suggest that CDCH and the other peptides released by the CDCs are also synthesized as part of a common precursor (Geraerts et al., 1985; Vreugdenhil et al., 1985). We have isolated and characterized a cDNA clone encoding the CDCH preprohormone in order to identify all the peptides, which may be involved in the regulation of egg laying in Lymnaea. Comparison of the CDCH preprohormone with the ELH preprohormone revealed a highly differential conservation of the peptides encoded within both preprohormones.

Materials and Methods

Construction and screening of cDNA libraries. Two cDNA libraries were constructed. The first cDNA library was made from mRNA isolated from the total CNS (CNS cDNA library). cDNA was made according to conventional methods (Huynh et al., 1985; Watson and Jackson, 1985), cloned in the Vector λgt10, and, after packaging, transfected to a C600 Hfl strain. Following this protocol we obtained 10^7 independent clones. The second cDNA library was prepared from mRNA isolated from the cerebral ganglia (CG cDNA library). This library was constructed in the Pst I site of the plasmid vector pBR322 by homopolymeric (poly dC) tailing of the cDNA with terminal deoxy nucleotidyl transferase (Boehringer-Mannheim). Approximately 1000 ampicilline-resistant clones were obtained.

The CNS cDNA library was first subjected to differential screening: 20,000 clones were adsorbed to 3 replica nitrocellulose filters and subsequently hybridized with radioactive cDNA (specific activity, 2×10^6 dpm/μg mRNA) synthesized from mRNA extracted from 3 different sources: the cerebral ganglia (positive probe), the total CNS minus the cerebral ganglia, and a non-neural tissue, the digestive gland (both negative probes). The filters were prehybridized, hybridized, washed and autoradiographed using standard techniques (Maniatis et al., 1983). Clones that hybridized with the cerebral ganglia probe only were isolated and purified by rescreening at a lower phage density. A collection of 150 cerebral ganglia positive clones were then individually screened with a synthetic oligonucleotide mixture [5'A(A/G)(A/G)TC(A/G)TC(G/A/T/C)GT(A/G/T)AT] corresponding to a part of the CDCH amino acid sequence (Ile-Thr-Asn-Asp-Leu; see Ebberink et al., 1985). This mixture was ^{32}P-labeled with T4 polynucleotide kinase (Boehringer-Mannheim) at a specific activity of 10^9 dpm/μg DNA. Nitrocellulose filters were prehybridized for 6 hr at 30°C in a solution containing 0.9 M NaCl, 0.1 M Tris-HCl (pH = 7.4), 0.01 M EDTA, 1% (wt/vol) Nonidet P40, 2× Denhardt solution (1 × Denhardt is 0.02% each of BSA, Ficoll, and polyvinylpyrrolidone), 0.2% (wt/vol) SDS, 100 μg/ml denatured salmon sperm DNA, and 5 μg/ml tRNA. The filters were then hybridized at 37°C for 18 hr with the ^{32}P-labeled oligonucleotide probe (10 ng/ml) in the above solution, washed at 24°C in 3 changes of 6× SSC, 0.1% (wt/vol) SDS and autoradiographed for 48 hr. Positive clones were designated as CDCH clones. The CG cDNA library was screened using one of the CDCH clones from the CNS cDNA library as a probe. Two positive clones designated as pLCA and pLCB were obtained.

DNA sequence analysis. Restriction endonuclease fragments were subcloned into the phages M13mp9 and M13mp10 and transformed into *E. coli* JM101 (Maniatis et al., 1983). Recombinant plaques were picked and single-stranded templates prepared and used for the dideoxy-chain termination sequencing described by Sanger et al. (1977). Ambiguities remaining after sequencing of both strands were solved by using the chemical modification and cleavage procedure (Maxam and Gilbert, 1980). The 5'- and 3'-ends of the cDNA were sequenced by using synthetic oligonucleotides derived from known sequences as a primer in the dideoxy-chain termination method.

RNA blot analysis. Total RNA was isolated from the cerebral ganglia using the method of Chirgwin et al. (1979). Size fractionation of RNA, subsequent transfer to Hybond filters (The Radiochemical Centre, Amersham), and hybridization were performed as described by Thomas

(1980). The complete Pst I fragment from pLCA encoding CDCH was used as a probe in the hybridization analysis. This fragment was ^{32}P-labeled by nick translation to a specific activity of 2×10^8 dpm/μg DNA.

In situ hybridization. Whole cerebral ganglia were fixed in Bouin and embedded in Paraplast. Sections of 7 μm were adhered to glutaraldehyde-activated slides. After removal of the Paraplast, sections were hydrolyzed in 0.2 N HCl for 10 min, briefly rinsed in double-distilled, RNAse-free H_2O, treated with 1% (wt/vol) Triton X100 for 10 min, rinsed in double-distilled, RNAse-free H_2O, fixed in 2% paraformaldehyde in PBS (4 min), rinsed in PBS with 0.2% glycerine, rinsed in PBS, upgraded in 100% ethanol, and air-dried. Sections were prehybridized in a moist chamber, containing 50% formamide, in a solution containing 50% formamide, 3× SSC, 5× Denhardt solution, 2 mM DDT, 10 μg/ml salmon sperm DNA, and 200 μg/ml tRNA for 1 hr at 37°C. After prehybridization, slides were upgraded in 100% ethanol and air-dried. The probe utilized during the hybridization was the same as described in the RNA blot analysis except that (α-^{35}S)dATP was used as a label during the nick translation. The hybridization conditions were performed under the same conditions as the prehybridization, including the nick-translated probe (specific activity, 2×10^8 dpm/μg; 10 ng/slide) instead of the carrier DNA, for 17 hr at 37°C. Slides, hybridized with nick-translated pBR322 DNA, served as a control. After hybridization, the slides were washed in 3 subsequent steps in 2× SSC (room temperature, 30 min), 0.1× SSC (room temperature, 30 min), and 0.1× SSC (60°C, 10 min). Finally, slides were upgraded in 100% ethanol, dipped in Kodak NTB 2 emulsion, and exposed for 24 and 48 hr.

Results

Isolation of CDCH cDNA clones

To isolate cDNA clones encoding CDCH and related peptides a 2-step procedure was used. Since the CDCs synthesize relatively high levels of CDCH (Ebberink et al., 1985), it seems very likely that the mRNA encoding the CDCH preprohormone is highly abundant in these cells. We therefore used differential screening techniques in the first step of the screening strategy. We screened a CNS cDNA library using cDNA prepared from cerebral ganglia mRNA as a positive probe and cDNA prepared from mRNA derived from CNS minus the cerebral ganglia and from the digestive gland as negative probes. In this way we isolated 150 clones out of 20,000 that specifically hybridized with cerebral ganglia cDNA. Since the cerebral ganglia also contain peptidergic neurons other than the CDCs, 20 out of these 150 clones were individually screened with a CDCH specific oligonucleotide mixture in the second step of the screening strategy. This screening resulted in the isolation of 7 positive clones, designated as CDCH clones. Restriction analysis showed that these CDCH clones contained inserts varying from 300 to 360 base pairs (bp). Subsequent analysis revealed that they contained a stretch of 108 nucleotides encoding the complete sequence of the 36 amino acids of CDCH (cf. Ebberink et al., 1985), with only one difference (see below). However, no AUG start codon was found upstream of the CDCH reading frame, suggesting that we were dealing with a partial cDNA clone. Moreover, RNA blot analysis (see below) confirmed that these cDNA clones were incomplete, since a CDCH positive transcript was found having a length of 1150 bases. Rescreening of the CNS cDNA library did not result in obtaining clones with longer inserts. We concluded that incomplete methylation of an Eco RI site in the CDCH cDNA clone had occurred during the construction of the CNS cDNA library, because all CDCH cDNA clones contained the same 5'-end, and differences in the size of the inserts were due to different lengths of the poly A tails only. Moreover, no sequences were found at the 5'-end of the inserts corresponding to the 12 nucleotides of the Eco RI linker used in the construction of the CNS cDNA library. We therefore constructed
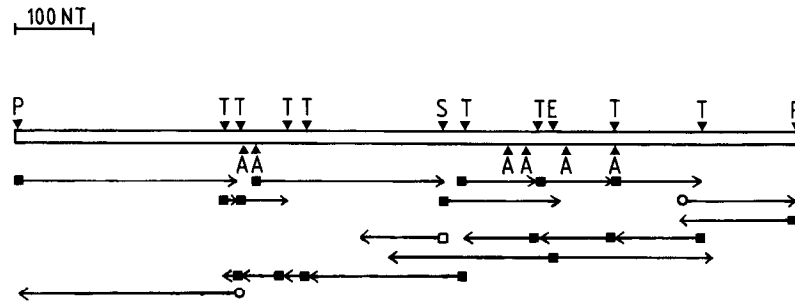
Figure 1. Partial restriction map and sequencing strategy for the CDCH-positive cDNA clone (pLCA). The DNA sequence was determined from the indicated restriction enzyme sites using the dideoxy method of Sanger et al. (1977) (*closed squares*) or by the chemical modification and cleavage procedure described by Maxam and Gilbert (1980) (*open squares*). The 5'- and 3'-ends were sequenced using synthetic oligonucleotides as primers in the dideoxy chain termination procedure (*open circles*). Arrows below the map represent the direction and extent of sequence determination. *A*, Sau 3A; *E*, Eco R1; *P*, Pst; *S*, Sal 1; *T*, Taq; *NT*, nucleotides.

a second cDNA library, however, from cerebral ganglia only. This cDNA library was screened with one of the CDCH clones isolated from the CNS cDNA library. Two positive clones were obtained having inserts of 1050 and 875 bp. We designated these clones as pLCA and pLCB, respectively.

Primary structure of the CDCH mRNA

The pLCA was completely sequenced according to the strategy outlined in Figure 1. Restriction analysis and partial sequence analysis showed that pLCB is identical with pLCA except for being shorter at its 5' end. The nucleotide sequence, comprising 1036 nucleotides, and the deduced amino acid sequence of pLCA are presented in Figure 2. RNA blot analysis (see below) revealed a CDCH positive transcript of 1150 bases, indicating that we cloned the full length cDNA assuming an average poly A tail of 100–150 nucleotides. The 5'-untranslated region shows a high A+T content. The coding region starts at the first ATG at position 188 and consists of 259 triplets; it contains many interspersed direct repeats, some of which correspond to repeated amino acid sequences (see below). The G+C content of the coding region is high (63%). The 3'-untranslated region is remarkably short (48 nucleotides) and is again rich in A+T residues. The sequence AATAAA, thought to be important in the process of poly A addition, is observed 17 nucleotides upstream of the poly A tail.

Primary structure of the CDCH preprohormone

The amino-terminal region of the CDCH precursor has the characteristics of a hydrophobic leader sequence. Taking into account the hydrophobicity plot of the CDCH precursor (not shown) and the fact that Ala is the most frequent amino acid preceding the cleavage site of leader sequences (Steiner et al., 1980), we predict that the leader sequence has a length of 34 amino acids and ends after Ala 34. Moreover, this proposed cleavage site fulfills the (-3,-1) rule for signal peptidase recognition (von Heyne, 1983). After removal of the proposed signal peptide a mature CDCH prohormone is generated with a calculated molecular weight of about 26,000 Da.

The 36 amino acids of CDCH are found near the carboxy terminus of the prohormone (residues 198 to 233) and are flanked on both sides by the dibasic sequence Lys-Arg, which is the most frequently used processing site in neuropeptide precursors (Steiner et al., 1980). The Gly-Lys-Arg sequence at the carboxy terminus of CDCH leads to amidation of CDCH, as this sequence is a common signal for processing and subsequent ami-

dation (Bradbury et al., 1982). The primary structure of CDCH is in good agreement with the peptide sequencing data (Ebberink et al., 1985). The only difference is a tryptophan residue at position 20 instead of the predicted lysine. Whether this discrepancy is due to an artifact or reflects polymorphism is at the present unclear.

In addition to the pairs of basic residues flanking CDCH, the CDCH prohormone contains 9 other di- or tribasic residues which may serve as cleavage sites. If all these sites are recognized, 12 peptides can be generated from the CDCH prohormone. Of these, 4 peptides are of particular interest since they exhibit a remarkable homology, corresponding with a sequence repetition within the preprohormone gene. Three of the predicted peptides are directly adjacent to each other (residues 103–125), whereas 4 pairs of Lys-Arg residues are flanking the pentapeptide sequences Arg-Leu-Arg-Phe-Asn, Arg-Leu-Arg-Ala-Ser, and Arg-Leu-Arg-Phe-His, respectively. This latter sequence is repeated exactly in the middle of a predicted 11-amino acid-long peptide (residues 142–152). CDCH and the 4 predicted peptides described above all have a basic character; however, the regions connecting these peptide domains have a strong acidic character. This implies that the CDCH prohormone comprises alternating negatively and positively charged regions. This structural organization may be important for tertiary folding, and hence for the correct processing of the preprohormone.

Expression of the CDCH gene

We characterized the nature of the CDCH RNA transcripts present in the cerebral ganglia by RNA blot analysis and *in situ* hybridization. A single abundant CDCH-positive transcript with a length of 1150 bases is present in the RNA fraction derived from the cerebral ganglia (see Fig. 3*A*). The cellular localization of this CDCH-positive transcript was determined by *in situ* hybridization using a ³⁵S-labeled Pst I insert from pLCA as a probe (see Fig. 3*B*). Hybridization was observed in all CDCs, and it was observed mainly in the cytoplasm. These results are in line with the immunocytochemical studies of Van Minnen and Vreugdenhil (1987) indicating that the CDCs express the CDCH gene.

Discussion

We have cloned and characterized a cDNA encoding a protein precursor that is cleaved to generate a set of small neuropeptides which are thought to govern egg-laying behavior in *Lymnaea*. The CDCH prohormone contains 11 potential cleavage sites

```
                                        1
                   ACACAGCCGGGCAGTGAGATTACGTGT

                                             100
GCGTTGTCAACGACGATTACGAAACGGGACAACCAGTTCACAGGAAGGACCACTCAGACTTCTGGTCTTCTCGCACCTTT

AGAGGCCTCTTTATCTTTATTACAAGCAAAACCAAATAAACACATTCTTTAAATAAGAATATTAACAACAGATTCAAAAG

              200
ATG'AAG'ATG'TCA'GGG'CTT'CTC'AGC'AAA'CCC'GAT'TAT'GGC'GTT'GTC'GGC'ATT'GTC'TTC'ACC'
Met-Lys-Met-Ser-Gly-Leu-Leu-Ser-Lys-Pro-Asp-Tyr-Gly-Val-Val-Gly-Ile-Val-Phe-Thr-
1                                                                              20

                                             300
GTC'GTC'TTC'TGC'TGT'TGG'TGT'TCC'TCC'TCG'ACG'ACG'CAC'GCC'CTG'TCG'ATC'GCG'GAG'CCA'
Val-Val-Phe-Cys-Cys-Trp-Cys-Ser-Ser-Ser-Thr-Thr-His-Ala-Leu-Ser-Ile-Ala-Glu-Pro-
                                                                               40

GGG'AGG'GAT'CGG'TAC'GAC'AAG'AGG'TCC'CCC'ACG'GGC'CAC'GGT'GTC'GAG'GTT'GTG'GAG'TCT'
Gly-Arg-Asp-Arg-Tyr-Asp-Lys-Arg-Ser-Pro-Thr-Gly-His-Gly-Val-Glu-Val-Val-Glu-Ser-
                                                                               60

                           400
GGC'GAG'GAC'TAC'GGG'TCG'AAC'AGA'CCC'CAA'CCC'GTT'TAC'GGC'GAT'GAG'GAC'GAA'GAA'GAC'
Gly-Glu-Asp-Tyr-Gly-Ser-Asn-Arg-Pro-Gln-Pro-Val-Tyr-Gly-Asp-Glu-Asp-Glu-Glu-Asp-
                                                                               80

TCC'GCG'GAC'GTG'TAC'GTC'GGC'TCC'GAC'GAA'AGC'TCC'AGC'GGC'GAA'AAA'ACG'AGG'CTC'ACG'
Ser-Ala-Asp-Val-Tyr-Val-Gly-Ser-Asp-Glu-Ser-Ser-Ser-Gly-Glu-Lys-Thr-Arg-Leu-Thr-
                                                                               100

              500
GCC'GCC'AAG'AGG'CGC'CTC'CGC'CTC'AAC'AAG'AGG'CGT'CTC'CGG'GCC'AGC'AAG'AGG'AGG'CTC'
Ala-Ala-Lys-Arg-Arg-Leu-Arg-Phe-Asn-Lys-Arg-Arg-Leu-Arg-Ala-Ser-Lys-Arg-Arg-Leu-
                                                                               120

                                    600
AGG'TTC'CAC'AAG'AGG'AGA'GTC'GAC'TCC'GCC'GAC'GAA'TCG'AAC'GAC'GAC'GGC'TTT'GAC'CGA'
Arg-Phe-His-Lys-Arg-Arg-Val-Asp-Ser-Ala-Asp-Glu-Ser-Asn-Asp-Asp-Gly-Phe-Asp-Arg-
                                                                               140

AAA'GCG'CGC'GAA'CCG'AGA'CTG'AGG'TTC'CAT'GAC'GTC'AGA'AAG'AGA'TCC'GCC'ACG'GCC'GAA'
Lys-Ala-Arg-Glu-Pro-Arg-Leu-Arg-Phe-His-Asp-Val-Arg-Lys-Arg-Ser-Ala-Thr-Ala-Glu-
                                                                               160

                        700
GAG'GGA'TCC'GAA'AAC'GCG'GAA'ATC'GAA'GAG'TCC'CAT'CTC'GGG'AAT'TCC'CGA'AGT'CGC'AGA'
Glu-Gly-Ser-Glu-Asn-Ala-Glu-Ile-Glu-Glu-Ser-His-Leu-Gly-Asn-Ser-Arg-Ser-Arg-Arg-
                                                                               180

TCG'GCC'GGC'TCC'GCA'CCG'AGC'TCG'GCT'AAC'GAG'GTC'CAA'CGT'TCC'AAA'CGA'CTT'TCG'ATC'
Ser-Ala-Gly-Ser-Ala-Pro-Ser-Ser-Ala-Asn-Glu-Val-Gln-Arg-Ser-Lys-Arg-Leu-Ser-Ile-
                                                                               200

              800
ACC'AAT'GAC'CTG'CGG'GCT'ATC'GCT'GAC'AGT'TAC'CTG'TAC'GAC'CAG'CAC'AAG'CTG'AGA'GAG'
Thr-Asn-Asp-Leu-Arg-Ala-Ile-Ala-Asp-Ser-Tyr-Leu-Tyr-Asp-Gln-His-Lys-Leu-Arg-Glu-
                                                                               220

                                             900
CGG'CAA'GAA'GAG'AAC'CTA'AGA'AGA'CGT'TTC'CTA'GAG'CTC'GGC'AAG'AGA'GGC'TCC'GCG'TTC'
Arg-Gln-Glu-Glu-Asn-Leu-Arg-Arg-Arg-Phe-Leu-Glu-Leu-Gly-Lys-Arg-Gly-Ser-Ala-Phe-
                                                                               240

TTC'GAC'CAT'ATC'CCG'ATA'ATC'TTT'GGG'GAA'CCT'CAG'TAT'GAT'TAC'CAG'CCT'TTT'AAA'TAA'
Phe-Asp-His-Ile-Pro-Ile-Ile-Phe-Gly-Glu-Pro-Gln-Tyr-Asp-Tyr-Gln-Pro-Phe-Lys-STOP
                                                                               259

                          1000
AATTATTTGCCTTAAAGTTTTGAAATAAAACCATAATTGGTCTCTTAAAAAAAAAAAAAAAAAAAAAAAA
```

Figure 2. Nucleotide and deduced amino acid sequence of the CDCH preprohormone mRNA. Nucleotide residues are numbered in the 5' to 3' direction above the nucleotide sequence. The putative initiator methionine codon is present at position 188 and is followed by a 777-nucleotide-long open reading frame. The amino acid residues are numbered below the amino acid sequence, beginning with the initiator methionine. The *dashed line* indicates the predicted primary sequence of CDCH. Basic residues that form potential cleavage sites are underlined with a solid line. The nucleotides corresponding to the polyadenylation consensus sequence are also underlined with a solid line.

A

B
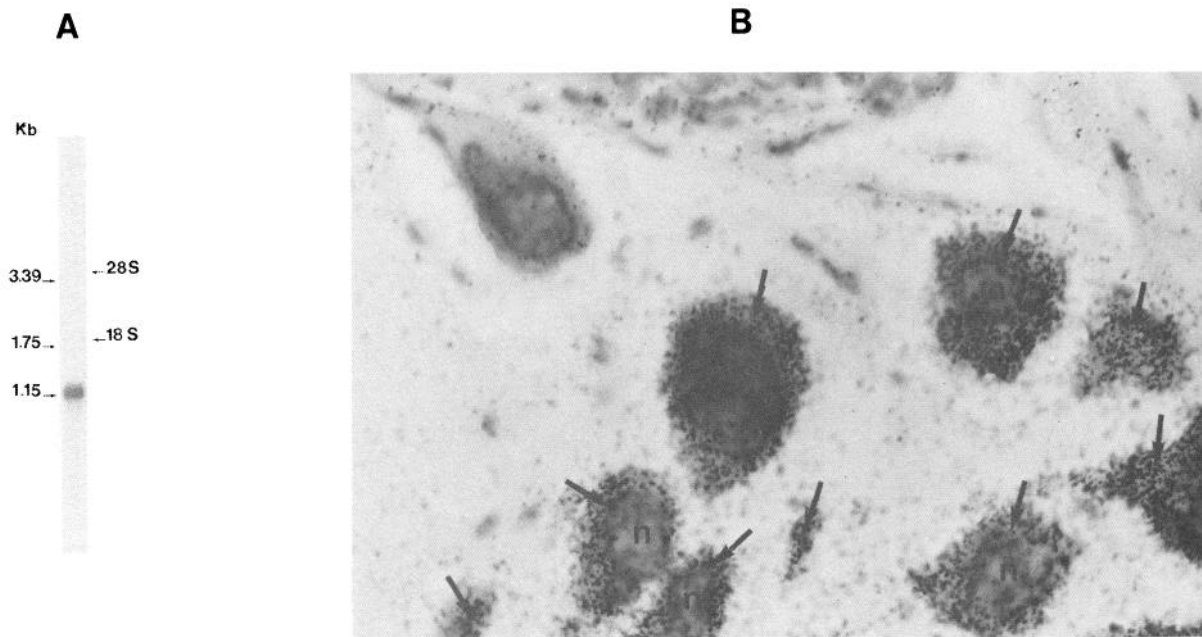
Figure 3. Expression of the CDCH gene in the cerebral ganglia. *A,* RNA blot analysis of the cerebral ganglia. Five micrograms of total RNA from the cerebral ganglia were fractionated on a 1.2% agarose gel and transferred to Hybond filter. The ^{32}P nick-translated Pst I insert of clone pLCA was used as a probe in the hybridization analysis. The positions of the 18S and 28S ribosomal RNAs are indicated. Yeast (*Saccharomyces carlsbergensis*) 17S rRNA (1750 bases) and 26S (3390 bases) rRNA were used as markers. *B, In situ* hybridization to sections of the cerebral ganglia. A ^{35}S-nick-translated Pst I fragment from pLCA was applied to 7 μm longitudinal sections of the cerebral ganglia. Slides were dipped in Kodak NTB 2 emulsion, exposed for 24 hr, and examined by light microscopy. *Arrows* indicate hybridization in the cytoplasm of the CDCs. *N,* nucleus. ×500.

(see Fig. 4*A*). If all these sites are recognized 12 peptides can be generated from the CDCH prohormone, a number close to that (10) released by electrically active CDCs (Geraerts et al., 1987b). Several experimental data fit in well with the hypothesis that the combined biological actions of this released set of peptides are necessary for the egg-laying action pattern in *Lymnaea* to occur. First, in contrast to similar experiments done with *A. californica* (Strumwasser, 1984; Strumwasser et al., 1987), injection of synthetic CDCH can elicit only the last part of the overt egg-laying behavior (Ter Maat et al., 1988b). Second, besides CDCH, other released peptides with a known function have been identified on the CDCH precursor, i.e., calfluxin and CDCA. The amino acid composition of purified calfluxin is identical with a 14-amino acid-long peptide that is encoded by the CDCH cDNA clone (Fig. 2, residues 126–139); in addition, a synthetic peptide corresponding to this region mimics the biological activity of native calfluxin, suggesting that (released)

calfluxin is synthesized as part of the CDCH prohormone (Dictus and Ebberink, 1988). The second identified peptide, CDCA, triggers the electrical activity of the CDCs (Ter Maat et al., 1988a). A synthetic nonapeptide corresponding to a 9-amino acid-long peptide, α-CDCP, encoded by the CDCH cDNA clone (Fig. 2, residues 144–152), is capable of exciting isolated CDCs at a concentration of 10^{-5} M (A.B. Brussaard, personal communication), indicating a structural relationship between CDCA and the encoded nonapeptide. We assume that the other peptides encoded by the CDCH cDNA clone are also released and that they are involved in the control of different aspects of the egg-laying fixed action pattern.

Comparison of the CDCH preprohormone of L. stagnalis *with the ELH preprohormones of* A. californica *and* A. parvula

The freshwater Lymnaeidae, having evolved from marine snails through land-based intermediates, are only distantly related to
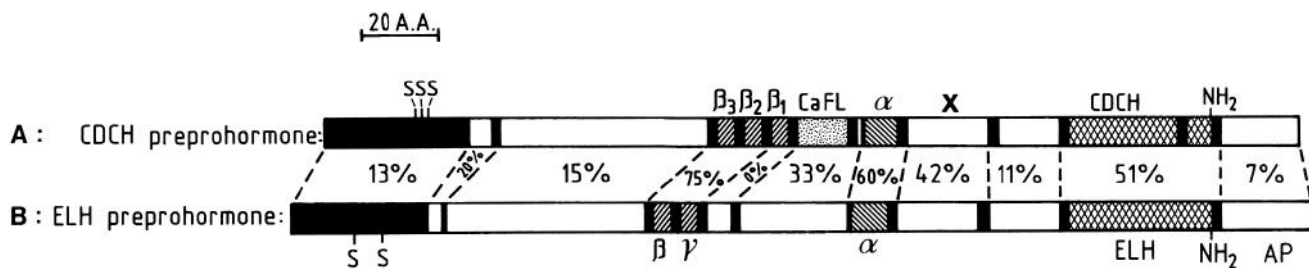
Figure 4. Comparison of the CDCH (*A*) and ELH (*B*) preprohormone. An S indicates the position of a cysteine residue. NH$_2$ represents a potential amidation signal. Percentages homology between corresponding regions are indicated. Gaps are introduced to achieve maximum homology. Furthermore, the positions of potential signal peptides are indicated (■), CDCH (▨), α-CDCP/α-BCP (▨), β_{1-3}-CDCP/β-BCP/γ-BCP (▨), calfluxin (CaFl, ▨), X-CDCP (*X*), and acidic peptide (AP). Known or potential cleavage sites are also shown (▮).

```
α-CDCP       :   RKAREP RLRF HDV RKR
α-BCP        :      RAP RLRF YSL RKR

β₁-CDCP      :       KR RLRF HKR
β-BCP        :       KR RLRF HKR

β₂-CDCP      :       KR RLR ASKR
γ-BCP        :       KR RLR FDRR

β₃-CDCP      :       KR RLRF NKR


CaF1         :   KRR------VDSADE------SNDDGFD--RK
CaF1(like)   :   RRFPTNAVSMSADENSPFDLSNEDGAVYQR

X-CDCP       :   KRSATAEEGSENAEIE-ESHLGNSRSRR
X-BCP        :   KRAAGGMEQSEGQNPETESHGRRKRS
```

Figure 5. Primary structure and the potential or known cleavage sites of homologous peptides of the CDCH and ELH preprohormones. The standard 1-letter code for amino acids is used. Identical amino acids are shown in *bold type* and homologous repetitive amino acid sequences are enclosed within a *box.* Potential or known cleavage sites are *underlined.* Gaps are introduced to achieve maximum homology.

the marine Aplysiidae. The evolutionary distance between these families has been estimated to be 350 million years (Moore and Pitrat, 1960). Nevertheless, beside certain differences, striking similarities can also be seen in the egg-laying behavior of these animals (Geraerts et al., 1987a). It is therefore interesting to compare the CDCH preprohormone with the ELH preprohormones. Because the homology between both ELH precursors of *A. californica* and *A. parvula* is high (Nambu and Scheller, 1986), we will limit ourselves to a comparison of the CDCH preprohormone and the ELH preprohormone of *A. californica.*

The overall homology between the CDCH preprohormone and the ELH preprohormone of *A. californica* is very low (28.2%) compared with homology figures reported for other neuropeptide precursors (e.g., Soma et al., 1984; Martens et al., 1985). Nevertheless, the structural organization of both preprohormones is very similar (see Fig. 4). They have long signal peptides as predicted, and nearly all the potential cleavage sites are conserved. A more detailed comparison gives a striking picture of homologous and divergent regions. In 3 regions significant homology is evident. The first region contains CDCH, which exhibits 51% homology with ELH. Peptide sequencing data have demonstrated a 44% homology between CDCH and ELH (Chiu et al., 1979; Ebberink et al., 1985). Our data demonstrate that, in addition, the processing sites (Lys-Arg) are completely conserved and that CDCH, like ELH, is amidated at its carboxy terminal because the sequence Gly-Lys-Arg serves as a combined proteolysis/amidation signal. The second region of interest contains α-CDCP, which exhibits 69% homology with its counterpart on the ELH preprohormone, α-BCP. α-CDCP has 5 uninterrupted amino acids in common with the 9-amino acid-long α-BCP (see Fig. 5). The cleavage sites flanking α-BCP are also completely conserved in the CDCH preprohormone. The third region contains 3 predicted pentapeptides, β_{1-3}-CDCP, which exhibit a high degree of homology among themselves and also with α-CDCP and α-BCP, and 75% homology with their counterparts on the ELH precursor, β-BCP and γ-BCP (see Fig. 5). As in the ELH preprohormone, the β_{1-3}-CDCP region contains related pentapeptides that are all flanked by dibasic amino acid residues. Our data also show the generation of an extra predicted pentapeptide in the CDCH preprohormone.

The significant homology found in the 3 regions suggests that the neuropeptides generated from them control similar aspects in the egg-laying fixed action patterns of both species. Indeed,

the well-established biological actions of CDCH and ELH (induction of ovulation, egg-mass and egg string formation) support this hypothesis. It is also likely that α-CDCP and α-BCP perform similar functions. The precise processing of α-CDCP is not known as yet. The conservation of the cleavage sites suggests that α-CDCP, as α-BCP (Rothman et al., 1983), is liberated from its precursor as a nonapeptide. However, the presence of an additional potential cleavage site (Arg-Lys) at the amino-terminal site of the α-CDCP region (see Figs. 4, 5) might give rise to a form of α-CDCP that contains 2 extra amino acids. Although the processing and function of peptides generated from the third homologous region are largely unknown, we assume that they play an important role in egg laying because of the following. (1) This region contains predicted pentapeptides with a clear repetitive character (see Fig. 5). (2) The third region exhibits the highest level of homology and contains a predicted peptide domain (β-BCP/β_1-CDCP; see Fig. 5) with even 100% homology. The related pentapeptides, β_2- and β_3-CDCP, have diverged at only 2 and 1 positions, respectively, compared with their counterpart, γ-BCP. (3) The CDCH preprohormone contains an extra pentapeptide. (4) All pentapeptide domains are flanked by Lys-Arg (except for γ-BCP), which strongly suggests that these peptides are liberated from their precursors as pentapeptides. The pentapeptides may have a role as a neurotransmitter, as proposed for β-BCP (Scheller et al., 1983; Rothman et al., 1985).

By introducing gaps in the peptide sequence of the CDCH preprohormone, significant homology was found in 2 additional regions (see Figs. 4, 5). In the region containing the predicted 14 amino acid peptide connecting α-CDCP with the β_{1-3}-CDCP region, 2 stretches of 4 and 5 amino acids are 100 and 80% homologous, respectively, with the 27 amino acid counterpart on the ELH preprohormone (see Fig. 5). In *Lymnaea,* this predicted 14 amino acid peptide most probably represents the primary structure of calfluxin (Dictus and Ebberink, 1988). The homology suggests a similar biological action as the calfluxin-related peptide in *Aplysia.* Indeed, preliminary experiments indicate that bag cell extracts are capable of inducing calcium movements in the accessory genital mass of *A. californica* (M. de Jong-Brink, personal communication). A second region with significant homology was found directly beyond the carboxy-terminal processing site of α-CDCP and α-BCP. The 23 amino acid peptide, X-CDCP, of the CDCH preprohormone shares 8

amino acid residues with its 19 amino acid counterpart on the ELH preprohormone (see Fig. 5), suggesting that these peptides also perform similar, but as yet unknown, functions.

No significant homology could be identified in other regions of the CDCH preprohormone. This is especially true for the amino-terminal part of the preprohormone. Also the 27-amino acid-long acidic peptide of the ELH preprohormone of *A. californica*, thought to play a role as a carrier protein (Scheller et al., 1983), has completely diverged to a less acidic *L. stagnalis* cognate of 23 amino acids. The divergence indicates that these predicted peptides have no physiological function (Soma et al., 1983) and might be involved in the folding and/or correct processing of the preprohormones.

The ELH gene of *A. californica* is a member of a small multigene family consisting of 4–5 distinct members (Mahon et al., 1985). In addition to the ELH gene, members of the family encode the A peptide- and B peptide-preprohormones (Scheller et al., 1983; Mahon et al., 1985). In *A. parvula*, the ELH family consists of only 2 genes, both of which encode ELH-like preprohormones (Nambu and Scheller, 1986). It is thought that *A. parvula* is more primitive than *A. californica*, and that the A peptide and B peptide genes arose less than 140 million years ago (Nambu and Scheller, 1986). Preliminary results of Southern blot analysis suggest the presence of a small multigene family in *L. stagnalis* (E. Vreugdenhil, unpublished observations). Considering the evolution of *L. stagnalis*, it will be interesting to characterize the genomic organization of this CDCH gene family. This characterization will give further insight into the evolution of peptide transmitters in general and will allow a better understanding of how egg laying and its accompanying behavior in *L. stagnalis* is controlled at the DNA level in particular.

References

Bradbury, A. F., M. D. A. Finnie, and D. G. Smyth (1982) Mechanism of C-terminal amide formation by the pituitary enzymes. Nature *298:* 686–688.

Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry *18:* 5294.

Chiu, A. Y., M. W. Hunkapillar, E. Heller, D. K. Stuart, L. E. Hood, and F. Strumwasser (1979) Neuropeptide egg-laying hormone of *Aplysia californica*: Purification and primary structure. Proc. Natl. Acad. Sci. USA *76:* 6656–6660.

Cobbs, J. S., and H. M. Pinsker (1982) Role of bag cells in egg deposition of *Aplysia brasiliana*. I. Comparison of normal anad elicited behaviors. J. Comp. Physiol. *147A:* 523–535.

Dictus, W. J. A. G., and R. H. M. Ebberink (1988) Purification and amino acid sequence of a novel neuropeptide (calfluxin), one of the peptides encoded by the egg-laying hormone precursor of *Lymnaea*. Mol. Cell. Endocrinol. (in press).

Dictus, W. J. A. G., M. de Jong-Brink, and H. H. Boer (1987) A neuropeptide (calfluxin) is involved in the influx of calcium into mitochondria of the albumen gland of the freshwater snail *Lymnaea stagnalis*. Gen. Comp. Endocrinol. *65:* 439–450.

Ebberink, R. H. M., H. Van Loenhout, W. P. M. Geraerts, and J. Joosse (1985) Purification and amino acid sequence of the ovulation hormone of *Lymnaea stagnalis*. Proc. Natl. Acad. Sci. USA *82:* 7767–7771.

Geraerts, W. P. M., C. P. Tensen, and T. M. Hogenes (1983) Multiple release of peptides by electrically active neurosecretory caudodorsal cells of *Lymnaea stagnalis*. Neurosci. Lett. *41:* 151–155.

Geraerts, W. P. M., P. Buma, and T. M. Hogenes (1984) Isolation of neurosecretory granules from the neurohaemal areas of peptidergic systems of *Lymnaea stagnalis*, with special reference to the ovulation hormone producing caudo-dorsal cells. Gen. Comp. Endocrinol. *52:* 212–217.

Geraerts, W. P. M., E. Vreugdenhil, R. H. M. Ebberink, and T. M. Hogenes (1985) Synthesis of multiple peptides from a larger precursor in the neuroendocrine caudodorsal cells of *Lymnaea stagnalis*. Neurosci. Lett. *56:* 241–246.

Geraerts, W. P. M., A. Ter Maat, and E. Vreugdenhil (1987a) The peptidergic neuroendocrine control of egg-laying behavior in *Aplysia* and *Lymnaea*. In *Invertebrate Endocrinology*, Vol. 2, H. Laufer and R. Downer, eds., pp. 377–468, Alan Liss, New York.

Geraerts, W. P. M., E. Vreugdenhil, and R. H. M. Ebberink (1987b) Bioactive peptides in molluscs. In *Invertebrate Peptide Hormones*, M. C. Thorndyke and G. Goldsworthy, eds., Cambridge U. P., Cambridge, UK (in press).

Goldschmeding, J. T., M. Wilbrink, and A. Ter Maat (1983) The role of the ovulation hormone in the control of egg-laying behaviour in *Lymnaea stagnalis*. In *Molluscan Neuro-Endocrinology*, J. Lever and H. H. Boer, eds., pp. 251–255, North-Holland, Amsterdam.

Huynh, T. V., R. A. Young, and R. W. Davis (1985) Construction and screening of cDNA libraries in λgt10 and λgt11. In *DNA Cloning*, Vol. 1, D. Glover, ed., IRL Press, Oxford, UK.

Jackson, J. F., E. Vreugdenhil, K. J. Johnson, and F. H. Hall (1986) Genes encoding molluscan gonadotropic hormones. Biochem. Soc. Symp. *52:* 83–90.

Jansen, R. F., and A. Ter Maat (1985) Ring neuron control of columellar motor neurons during egg-laying behavior in the pond snail. J. Neurobiol. *16:* 1–14.

Kits, K. S. (1980) States of excitability in the ovulation hormone producing neuroendocrine cells of *Lymnaea stagnalis* (Gastropoda) and their relation to the egg-laying cycle. J. Neurobiol. *11:* 397–410.

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 family in *Aplysia*. J. Neurosci. *5:* 1872–1880.

Maniatis, T., E. G. Fritsch, and J. Sambrook (1983) *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Martens, G. J. M., O. Civelli, and E. Herbert (1985) Nucleotide sequence of cloned cDNA for pro-opiomelanocortin in the amphibian *Xenopus laevis*. J. Biol. Chem. *260:* 13685–13689.

Maxam, A. M., and W. Gilbert (1980) Sequencing end labeled DNA with base specific chemical cleavages. Methods Enzymol. *65:* 499–560.

Moore, R. C., and C. W. Pitrat (1960) *Treatise on Invertebrate Paleontology, Part 2, Mollusca 1*, Geological Society of America Inc./University of Kansas Press.

Nambu, J. R., and R. H. Scheller (1986) Egg-laying hormone genes of *Aplysia*: Evolution of the ELH gene family. J. Neurosci. *6:* 2026–2036.

Rothman, B. S., E. Mayeri, R. O. Brown, P. M. Yuan, and J. E. Shively (1983) Primary structure and neuronal effects of α-bag cell peptide, a second candidate neurotransmitter encoded by a single gene in bag cell neurons of *Aplysia*. Proc. Natl. Acad. Sci. USA *80:* 5753–5757.

Rothman, B. S., R. H. Scheller, and E. Mayeri (1985) The bag cell neurons of *Aplysia* as a possible peptidergic multi-transmitter system: From genes to behavior. In *Gene Expression in Brain*, C. Zomzely-Neurath and W. A. Walker, eds., pp. 235–274, Wiley, New York.

Roubos, E. W. (1984) Cytobiology of the ovulation neurohormone producing caudodorsal cells of the snail *Lymnaea stagnalis*. Int. Rev. Cytol. *89:* 295–346.

Sanger, F., S. Nicklen, and A. R. Coulson (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74:* 5463–5467.

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.

Soma, G., N. Kitahara, T. Nishizawa, H. Nanami, C. Kotake, H. Okazaki, and T. Andoh (1984) Nucleotide sequence of a cloned cDNA for pro-opiomelanocortin of chum salmon. Nucleic Acids Res. *12:* 8029–8041.

Steiner, D. F., P. S. Quinn, J. C. Shu, J. March, and J. H. S. Tager (1980) Processing mechanisms in the biosynthesis of proteins. Ann. NY Acad. Sci. *343:* 1–16.

Strumwasser, F. (1984) The structure of the commands for a neuropeptide-mediated behavior, egg laying, in an opisthobranch mollusc. In *Biosynthesis, Metabolism and Mode of Action of Invertebrate Hormones*, J. Hoffmann and M. Porchet, eds., pp. 36–43, Springer, Heidelberg.

Strumwasser, F., D. L. Schiller, and S. B. H. Kent (1987) Synthetic neuropeptide egg-laying hormone (ELH) of *Aplysia californica* induces normal egg-laying. Structure-activity studies. Soc. Neurosci. Abstr. *13:* 38.

Ter Maat, A., J. C. Lodder, and M. Wilbrink (1983) Induction of egg laying in the pond snail *Lymnaea stagnalis* by environmental stimulation of the release of ovulation hormone from the caudo-dorsal cells. Int. J. Invert. Reprod. *6:* 239–247.

Ter Maat, A., W. P. M. Geraerts, R. F. Jansen, and N. P. A. Bos (1988a) Chemically mediated positive feedback generates long lasting after discharge in a molluscan neuroendocrine system. Brain Res. *438:* 77–82.

Ter Maat, A., A. W. Pieneman, J. T. Goldschmeding, T. Smelik, and G. P. Ferguson (1988b) Spontaneous and induced egg-laying behavior of the pond snail *Lymnaea.* J. Comp. Physiol. (in press).

Thomas, P. S. (1980) Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA *77:* 5201–5205.

Van Minnen, J., and E. Vreugdenhil (1987) The occurrence of gonadotropic hormones in the CNS and reproductive tract of *Lymnaea stagnalis*: An immunocytochemical and *in situ* hybridization study.

In *Neurobiology: Molluscan Models,* Mon. Royal Netherlands Academy of Arts and Sciences, H. H. Boer, W. P. M. Geraerts, and J. Joosse, eds., pp. 62–68, North-Holland, New York.

von Heyne, G. (1983) Patterns of amino acids near signal sequence cleavage sites. Eur. J. Biochem. *1333:* 17–21.

Vreugdenhil, E., W. P. M. Geraerts, J. F. Jackson, and J. Joosse (1985) The molecular basis of the neuroendocrine control of egg-laying behaviour in *Lymnaea stagnalis.* Peptides *6:* 465–470.

Watson, C. J., and J. F. Jackson (1985) An alternative method for the synthesis of double stranded cDNA for cloning in phage and plasmid vectors. In *DNA Cloning,* Vol. 1, D. Glover, ed., IRL Press, Oxford, UK.

Wijdenes, J., R. van Elk, and J. Joosse (1983) Effects of two gonadotropic hormones on polysaccharide synthesis in the albumen gland of *Lymnaea stagnalis* studied with the organ culture technique. Gen. Comp. Endocrinol. *51:* 263–271.