FEBS 21320 FEBS Letters 441 (1998) 397–403

The shrimp hyperglycemic hormone-like neuropeptide is encoded by multiple copies of genes arranged in a cluster

P.-L. Gu, S.-M. Chan*

Department of Zoology, The University of Hong Kong, Pokfulam Road, Hong Kong, China

Received 16 October 1998; received in revised form 23 November 1998

Abstract The crustacean hyperglycemic hormone (CHH) plays an important role in the regulation of glucose metabolism. We have cloned and sequenced several cDNAs encoding the preproCHH-like of the shrimp, *Metapenaeus ensis*. The preproCHH-like peptide of the shrimp consists of a signal peptide, a CHH precursor-like peptide (CPRP) and the CHH-like peptide. Comparative analysis revealed that the signal peptide and the CPRP of the shrimp peptide are the shortest among all the CHHs reported. MeCHH-like is expressed in the eyestalk, but it is not expressed in the heart, hepatopancreas, muscle, nerve cord and pre-hatch embryo. To study the structural organization of the shrimp CHH-like gene, we have screened the genomic DNA library constructed from one shrimp. Three groups of overlapping genomic clones have been isolated. The results from both genomic Southern blot analysis and library screening indicate that the shrimp genome contains at least six copies of the CHH-like genes arranged in a cluster on the chromosome. The size of the CHH-like genes is 1.5–2.1 kb. DNA sequence determinations indicate that the CHH-like genes share 98–100% amino acid sequence identity. There are three exons and two introns in each CHH-like gene. The first intron separates the signal peptide and the second intron separates the mature peptide in the coding region. The 150–200 bp of the upstream 5′ flanking region of the CHH-like genes contains promoters with characteristics similar to most eukaryotic genes. Several putative *cis*-acting elements are also identified in the first 400 bp 5′ end upstream region. The organization of the shrimp CHH-like genes is similar to that of the molt inhibiting hormone gene of the same shrimp and the crab, *Charybdis feriatus*.

Key words: Crustacean hyperglycemic hormone gene; Gene cluster; Shrimp

1. Introduction

The medulla terminalis X-organ (MTXO) of the crustacean eyestalk synthesizes the crustacean hyperglycemic hormone (CHH) which regulates glucose metabolism [1]. The role that CHH plays in the regulation of glucose metabolism has been studied in a number of crustaceans. Because of the relative abundance of CHH and the simplicity of using bioassays to identify this neuropeptide, the isolation, purification, amino acid sequence determination and bioassay analysis for the hyperglycemic activity have been reported for many species [2,3]. The amino acid sequence of CHHs for several crustaceans has been reported [4–8]. Like other neurohormones, neuropeptides of the CHH family all possess 72–74 amino acids and six cysteine residues in identical position [1]. Although there is a high degree of similarity in the amino acid sequence of the CHHs of similar crustaceans, the nucleotide and amino acid sequences among different crustaceans display only 40–60% identity. Research in this area is complicated by the presence of more than one form of CHH. The multiple isoforms or different CHHs commonly observed in single species of crustacean suggest that these neurohormones may be encoded by several genes. Although extensive studies on the CHH have been performed, there is no information on the structural organization of the CHH gene in the shrimp. This information is important for our understanding of the mechanism that regulates the expression of these genes and may provide insight into the evolution of this group of important neuropeptides. This paper reports the isolation and characterization of a cDNA encoding the putative CHHs of the shrimp, *Metapenaeus ensis*. We have also extended our study to the structural organization of the shrimp CHH-like gene and provide evidence that the shrimp CHH-like neuropeptide is encoded by multiple copies of genes arranged in a cluster on the same chromosome segment.

2. Materials and methods

2.1. RNA preparation and RT-PCR

Various tissues were dissected and extracted for RNA [9] within 2 min after handling in order to avoid stress to the animals. The quality of the total RNA was monitored after running it on 1.5% denatured RNA gel. Intact RNA (1 μg) was used in the reverse transcription reaction for the synthesis of the first strand cDNA. Reverse transcription (RT) was performed in a transcription reaction buffer (50 mM Tris-HCl, 8 mM $MgCl_2$, 30 mM KCl, 2 mM each of dNTP, 10 mM DTT), 2 pmol of oligo(dT)17 primer, and 1 unit of reverse transcriptase (Promega, Madison, WI, USA). The reaction mixture was incubated at 42°C for 2 h. For RT-PCR, we designed forward (S1) and reverse primers (S2) based on the known CHH nucleotide sequence of the shrimp, *Penaeus japonicus* [10] (S1: AGTCTGTTCGACCCGTCGTGC and S2: CTACTTCCCAACCACCTGGAC). The final PCR mix (30 μl/reaction) consisted of 10 mM Tris-HCl, pH 8.0, 1.5 mM $MgCl_2$, 50 mM KCl, 0.5 pmol primer S1 and S2 and 2.5 μl reaction mix from the reverse transcription as described above. The PCR conditions consisted of 35 cycles each of denaturing at 95°C for 1 min, annealing at 60°C for 1 min and extension at 72°C for 1 min. At the end of the last cycle, the PCR mix was incubated at 72°C for another 10 min for the completion of DNA synthesis. PCR products were analyzed on 1.5–2.0% agarose gels to determine specific DNA amplification. Targeted DNA fragments of expected sizes (approximately 220 bp) were excised from the gel and DNA fragments were subcloned into the pBluescript cloning vector (Stratagene, La Jolla, CA, USA). DNA sequence determination on both strands was performed using a T7 DNA polymerase sequencing kit (Pharmacia, Sweden). DNA and amino acid sequences derived from the results were compared and analyzed with the GenBank database.

*Correspondent author. Fax: (852) 2857-4672.
E-mail: chansm@hkucc.hku.hk

The DNA sequences of the MeCHH-like cDNA and four (43-1, 43-2, 43-3 and 43-4) out of six CHH-like genes have been submitted to the GenBank sequence database with the accession numbers AF109775, AF109776, AF109777, AF109778 and AF109779 respectively.

PII: S0014-5793(98)01573-7

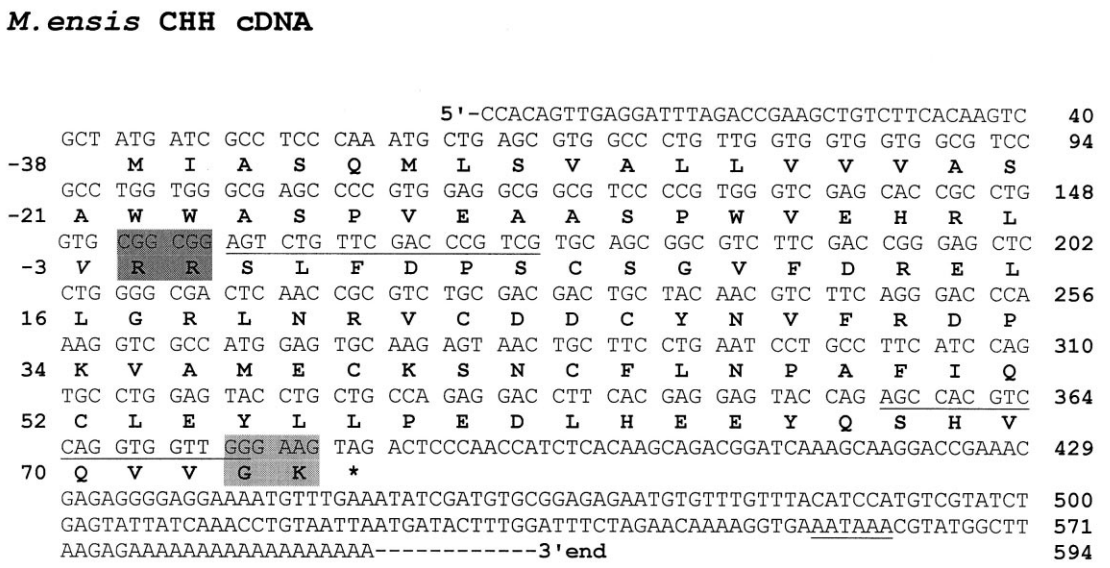

Fig. 1. Nucleotide and deduced amino acid sequence of the preproCHH-like cDNA of the shrimp *Metapenaeus ensis*. The putative polyadenylation signal (AATAAA) is in bold letters. The position of primers (S1, forward and S2, reverse) used in the initial RT-PCR is underlined. Numbers on the left and right margins represent amino acid and nucleotide positions respectively. The darker box indicates the processing site and the lighter box indicates the region for amidation site. The nucleotide sequence shown in bold italic indicates the position of the *Sst*I restriction enzyme site.

Northern blot analysis was used to determine tissue-specific expression of mRNA encoding CHH-like peptide. RNA isolated from different tissues was separated on 1.5% formaldehyde agarose gel and transferred onto a nitrocellulose membrane. The membrane contained hybridization buffer with a randomly primed cDNA probe of CHH-like from *M. ensis*. Hybridization was performed at 42°C overnight with a buffer containing 50% formamide. The final washing was performed at 60°C in 0.2×SSPE and 0.1% SDS. For the detection of *M. ensis* CHH-like mRNA, the PCR amplified cDNA probe was used. In both cases, high stringency (0.1×SSC and 0.1% SDS wash at 65°C) washes were performed to eliminate cross-hybridization of the probe to other neuropeptide genes.

2.2. cDNA library construction and screening

Using 5 μg of poly(A)-rich RNA, an eyestalk cDNA library was constructed in the vector λZap-II according to the instruction of the manufacturer (Stratagene). The cDNA library was screened for CHH using a gene specific probe derived from RT-PCR and the probe was labeled by [α-^{32}P]dCTP. For library screening, the probe was derived from the PCR subclones as mentioned before. Randomly primed DNA probe was synthesized and used for screening the shrimp eyestalk cDNA library. After the third round screening, potential positive phage plaques were purified and the recombinant pBK-CMV phagemids were rescued from the bacteriophage clones by in vivo excision according to the instruction of the manufacturer (Stratagene). DNA

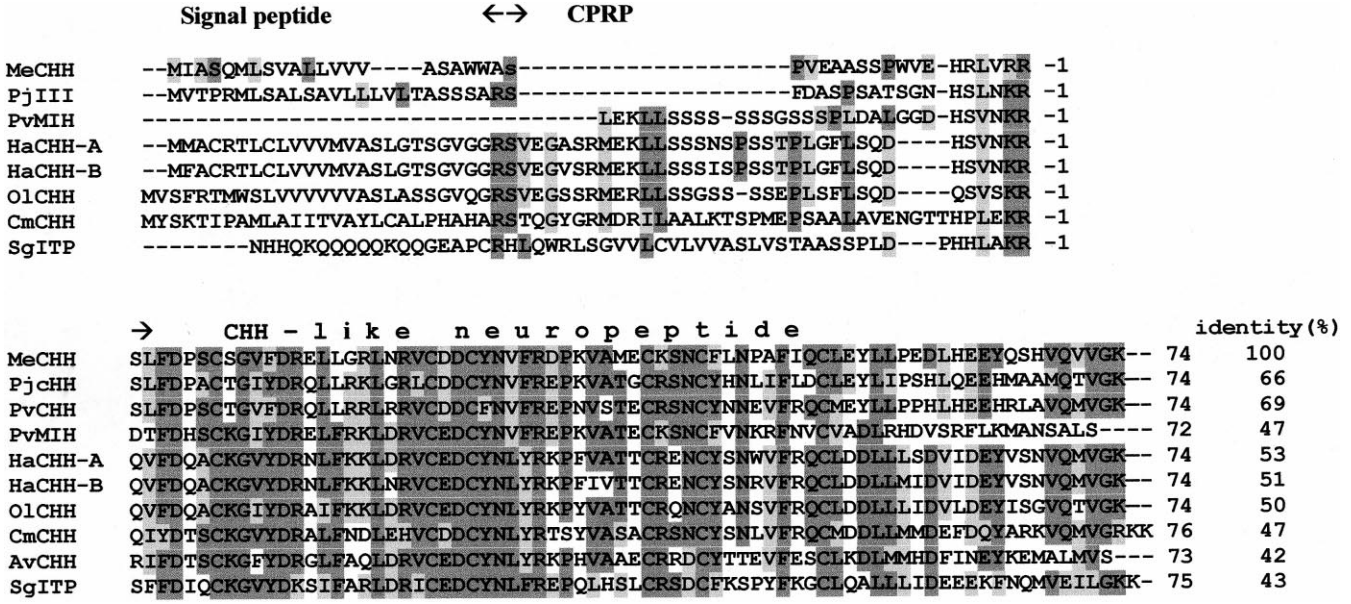

Fig. 2. Amino acid sequence comparison of the shrimp CHH-like peptides with other crustacean CHHs. Alignment of the CHH amino acid sequence for shrimp with other crustaceans and the ITP of an insect. The compared sequences included the CHH of the shrimp *Penaeus japonicus* [12] and MIH-like of *Penaeus vannamei* [22], the lobster *Homarus americanus* [14], the crab *Carcinus maenas* [8], the crayfish *Orconectes limonus* [17] and the ITP of the locust *Schistocerca gregaria* [15]. Identical amino acid residues are shaded in darker color and similar amino acids are shaded in light gray. Inserts (-) have been added to maximize sequence identity. The percentage identity indicates the overall amino acid identity for the mature peptide.

sequence determination was performed using a T7 DNA sequencing kit (Pharmacia).

2.3. Cloning of the shrimp CHH-like gene

A shrimp genomic DNA library was constructed using a genomic library construction kit (Promega). High molecular weight genomic DNA from one shrimp was prepared according to Maniatis et al. [11]. The procedure for library construction followed the manufacturer's instructions. The procedure for the genomic library screening was similar to that described for cDNA library screening. Potential positive clones were analyzed by Southern blot after restriction enzyme digestion and agarose gel electrophoresis. DNA fragments containing CHH-like genes were subcloned into a plasmid vector (pBluescript, Stratagene). The subclones were also subjected to DNA sequence determination to confirm the presence of CHH specific sequence.

3. Results

3.1. Cloning of shrimp preproCHH-like cDNA

RT-PCR product of expected size (220 bp) was subcloned into pBluescript KS vector (Stratagene) and the partial cDNA sequence (data not shown) was shown to have coding sequence homology to the shrimp, *P. japonicus* [11]. The partial cDNAs were used as probes to screen the shrimp eyestalk cDNA library. Three cDNA clones of identical coding sequence were obtained. The largest cDNA revealed an open reading frame of 336 nucleotides encoding a neuropeptide of 112 amino acids (Fig. 1). A putative polyadenylation signal (AATAAA) was found 80 bp 3′ from the translation termination codon (TGA). Hydropathicity analysis of the deduced neuropeptide suggested that the shrimp CHH-like peptide consisted of a hydrophobic region at the N-terminal end of the deduced peptide (data not shown). Therefore, this region most likely represented the signal peptide of the hormone [12]. A relatively short CHH precursor-related peptide (CPRP) was also found in the MeCHH-like prepropeptide. The shrimp mature peptide showed the highest overall homology to the CHH of the shrimp, *P. japonicus*, followed by the CHHs of other crustaceans (Fig. 2). The overall amino acid identity of MeCHH-like peptide was 42–69% as compared to other CHHs. Furthermore, it showed 42% amino acid identity as compared to the ion transport protein of the locust [13]. However, the position that the processing site of CHH (Lys-Arg) in other crustaceans was replaced by the amino acid residue Arg-Arg in the shrimp CHH-like peptide.

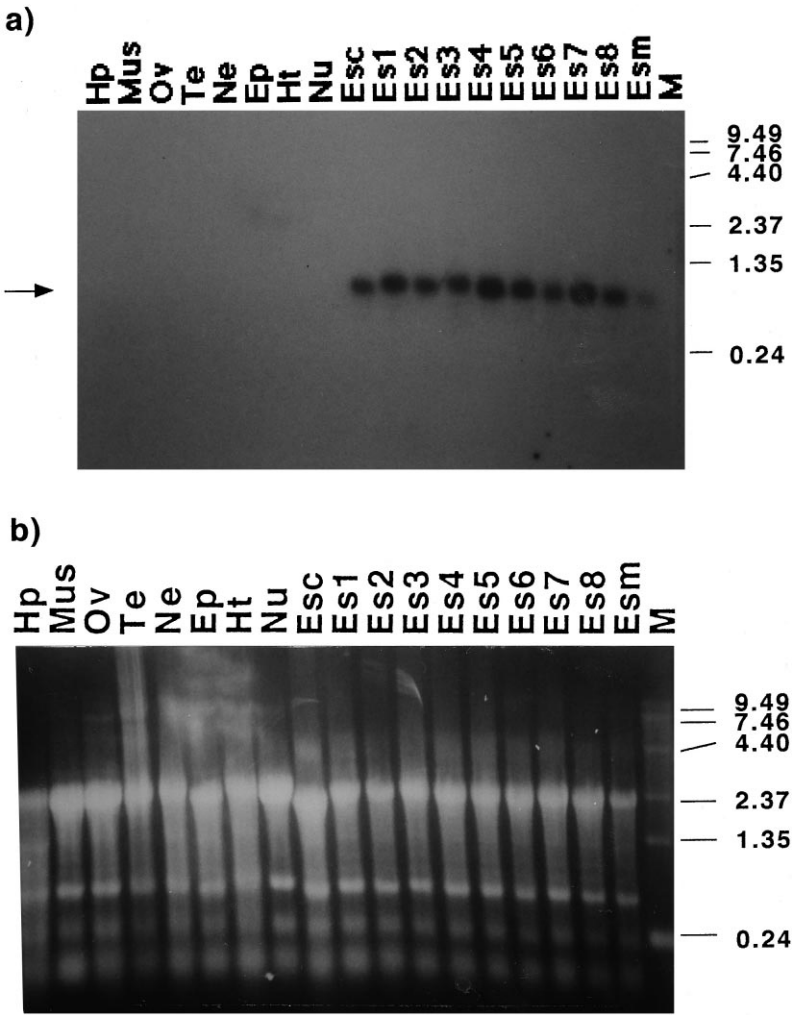

Fig. 3. Northern blot analysis of the shrimp CHH-like mRNA expression. a: Tissue distribution of MeCHH-like. Lanes are Hp (hepatopancreas); Mus (muscle); Ov (ovary); Te (testis); Ne (nerve cord); Ep (epidermis); Ht (heart); Nu (pre-hatch nauplius); Esc (eyestalk of intermolt female; Es1–Es8: eyestalks of different females at different vitellogenic stages I–V) and Esm (eyestalk of adult male); M (RNA size marker, numbers on the right indicate RNA size in kb). The CHH-like probe is derived from the original RT-PCR clone. The arrow indicates the CHH-like RNA hybridized with the probe. b: Ethidium bromide staining of the same gel before RNA transfer to show the integrity of RNA and the relative concentration of the RNA (average 10 μg) loaded in each lane.

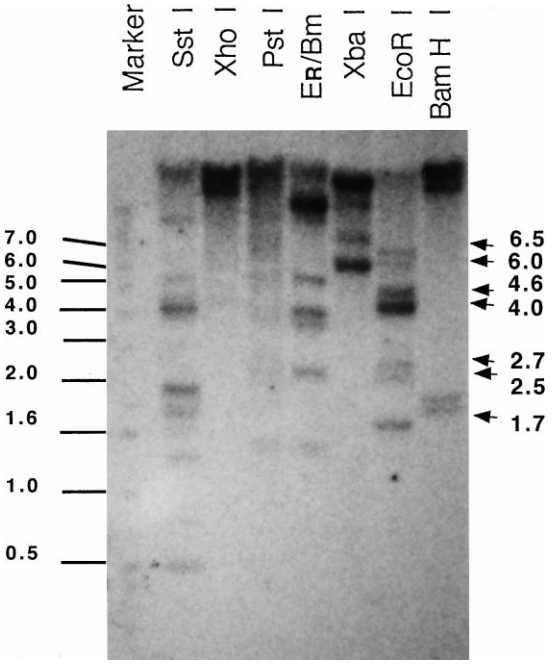

Fig. 4. Southern blot detection of the MeCHH-like gene in the shrimp genome. Shrimp genomic DNA was digested by *Sst*I, *Xho*I, *Pst*I, *Xba*I, both *Eco*RI and *Bam*HI (ER/BM), *Eco*RI, and *Bam*HI. The membrane was hybridized to a CHH-specific probe as described above. The arrows on the right indicate the sizes of the DNA fragment hybridized to the CHH-like probe when the genomic DNA was digested by *Eco*RI. The DNA marker was a 1 kb ladder labeled with radioactive [γ-^{32}P]ATP by a T4 phosphonucleotide kinase.

3.2. Expression of the shrimp CHHs

The size of the cDNA detected in the Northern blot analysis was close to the size of the cDNA (Fig. 3) which suggested that full-length cDNA of the shrimp was cloned. MeCHH-like peptide is expressed in the eyestalks of female and male. It is not detected in the hepatopancreas, muscle, ovary, testis, nerve cord, epidermis, heart, and pre-hatch nauplius. During the reproductive cycle, a much higher level of mRNA was detected in the eyestalk of female shrimp undergoing vitellogenesis (Fig. 3).

3.3. Organization of the shrimp CHH-like gene

When shrimp genomic DNA was digested with *Eco*RI, more than eight DNA fragments (i.e. 6.4 kb, 4.0 kb, 3.8 kb, 3.4 kb, 2.7 kb, etc.) were labeled by the probe (Fig. 4). Screening of an unamplified shrimp genomic DNA library (constructed from a single shrimp) with CHH-like specific cDNA probe produced three positive clones (λ43, λ47 and λ52). Each clone carried 3–4 genes specific for CHH-like peptide (Fig. 5a). *Eco*RI restriction digestion of these genomic clones released the corresponding DNA fragments as labeled by the MeCHH-like probe in the Southern blot (data not shown). After restriction digestions and Southern blot analysis, a gene cluster for the shrimp CHH-like was reconstructed (Fig. 5b). DNA sequence determination on four of the six genes revealed that these genes were highly conserved with only a few changes of nucleotide sequence in the coding region. For example, the prepropeptide showed an additional five amino acids in the region of the signal peptide (Fig. 6) and the deduced amino acid sequence of the mature peptide was unchanged. The size of these CHH-like genes ranged from 1.5 to 2.0 kb. The CHH-like genes for *M. ensis* consist of three exons and two introns. The first intron separates the signal peptide and the second intron separates the mature peptide in the coding region (Fig. 5c). The size of intron 1 for these CHH genes is 300–350 bp. The size of intron 2, however, ranged from 200 to 700 bp for all six genes. The size of the shrimp CHH-like gene spans approximately 2 kb (Fig. 5c). The intron and exon boundaries and the splice site for the donor and acceptor consensus sequence (GT-----AG) for the CHH genes were also conserved. The DNA se-

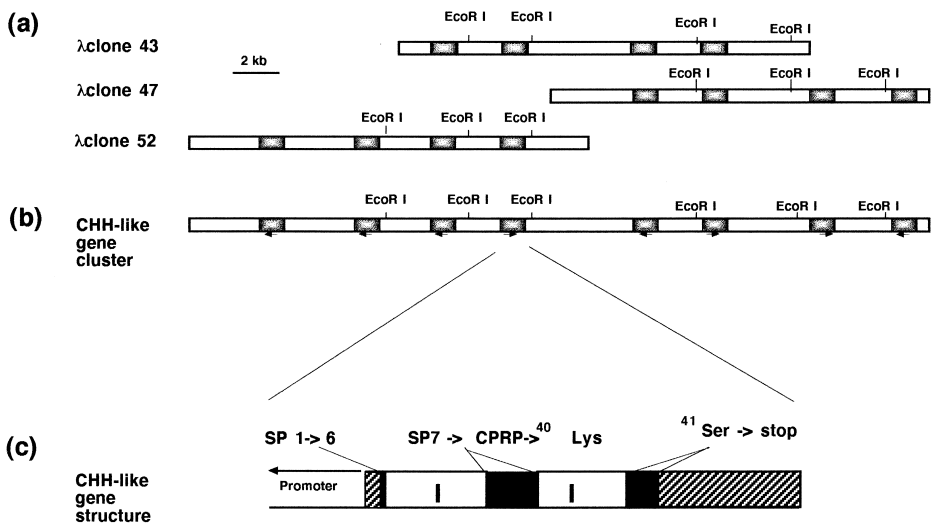

Fig. 5. Organization of the shrimp CHH-like gene cluster in the shrimp genome. a: *Eco*RI restriction map of the shrimp genomic DNA containing the CHH-like gene cluster in three different genomic clones (λ43, λ47 and λ52). The CHH genes are represented in diagonally shaded boxes. b: The reconstructed CHH-like gene cluster spans 20 kb in length. c: Enlarged sketch for the structure of individual CHH-like gene. The 5′ and 3′ end non-coding regions are in diagonally shaded boxes. The black boxes represent the exon that carries coding sequences and the white boxes represent the introns (I). Exon 1 consists of the 5′ non-coding region and DNA sequence encoding the first six amino acid residues of the signal peptide (SP1-6). Exon 2 carries the coding sequence for Met^{7} (SP7) to Lys^{41} (MP41) of the mature CHH-like peptide. Exon 3 carries the coding seqeunce of the rest of the CHH and the 3′ untranslated region (diagonally shaded). The introns are represented by white boxes.

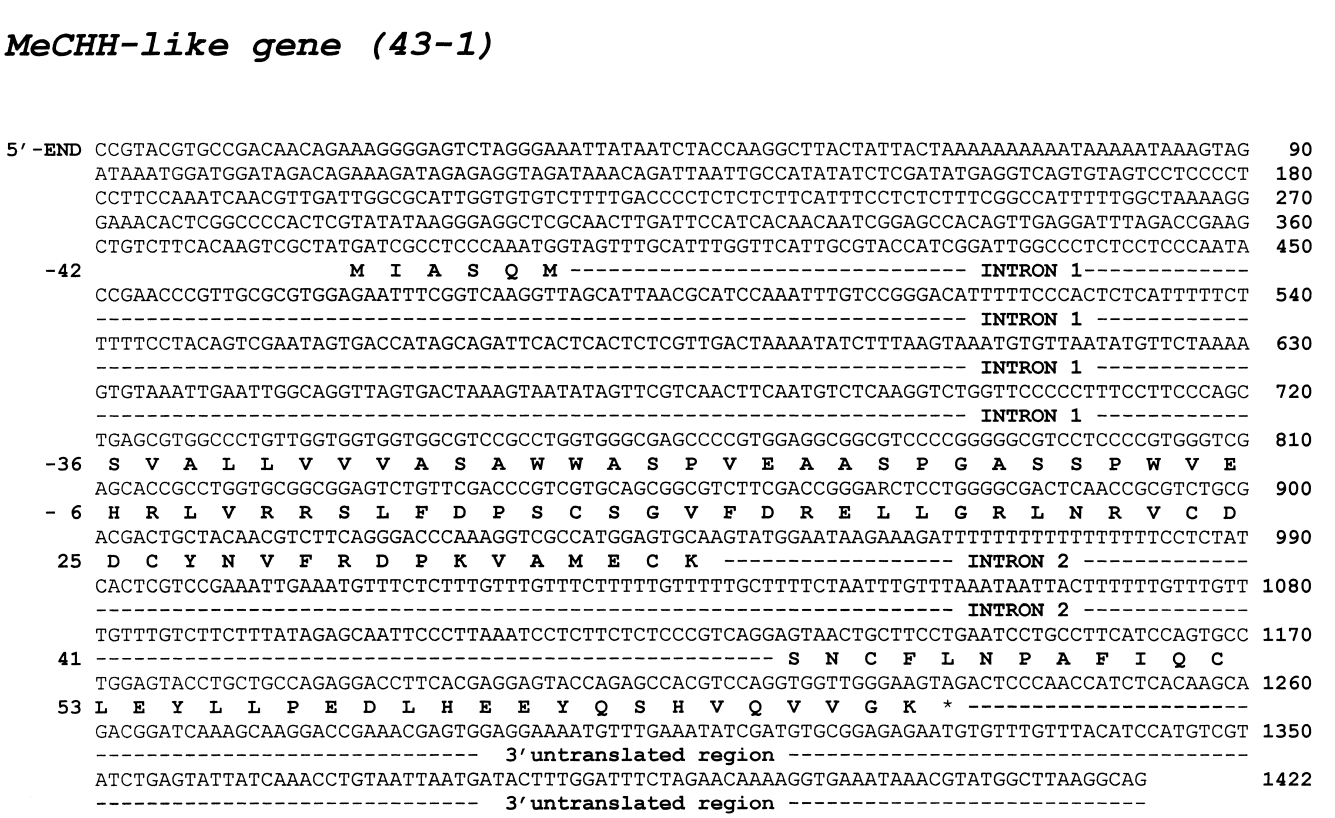

Fig. 6. Complete DNA sequence of the shrimp CHH-like gene (43-1). The numbers on the left and right indicate the positions of the amino acid and nucleotide respectively. The intron and 3′ untranslated region are represented by a string of dashes. Approximately 0.45 kb of the upstream region containing the putative promoter of the shrimp CHH-like gene was sequenced.

quence for one of the genes is shown in Fig. 6. We also determined the DNA sequences of 1.5–2 kb upstream region of 3 CHH-like genes (data not shown) and analysis shows that the 400 bp 5′ upstream region of these CHH-like genes show a high degree of similarity. Several potential recognition sites for transcription factors have been identified (Fig. 7). Putative recognition sites for CAAT, SP1, TATA are all located in the first 150 bp of the proposed CHH gene promoter. Furthermore, specific sequences for other transcription factors (i.e. Pit-1a, CREB, COUP and RXR) have also been identified.

4. Discussion

4.1. Shrimp preproCHH-like cDNA

MeCHH-like cDNA is relatively small in size (617 bp) as compared to those of the lobster [14] and crayfish [5,7,15].

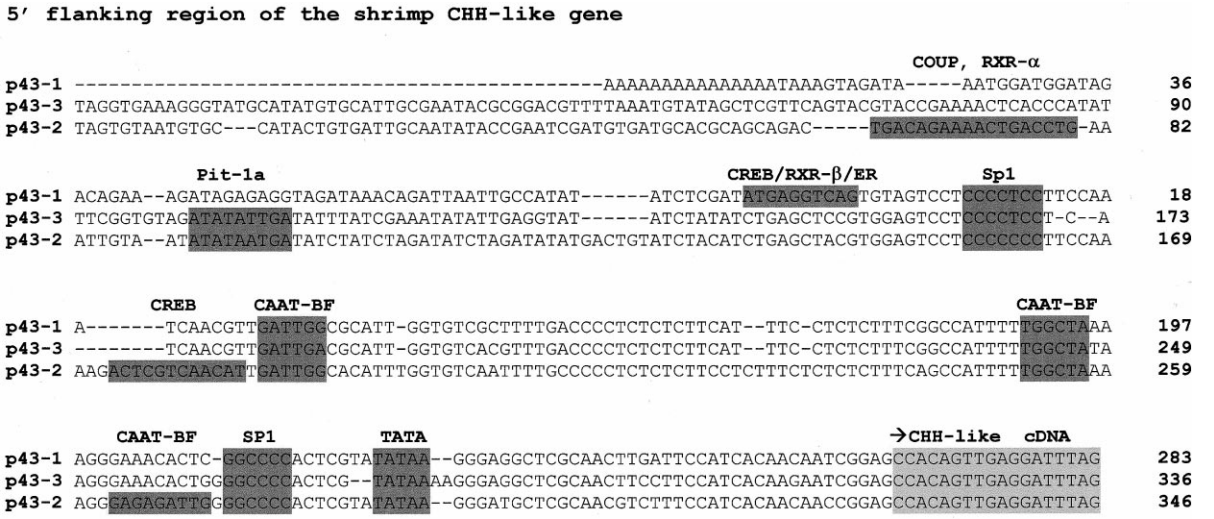

Fig. 7. Sequence analysis of the 5′ end promoter region of the shrimp CHH-like gene. Alignment of the putative promoter region of three of the CHH-like genes. The proposed promoter is underlined and putative binding sites for transcription factors are indicated by a gray box. Computer-generated putative transcription factor binding sites including TATA, CAAT binding factor (CAAT-BF) and SP1. The putative transcription factor binding sites for CREB, Pit-1a, COUP, RXR etc. are indicated.

The cDNA encoded a smaller precursor peptide with the putative CPRP and six cysteine amino acid residues were located at conserved positions. The preprohormone also lacks a glycine residue at position 10 of the mature peptide and appears to be different from those of the lobster and crayfish. The preproCHH-like molecule of *M. ensis* is more closely related to the Pj-SG-III in terms of nucleotide and amino acid sequence. Although the deduced amino acid sequences for the mature peptide of the CHH-like genes are identical, the signal peptide of all MeCHH-like genes has an addition of five amino acid residues. Because the eyestalk cDNA library was constructed from different juveniles, this suggested that individual polymorphism may occur.

4.2. Expression of shrimp preproCHH-like cDNA

Since total RNA from eyestalks of individual shrimp was used in Northern blot analysis, the expression of MeCHH-like mRNA was relatively high. RNA for CHH-like peptide could also be detected in the eyestalk continuously throughout the molting cycle (data not shown). Although CHH is known to be produced by the MTXO of the eyestalk, a CHH peptide has also been detected in the pericardial organs (PO), and the mandibular organ of different crustaceans (unpublished). These tissues may represent additional sites of production of CHH. In this context it is of interest to note that following eyestalk removal in *Homarus americanus*, small but significant levels of CHH remain in the hemolymph, suggesting additional sites of synthesis and release of CHH apart from the eyestalk. [16]. The increase in the expression of the MeCHH-like mRNA during ovarian maturation suggested that the peptide may be involved in the gonad maturation. This is in agreement with the report for the CHH mRNA level in the lobster [17].

4.3. Structural organisation of the CHH-like gene

The gene for MeCHH-like appears functional as there is no detrimental missense mutation in the coding region or in the promoter region of the shrimp CHH gene. Except for the additional sequence of CPRP present in the preprohormone, the shrimp CHH-like gene is similar to the MIH gene of the crab (Chan et al., 1999, in press) and the MIH gene of the same shrimp [18]. For example, intron 1 separates the signal peptide between positions −7 (Met) and −6 (Leu) of the preproCHH-like and intron 2 separates the mature peptide between amino residues 40 (Lys) and 41 (Ser) in the shrimp CHH-like gene. In the locust, *Schistocerca gregaria*, an ion transport protein (ITP) cDNA has been cloned and characterized. In addition to the high overall amino acid similarity (60%) of the mature peptide to the CHHs [21], a long form (L-ITP) has been identified. This L-ITP has an additional 40 amino acid residues inserted in the same relative position that corresponds to the coding sequence of the CHH (i.e. after Arg^{40}). Although the complete gene organization for ITP is not known, it appears that both the insect ITP and the crustacean CHH may share similar gene organization. Since the crustacean neurohormone may derive from a common ancestor gene, we believe that the CHH genes for other crustaceans may share a similar structural organization.

Previous studies in the crab *Carcinus maenas* [8], in the crayfish *Orconectes limosus* [15,21] and in the lobster *H. americanus* [14] have suggested the presence of more than one CHH gene. For example, in the crayfish, two different cDNAs encoding identical amino acid sequences for CHH peptide have been isolated; in the shrimp, *P. japonicus*, a number of sinus gland neuropeptides showing CHH activity have been purified [10]. The existence of isoforms in the crayfish, *O. limosus*, may be the result of gene duplication [15]. This study therefore provides direct evidence for the presence of multiple genes for the MeCHH-like molecule. It is therefore likely that other crustaceans also possess a CHH gene cluster. The presence of gene clusters encoding neurohormones has been reported for both invertebrates and vertebrates. For example, the brain secretory peptide (bombyxin) of the silkworm, *Bombyx mori*, is encoded by 38 genes [19] and the growth hormone of the human is encoded by five copies of genes [20]. Most of these genes are also arranged in a cluster on the same chromosome. The expression of these genes is regulated in a spatial and temporal manner. Since more than six shrimp genomic DNA fragments hybridized to the MeCHH-like probe, we speculate that the shrimp genome may include more than six CHH-like genes. The presence of enhancer-like element (such as for nuclear hormone receptors) in some of the CHH-like genes suggests that steroids such as ecdysone and their receptor may also bind to some of the CHH-like genes and regulate their expression.

In conclusion, this is the first report describing the complete structure of a neuropeptide gene in shrimp. Multiple forms of CHH may represent expression of genes for CHHs which differ in structure by only a few amino acids. Based on our study of the CHH-like gene, and some data on the MIH, we speculate that the gene encoding the neurohormones of other crustaceans may share a very similar organization. The results of this study may provide inferences for the evolution of this important gene family in crustaceans.

Acknowledgements: This work was supported in part by a RCG grant and a HKU institutional grant (CRCG) awarded to S.M.C.

References

[1] De Kleijn, D.P.V. and Van Herp, F. (1995) Comp. Biochem. Physiol. 112B, 573–579.
[2] Webster, S.G. and Keller, R. (1986) J. Comp. Physiol. 156B, 617–624.
[3] Sefiani, M., Le Caer, J.-P. and Soyez, D. (1996) Gen. Comp. Endocrinol. 103, 41–53.
[4] Keller, R. and Wunderer, G. (1978) Gen. Comp. Endocrinol. 103, 41–53.
[5] Kegel, G., Reichwein, B., Tensen, C.P. and Keller, R. (1991) Peptides 12, 909–913.
[6] Chang, E.S., Prestwich, G.D. and Bruce, M.J. (1990) Biochem. Biophys. Res. Commun. 171, 818–826.
[7] Huberman, A., Aguilar, M.B., Brew, K., Shabanowitz, J. and Hunt, D.F. (1993) Peptides 14, 7–16.
[8] Kegel, G., Reichwein, B., Weese, S., Gaus, G., Peter-Katalinic, J. and Keller, R. (1989) FEBS Lett. 255, 10–14.
[9] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[10] Ohira, T., Watanaba, T., Nagasawa, H. and Aida, K. (1997) Mol. Mar. Biol. Biotechnol. 6, 59–63.
[11] Maniatis, T., Sambrook, J. and Fritsch, E.F. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[12] Von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4691.
[13] Meredith, J., Ring, M., Macins, A., Marschall, J., Cheng, N.N., Theilmann, D., Brock, H.W. and Phillips, J.E. (1996) J. Exp. Biol. 199, 1053–1061.
[14] Tensen, C.P., De Kleijn, D.P.V. and Van Herp, F. (1991) Eur. J. Biochem. 200, 103–106.

[15] De Kleijn, D.P.V., Janssen, K.P.C., Martens, G.J.M. and Van Herp, F. (1994) Eur. J. Biochem. 224, 623–629.
[16] Chang, E.S., Keller, R. and Chang, S.A. (1998) Gen. Comp. Endocrinol. 111, 359–366.
[17] De Kleijn, D.V.P., Janssen, K.P.C., Waddy, S.L., Hegemann, R., Lai, W.Y., Martens, G.J.M. and Van Herp, F. (1998) J. Endocrinol. 156, 291–298.
[18] Gu, P.L. and Chan, S.-M. (1998) Mol. Mar. Biol. Biotechnol. 7, 14–22.
[19] Kondo, H., Ino, M., Suzuki, A., Ishizaki, H. and Iwami, M. (1996) J. Mol. Biol. 259, 926–937.
[20] Chen, E.Y., Liao, Y.-C., Smith, D.H. and Barrera-Saldana, H.A. (1995) Genome 4, 479–497.
[21] Soyez, D., Van Herp, F., Rossier, J., Le Caer, J.P., Tensen, C.P. and Lafont, R. (1994) J. Biol. Chem. 269, 18259–18298.
[22] Sun, P. (1994) Mol. Mar. Biol. Biotechnol. 3, 1–6.