Molecular Brain Research, 14 (1992) 7–12

BRESM 70416

Characterization of a cDNA clone encoding molluscan insulin-related peptide V of *Lymnaea stagnalis*

August B. Smit[a], Steven F.T. Thijsen[a], Wijnand P.M. Geraerts[a], Irene Meester[a], Harm van Heerikhuizen[b] and Joos Joosse[a]

[a]*Department of Biology, Vrije Universiteit, Amsterdam, (The Netherlands) and* [b]*Department of Biochemistry, Vrije Universiteit, Amsterdam (The Netherlands)*

(Accepted 10 December 1991)

Key words Insulin-related peptide; Neuropeptide; Central nervous system, Mollusc; *Lymnaea stagnalis*

A cDNA clone encoding molluscan insulin-related peptide V (MIP V) was isolated from a cDNA library of the central nervous system (CNS) of the freshwater snail, *Lymnaea stagnalis*, using a heterologous screening with a previously identified MIP II cDNA The MIP V cDNA encodes a preprohormone resembling the organization of preproinsulin, with a putative signal sequence, and an A and B chain, however, in this case connected by two distinct C peptides, Cα and Cβ, instead of one single C peptide This phenomenon, which is shared by the MIP II precursor, represents a new development in the prohormone organization of peptides belonging to the insulin superfamily The A and B chains of MIPs V, I and II, differ remarkably in primary structure; in contrast, the Cα peptide domains are almost identical. MIP V has only limited sequence similarity with insulins and related peptides. Both MIP V and I exhibit structural features, which make them a unique class of the insulin superfamily. The MIP I, II and V genes are expressed in a single type of neuron· the growth controlling neuroendocrine light green cells of the *Lymnaea* CNS.

INTRODUCTION

Insulins and related peptides, i.e. relaxins and insulin-like growth factors I and II (IGFs I and II), form a superfamily of regulatory peptides[26]. In vertebrates, they signal essential steps in growth, development, reproduction and metabolism[6,10,13]. In most invertebrate groups, there is substantial immunocytochemical evidence for the presence of insulin or related molecules[15], however, structural data are scarce and only available for insects [bombyxins, *Bombyx mori*[16]; locusta insulin-related peptide (LIRP), *Locusta migratoria*[18]], and molluscs [(MIPs), *Lymnaea stagnalis*[24,25]]. In contrast to the synthesis of insulin in the pancreas of vertebrates, the insulin-related molecules of invertebrates are not produced in endocrine cells, but in neuroendocrine cells of the central nervous system (CNS). In *L. stagnalis*, insulin-related peptides are produced by the light green cells (LGCs), peptidergic neurons, which are present as four clusters of together ~150 cells in the cerebral ganglia[14,32]. The LGCs are involved in the regulation of growth of both the soft body parts[11] and the shell[7], as well as of protein and carbohydrate metabolism[8,28,29]. We have recently identified two LGC-specific cDNA clones encoding molluscan insulin-related peptides (MIPs)[24,25]. The MIPs belong to a small family of insulin-related peptides. Here, we report the isolation and characterization of a cDNA encoding a MIP I, II related preprohormone, designated preproMIP V, which is also expressed by the LGCs. Overall, preproinsulin and the MIP V preprohormone share a similar pattern of organization, with the exception of an extra C peptide domain, which is present in preproMIP V. Although MIP I, II and V differ substantially, they have structural features in common that are absent in other insulins and related peptides. This makes them a unique class of the insulin superfamily.

MATERIALS AND METHODS

Screening of the Lymnaea cDNA library

We screened 20,000 clones of an amplified cDNA library in λgt10 of the CNS of *L. stagnalis*[31], using Hybond-N membranes (Amersham International Corporation) at a density of 5,000 pfu/135 mm^2 filter Clones were purified by rescreening at a lower plaque density Hybridization of the membranes was performed in 6× SSC (1× SSC. 0.15 M NaCl and 0.015 M sodium citrate) for 16 h at 65°C The filters were washed in 1× SSC for 45 min at 65°C. Autoradiography was done on preflashed Kodak X-OMAT using an intensifying screen

MIP II cDNA labeling

Radiolabeled cDNA was synthesized by primer extension on a single-stranded M13 MIP II cDNA clone[25] using the Klenow frag-

Correspondence. A.B. Smit, Department of Biology, Vrije Universiteit, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands Fax: (31) (20) 6429202

ment of DNA polymerase I, and an M13 specific primer A mixture of 0.1 pmol of M13 single-stranded DNA and 0 5 pmol M13 specific primer (5′-TGACCGGCAGCAAAATG-3′) was heated to 95°C and allowed to cool to 20°C To the mixture 10 μCi [α-^{32}P]-dATP and 7 U Klenow enzyme (Boehringer Mannheim) were added and the reaction was carried out for 20 min at 20°C A chase reaction was performed by adding 0 2 mM dATP for 10 min at 20°C Next, free label was separated from synthesized DNA on a Sephadex G-50 column The specific activity of the probe was 1 × 10^8 dpm/μg

Size determination of preproMIP V mRNA

RNA was isolated according to the method of Chirgwin et al [4] About 12 μg of RNA from the cerebral ganglia was glyoxylated, fractionated on a 1 6% agarose gel, transferred to a Hybond-N filter, and hybridized for 16 h at 65°C with the 5′ end labeled oligonucleotide probe 5′-CTCGTAAAATACATCAACTGTGCATAC-ATTATA-3′ (spec act 10^9 dpm/μg), specific for the MIP V mRNA Filters were washed in 2× SSC for 20 min at 65°C. Glyoxylated yeast (*Saccharomyces cerevisiae* ribosomal RNAs, 26S (3400 bases) and 17S (1800 bases), were used as size markers

Nucleotide sequence analysis

An *Eco*RI fragment from λgt10 was subcloned in M13mp19 and sequenced in both orientations according to the method of dideoxy chain termination[23], using either Klenow polymerase or sequenase Reactions were performed with standard nucleotide mixtures, or with dITP as a substitute for dGTP. Sequenase and reaction mixtures were obtained from USB

```
                                                        5'-CUAAAAUCCAAA    12

1 ──► SIGNAL SEQUENCE
AUG GCC GGC GUG CGC CUC GUU UUC ACC AAG GCC UUC AUG GUC ACA    57
met ala gly val arg leu val phe thr lys ala phe met val thr

                    20
GUG CUU CUC ACC CUG UUG CUG AAC AUU GGC GUG AAA CCU GCA GAG   102
val leu leu thr leu leu leu asn ile gly val lys pro ala glu

    ┌──► B CHAIN                  40
GGU CAA UUU UCA GCU UGC AGU UUC UCA AGC AGA CCA CAC CCC CGU   147
gly gln phe ser ala cys ser phe ser ser arg pro his pro arg

                                                        60
GGG AUU UGU GGC UCG GAU CUG GCC GAC CUC CGG GCG UUC AUA UGU   192
gly ile cys gly ser asp leu ala asp leu arg ala phe ile cys

                                          ──► Cβ PEPTIDE
UCG AGG CGC AAC CAA CCG GCA AUG GUC AAG AGG GAC GCA GAA ACA   237
ser arg arg asn gln pro ala met val lys arg asp ala glu thr

                    80                    ──► Cα PEPTIDE
GGC UGG UUA CUG CCU GAA ACC AUG GUC AAG AGG AAU GCA CAA ACA   282
gly trp leu leu pro glu thr met val lys arg asn ala gln thr

                                       100
GAU CUG GAC GAC CCA CUG CGU AAC AUA AAG CUG AGC AGC GAG AGC   327
asp leu asp asp pro leu arg asn ile lys leu ser ser glu ser

                                  ──► A CHAIN                 120
GCU UUG ACC UAU CUG ACC AAG AGA CAA CGU ACA ACC AAC UUA GUG   372
ala leu thr tyr leu thr lys arg gln arg thr thr asn leu val


UGU GAA UGC UGU UAU AAU GUA UGC ACA GUU GAU GUA UUU UAC GAG   417
cys glu cys cys tyr asn val cys thr val asp val phe tyr glu

            139
UAC UGC UAU UAA UGGAGUUUGACAAUUAAAAAAAAUAGAUCUUUGAAUAUUUGGC   472
tyr cys tyr stop

CGAUUUCUUAACCAAAAAAAAA -3'                                    494
```

Fig 1 Nucleotide sequence of preproMIP V cDNA and derived amino acid sequence The putative proteolytic processing sites (Lys-Arg) are boxed, and the various peptide domains are indicated the signal peptide, A and B chains, and the Cα and Cβ peptides. The number of nucleotides is indicated at the end of each line The predicted amino acid sequence of preproMIP V is numbered designating the first methionine as position 1.

In situ hybridization

The cellular localization of preproMIP V mRNA was studied in sections of the cerebral ganglia of *L. stagnalis* by in situ hybridization using the MIP V-specific oligonucleotide 5′-CTCGTAAAATA-CATCAACTGTGCATACATTATA-3′, labeled at the 3′ end with [α-^{35}S]dATP using dideoxy terminal nucleotidyl transferase Conditions molecular ratio of oligonucleotide over [α-^{35}S]dATP was 1 10; [α-^{35}S]dATP was obtained from Dupont-New England Nuclear, dideoxy terminal nucleotidyl transferase was obtained from Gibco/Bethesda Research Laboratoria, reaction conditions as described by the supplier Hybridization conditions 4 ng oligonucleotide (spec act. 2 5 × 10^9 dpm/μg) was added per slide; hybridization in 3× SSC at 37°C, and washing up to 0 5× SSC at 45°C, other conditions as described in ref 25

RESULTS AND DISCUSSION

Screening of the λgt10 cDNA library of the *Lymnaea* CNS with the MIP II cDNA, yielded 20 positive cDNA clones, of which the longer cDNAs were subcloned as *Eco*RI fragments in M13mp19. Sequence analysis, in both directions, of one of these clones (551 bp), revealed a single open reading frame encoding a preprohormone of 139 amino acids with the characteristics of preproinsulin and other precursors of insulin-related peptides (Fig. 1). Because the protein sequence found, was most structurally related with the previously described preprohormones of MIP I and II[24,25], we called the newly pre-

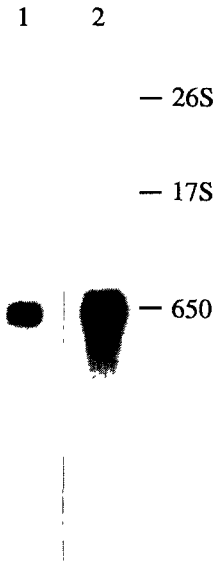

Fig 2 Size determination of preproMIP V mRNA as determined by Northern blotting Total RNA (12 μg), isolated from cerebral ganglia was size-fractionated on a 1.6% agarose gel, blotted to Hybond-N and hybridized to ^{32}P-labeled oligonucleotide probes, specific for either MIP V, (Lane 1) or MIP II (Lane 2), 26S (3400 bases) and 17S (1800 bases) indicate the positions of yeast rRNAs (transcript size markers). The size of the MIP I and V transcripts is 650 bp

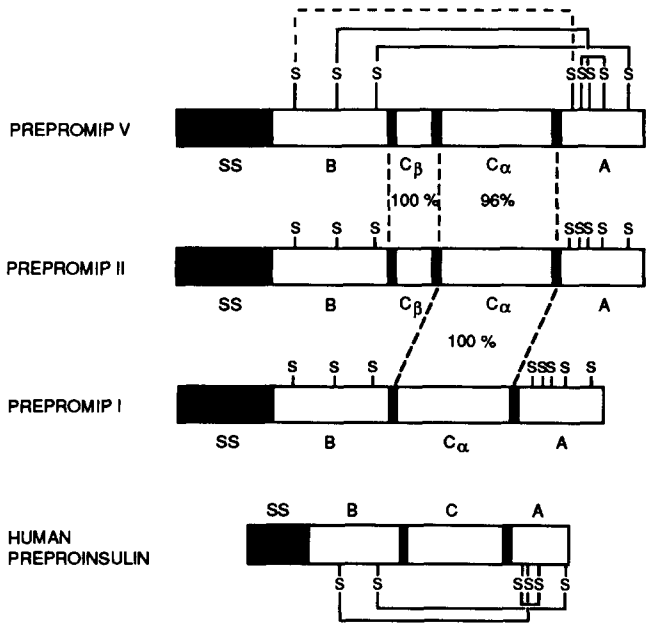

Fig. 3. Schematic representation of the MIP V, II and I precursors, as predicted from the identified cDNAs, and of the human insulin precursor Notice the conservation of the overall organization of the MIP V and II precursor, and the almost complete sequence conservation of the Cα peptides In MIP V and in human insulin, the three homologous disulphide bridges are indicated by a solid line The putative extra disulphide bridge, present in the MIPs, is indicated by a dotted line SS, signal sequence, B, B chain; C, C peptide of human insulin; Cα and Cβ, MIP C peptides, A, A chain; -S, indicates the position of a cysteine residue. Vertical bars indicate the position of dibasic (Lys-Arg) proteolytic processing sites.

dicted protein, with respect to numbering of putative genes, preproMIP V.

The cDNA clone contains, in addition to the 417 nucleotide long coding region, 12 nucleotides of the 5′ untranslated leader sequence and 65 nucleotides of the 3′ untranslated region. The length of the preproMIP V cDNA (494 nucleotides), as obtained by sequence analysis, corresponds reasonably well with the length of ~650 nucleotides of the MIP V transcript as determined by Northern blotting (Fig. 2), taking into account a poly(A) stretch with an average length of ~100 nucleotides. A signal for polyadenylation[20] may be present at 463 (AUUAAU), although such a signal has, to our knowledge, not been reported before.

PreproMIP V consists of a signal sequence, A and B chains, as well as two putative C peptides, the Cα and Cβ peptides. A comparison of the overall organization of preproMIP V with preproMIP I, II, and human preproinsulin is given in Fig. 3. We estimate that initiation of translation will take place at the Met residue on position 1. Alternatively, the Met at position 13 could function as initiation site. This option is, however, not very likely since only the Met at position 1 has the appropriate consensus sequence for the initiation of protein synthesis[17]. In accordance with this, the primary structure of the cDNA encoding the MIP V precursor predicts Gly at position 31 as the most likely cleavage site giving rise to a signal peptide of 31 residues with a clearly hydrophobic character[30]. Furthermore, it can be predicted that proMIP V can be endoproteolytically cleaved at three Lys-Arg processing sites (residues 69/70, 84/85 and 111/112). These predictions were recently confirmed by protein sequencing (K.W. Li, unpublished data).

MIP V, the insulins and related peptides share the amino acids that are necessary in adopting the basic insulin conformation (Fig. 4a). In the A and B chains of MIP V, among the 8 cysteines, 6 are present at positions typical for the insulin superfamily, suggesting that the three characteristic disulphide bridges in the MIP molecules have been conserved (A6, A7, A11, A20, B7 and B19). In addition, the important hydrophobic core residues of the globular insulin structure are either conserved as identical residues (A11, A3, A19, A20, B11, B19; Arg at B15 is an exception), or are replaced by residues with an equally hydrophobic character (residues A2, A16, B6, B12 and B18). In addition, α-helices, which are present in the A and B chains of vertebrate insulins, can be predicted for MIP V[5]. The A and B chains of both MIP I and V are N-terminally extended, and the B chain extension contains an extra cysteine at B-6, which may form a third interchain disulphide bridge with an extra cysteine at A4. Due to these properties, the MIPs are the most complexly folded molecules to date of the insulin superfamily. Thus, although MIP V has important characteristics in common with various vertebrate insulins, IGFs I and II, relaxins, bombyxins, and LIRP, it possesses unique features not found in any of these peptides. This indicates that MIPs I, II and V form another branch of an ancient superfamily of peptides regulating growth and metabolism. The alignment of the A and B chains of the MIPs with those of other members of the insulin superfamily shows that, besides residues strictly necessary for the insulin conformation, there is low amino acid homology (Fig. 4b). The highest sequence homology of MIP V is observed with MIP II, MIP I, and the bombyxins B1 and A1.

The Cα and Cβ peptides are cleaved out separately of proMIP V, yielding a mature MIP V (K.W. Li, unpublished sequence data), like in the case of insulin. The presence of two C peptides is, apart from the MIP II prohormone, not documented for any of the members of the insulin superfamily. Also, the MIP V Cα and Cβ peptides show no sequence homology with C peptides of other proinsulins, or related prohormones of vertebrates and invertebrates. With respect to the vertebrate C peptides this is not unexpected since these peptides exhibit a strong divergence amongst themselves. Interestingly,

A CHAINS

		-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
INVERTEBRATE INSULIN-RELATED MOLECULES																											
MIP V	Lymnaea	Q	R	T	T	N	L	V	C	E	C	C	Y	N	V	C	T	V	D	V	F	Y	E	Y	C	Y	
MIP I	Lymnaea	Q	G	T	T	N	I	V	C	E	C	C	M	K	P	C	T	L	S	E	L	R	Q	Y	C	P	
MIP II	Lymnaea	Q	R	T	T	N	L	V	C	E	C	C	F	N	Y	C	T	P	D	V	V	R	K	Y	C	Y	
Bombyxin A1	Bombyx	-	-	-	-	G	I	V	D	E	C	C	L	R	P	C	S	V	D	V	L	L	S	Y	C	-	
Bombyxin B1	Bombyx	-	-	-	-	G	V	V	D	E	C	C	F	R	P	C	T	L	D	V	L	L	S	Y	C	G	
Bombyxin IV	Bombyx	-	-	-	-	G	V	V	D	E	C	C	I	Q	P	C	T	L	D	V	L	A	T	Y	C	-	
LIRP	Locusta	-	-	T	R	G	V	F	D	E	C	C	R	K	T	C	S	I	S	E	L	Q	T	Y	C	G	
VERTEBRATE INSULIN-RELATED MOLECULES																											
IGF I	Human	-	-	-	-	G	I	V	D	E	C	C	F	R	S	C	D	L	R	R	L	E	M	Y	C	A	...
IGF II	Human	-	-	-	-	G	I	V	E	E	C	C	F	R	S	C	D	L	A	L	L	E	T	Y	C	A	....
Relaxin	Human	R	P	Y	V	A	L	F	E	K	C	C	L	I	G	C	T	K	R	S	L	A	K	Y	C	-	
VERTEBRATE INSULINS																											
Insulin	Human	-	-	-	-	G	I	V	E	Q	C	C	T	S	I	C	S	L	Y	Q	L	E	N	Y	C	N	
Insulin	Rat I	-	-	-	-	G	I	V	D	Q	C	C	T	S	I	C	S	L	Y	Q	L	E	N	Y	C	N	
Insulin	Hagfish	-	-	-	-	G	I	V	E	Q	C	C	H	K	R	C	S	I	Y	N	L	Q	N	Y	C	N	

B CHAINS

		-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
INVERTEBRATE INSULIN-RELATED MOLECULES																																										
MIP V	Lymnaea	Q	F	S	A	C	S	F	S	S	R	P	H	P	R	G	I	C	G	S	D	L	A	D	L	R	A	F	I	C	S	R	R	N	Q	P	A	M	V	-	-	-
MIP I	Lymnaea	Q	F	S	A	C	N	I	N	D	R	P	H	R	R	G	V	C	G	S	A	L	A	D	L	V	D	F	A	C	S	S	S	N	Q	P	A	M	V	-	-	-
MIP II	Lymnaea	Q	-	S	S	C	S	L	S	S	R	P	H	P	R	G	I	C	G	S	N	L	A	G	F	R	A	F	I	C	S	N	Q	N	S	P	S	M	V	-	-	-
Bombyxin A1	Bombyx	-	-	-	-	-	-	-	Q	Q	P	Q	R	V	H	T	Y	C	G	R	H	L	A	R	T	L	A	D	L	C	W	E	A	G	V	D	-	-	-	-	-	-
Bombyxin B1	Bombyx	-	-	-	-	-	-	-	-	-	Q	E	V	A	R	T	Y	C	G	R	H	L	A	D	T	L	A	D	L	C	F	G	V	E	-	-	-	-	-	-	-	-
Bombyxin IV	Bombyx	-	-	-	-	-	-	-	-	E	A	N	V	A	H	H	Y	C	G	R	H	L	A	N	T	L	A	D	L	C	W	D	T	S	V	E	-	-	-	-	-	-
LIRP	Locusta	-	-	-	-	-	-	S	G	A	P	Q	P	V	A	R	Y	C	G	E	K	L	S	N	A	L	K	L	V	C	R	G	N	Y	N	T	M	F	-	-	-	-
VERTEBRATE INSULIN-RELATED MOLECULES																																									
IGF I	Human	-	-	-	-	-	-	-	-	-	-	-	G	P	E	T	L	C	G	A	E	L	V	D	A	L	Q	F	V	C	G	D	R	G	F	Y	F	N	K	P	T	G
IGF II	Human	-	-	-	-	-	-	-	-	A	Y	R	P	S	E	T	L	C	G	G	E	L	V	D	T	L	Q	F	V	C	G	D	R	G	F	Y	F	S	R	P	A	S
Relaxin	Human	-	-	-	-	-	-	-	K	W	K	D	D	V	I	K	L	C	G	R	E	L	V	R	A	Q	I	A	I	C	G	M	S	T	W	S	K	R	S	L	-	-
VERTEBRATE INSULINS																																									
Insulin	Human	-	-	-	-	-	-	-	-	-	-	F	V	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	T	P	K	T	-
Insulin	Rat I	-	-	-	-	-	-	-	-	-	-	F	V	K	N	H	L	C	G	P	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	T	P	K	S	-
Insulin	Hagfish	-	-	-	-	-	-	-	-	-	-	R	T	T	G	H	L	C	G	K	D	L	V	N	A	L	T	I	A	C	G	V	R	G	F	F	Y	D	P	T	K	M

A Chains	MIP I	MIP II	MIP V
MIP V	52	76	100
MIP I	100	56	52
MIP II	56	100	76
Bombyxin A1	50	45	50
Bombyxin B1	52	52	50
LIRP	43	30	30
Human insulin	42	28	29
Human relaxin	33	33	29
Human IGF I	52	38	33

B Chains	MIP I	MIP II	MIP V
MIP V	68	74	100
MIP I	100	53	68
MIP II	53	100	74
Bombyxin A1	18	21	23
Bombyxin B1	29	29	33
LIRP	13	13	13
Human insulin	17	17	20
Human relaxin	16	16	16
Human IGF I	22	22	30

Fig. 4. Amino acid sequence comparison of MIPs V, I and II, selected insulins and other insulin-related peptides. a: the alignment of the MIP sequences and other members of the insulin superfamily. Notice the strict conservation of the cysteine residues (including the extra cysteines A4 and B –6), as well as the conservation of character of the substituted amino acids (e.g., A2, A16 and B –4, B6, B14, B18). Numbers 1 indicate the first residues of the A and B chains of human insulin. The sequences of the MIPs have been predicted from the cDNAs. b: percentages of amino acid residues shared by the A and B chains of MIPs I, II and V, and of these MIPs and insulins and other insulin-related peptides. The percentage of sequence identity among insulins and related peptides from vertebrates ranges from 55 to 95%.

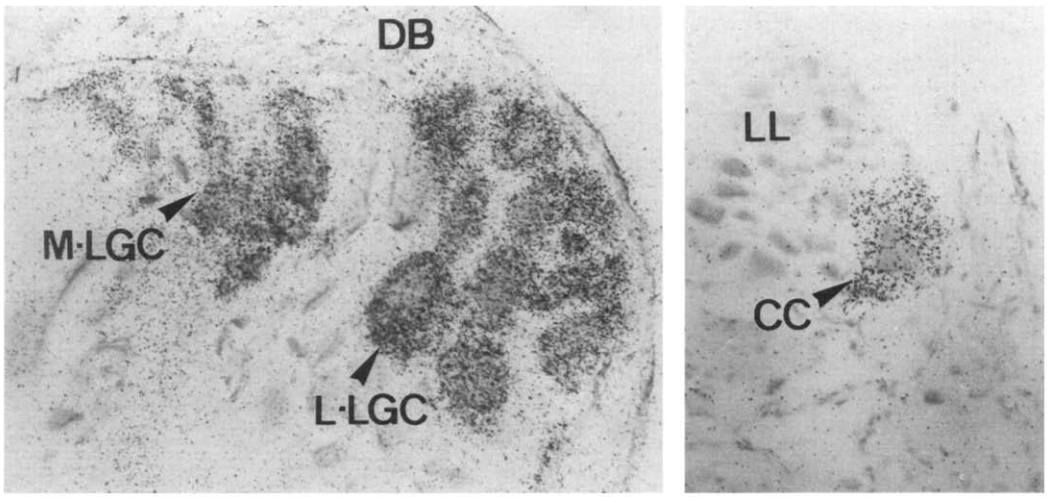

Fig. 5. Localization of preproMIP V mRNA in sections of the cerebral ganglia of *L. stagnalis* by in situ hybridization using a ^{35}S-labeled MIP oligonucleotide. Left: hybridization of MIP V mRNA in the lateral (L-LGC) and median (M-LGC) group of LGCs in the left cerebral ganglion. Magnification ×95. Right: hybridization of MIP mRNA in the canopy cell of the lateral lobe of the left cerebral ganglion. Magnification ×210. DB, the non-nervous female gonadotropin-producing dorsal bodies.

the MIP V Cα peptide is almost identical to the MIP II and MIP I Cα peptide, which suggests an important physiological role for the Cα peptide in *Lymnaea*. Suggestions as to the evolutionary origin of the Cβ peptides, as well as to the remarkable sequence identity of the Cα peptides must await further studies on the genomic organization of the MIP genes.

In situ hybridization on histological sections of the *Lymnaea* CNS, using a radiolabeled probe specific for MIP V, shows that the LGCs in the cerebral ganglia are the main neurons to express the MIP V gene (Fig. 5, left side). In addition to the LGCs, the MIP V gene is expressed in the canopy cells, which are present as single neurons in the lateral lobes, small ganglia attached to the cerebral ganglia[12,27]. This is an interesting observation since it has been suggested, on the basis of endocrinological evidence, that the canopy cells are involved in the control of the synthesis and release activities of the LGCs[12,22]. The pattern of MIP V gene expression is similar to that of the MIP I gene.

In insulin, residues in the A chain (A1, A5, A8, A19 and A21) and the spatially adjacent B chain residues (B12, B16 and B24–B26) are implicated in receptor interactions[3,21]. The residues in MIP V at these positions differ strongly from those found in insulin, or other members of the insulin superfamily. In addition, MIPs I, II and V, the environment of these residues is probably altered, due to the presence of an extra disulphide bridge. It is likely, therefore, that both MIP I, II and V will bind to a receptor, which is different, at least in the extracellular domain, from the mammalian type of insulin receptor.

MIP I, II and V diverge also in the presumed receptor binding domain (e.g., residues A8, A21, B16, B24 and B26). Due to these differences, each MIP may bind a distinct (sub)type of receptor, or alternatively, the same type with different affinities. Accordingly, the MIPs may be involved in the control of distinct yet interrelated processes associated with growth and metabolism. Since the LGCs can synthesize several different MIPs, differential production and release of these MIPs in response to different environmental stimuli would considerably enhance the information handling capacity of these neurons.

Interestingly, in molluscs and insects, insulin-related peptides, are produced by neuroendocrine cells of the CNS. The evidence that insulin is produced by neurons in the CNS of vertebrates is controversial[12]. However, in view of the evidence that pancreatic islet cells are derived from neuroectodermal precursor cells[1], it seems possible that in the Archaemetazoa, a stemgroup from which both vertebrates and invertebrates have arisen, the evolution of the insulin superfamily may have been within primitive neuroendocrine cells, probably associated with the digestive system. Interestingly, in *Lymnaea* and in other invertebrates, there is evidence for the presence of insulin in the gut[9], however the evidence stems only from immunological experiments.

An intriguing aspect of the evolution of the insulin superfamily concerns the striking difference in the degree of sequence conservation of its members. Insulins from over 50 vertebrate species have now been identified and described in detail, and the data indicate that insulin is a conserved molecule both in structure and function. In addition to insulin, the IGFs, and to a lesser extent also the bombyxins are conserved, suggesting a low acceptance of mutational change In contrast, MIPs I, II and V, like the relaxins, are highly divergent, indicating a high degree of mutational acceptance. Whether a specific molecular mechanism is responsible for the MIP and relaxin sequence diversities remains an unresolved question.

Acknowledgements This work was supported by the Foundation for Fundamental Biological Research (BION), which is subsidized by the Netherlands Organization for Scientific Research (NWO)

ABBREVIATIONS

CNS	central nervous system
IGF I	insulin-like growth factor I
IGF II	insulin-like growth factor II
LGCs	light green cells
λgt10	lambda bacteriophage gt10
LIRP	locust insulin-related peptide
MIP I	molluscan insulin-related peptide I
MIP II	molluscan insulin-related peptide II
MIP V	molluscan insulin-related peptide V
pfu	plaque forming unit

REFERENCES

1 Alpert, S , Hanahan, D and Tertelman, G , Hybrid insulin genes reveal a developmental lineage for pancreatic endocrine cells and imply a relationship with neurons, *Cell*, 53 (1988) 295–308

2 Baskin, D G , Figlewicz, D.P , Woods, S C , Porte, D. and Dorsa, D H , Insulin in the brain, *Annu Rev Physiol* , 49 (1987) 335–347

3 Blundell, T.L , Dodson, G.G., Hodgkin, D C and Mercola, D A , Insulin· the structure in the crystal and its reflection in chemistry and biology, *Adv Protein Chem* , 26 (1972) 279–402

4 Chirgwin, J M , Przybyla, A E , MacDonald, R.J and Rutter, W J , Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease, *Biochemistry*, 18 (1979) 5295.

5 Chou, P Y. and Fasman, G D , Prediction of protein conformation, *Biochemistry*, 13 (1974) 222–245

6 Czech, M P , in M P. Czech (Ed), *Molecular Basis of Insulin*

Action, Plenum, New York, 1985, p 473
7 Dogterom, A A , van Loenhout, H and van der Schors, R C , The effect of the growth hormone of *Lymnaea stagnalis* on shell formation, *Gen Comp Endocrinol* , 39 (1979) 63–68
8 Dogterom, A A , The effect of the growth hormone of the freshwater snail *Lymnaea stagnalis* on biochemical composition and nitrogen wastes, *Comp Biochem Physiol* , B65 (1980) 163–167
9 Ebberink, R H M , Smit, A B and van Minnen, J , The insulin family evolution of structure and function in vertebrates and invertebrates, *Biol Bull* , 177 (1989) 176–182
10 Froesch, E R , Schmidt, C , Schwander, J and Zapf, J , Actions of insulin-like growth factors, *Annu Rev Physiol* , 47 (1985) 443–467
11 Geraerts, W P M , Control of growth by the neurosecretory hormone of the light green cells in the freshwater snail *Lymnaea stagnalis*, *Gen Comp Endocrinol* , 29 (1976) 61–71
12 Geraerts, W P M , The role of the lateral lobes in the control of growth and reproduction in the hermaphrodite freshwater snail *Lymnaea stagnalis*, *Gen Comp Endocrinol* , 29 (1976) 97–108
13 Girbau, M , Gomez, J A , Lesniak, M A and de Pablo, F , Insulin and Insulin-like Growth Factor I both stimulate metabolism, growth and differentiation in the postneurula chick embryo, *Endocrinology*, 121 (1987) 1477–1482
14 Joosse, J , Dorsal bodies and dorsal neurosecretory cells of the cerebral ganglia of *Lymnaea stagnalis*, *Arch Neerl Zool* , 15 (1964) 1–103
15 Joosse, J and Geraerts, W P M , In A S M Saleuddin and K M Wilbur (Eds), *The Mollusca, Vol 4*, Academic Press, New York, 1983, pp. 317–406
16 Kawakami, A , Iwami, M , Nagasawa, H , Suzuki, A and Ishizaki, H , Structure and organization of four clustered genes that encode Bombyxin, an insulin-related brain secretory peptide of the silkmoth, *Bombyx mori*, *Proc Natl Acad Sci USA*, 84 (1989) 6843–6847
17 Kozak, M , Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes, *Cell*, 44 (1986) 283–292
18 Lagueux, M , Lwoff, L , Meister, M , Goltzené, F and Hoffmann, J , cDNAs from neurosecretory cells of brains of *Locusta migratoria* (Insecta, orthoptera) encoding a novel member of the superfamily of insulins, *Eur J Biochem* , 187 (1990) 249–254
19 Maniatis, T , Fritsch, E G and Sambrook J , In *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1982
20 Manley, J L , Polyadenylation of mRNA precursors, *Biochem Biophys Acta*, 950 (1988) 1–12
21 Pullen, R A , Lindsay, D G , Wood, S P , Tickle, I.J , Blundell, T L , Wollmer, A , Krail, G , Brandenburg, D , Zahn, H , Gliemann, J and Gammeltoft, S , Receptor-binding region of insulin, *Nature*, 259 (1976) 369–373
22 Roubos, E W , Geraerts, W P M , Boerrigter, G H and van Kampen, G P J , Control of the activities of the neurosecretory light green cells and caudodorsal cells and of the endocrine dorsal bodies by the lateral lobes in the freshwater snail *Lymnaea stagnalis*, *Gen Comp Endocrinol* , 40 (1980) 446–454
23 Sanger, F , Nicklen, S and Coulson, A R., DNA sequencing with chain-terminating inhibitors, *Proc Natl Acad. Sci USA*, 74 (1977) 5463–5467
24 Smit, A B , Vreugdenhil, E , Ebberink, R H M , Geraerts, W P M , Klootwijk, J and Joosse, J , Growth-controlling molluscan neurons produce the precursor of an insulin-related peptide, *Nature*, 331 (1988) 535–538
25 Smit, A B , Meester, I , Geraerts, W P.M , van Heerikhuizen, H and Joosse, J , Characterization of a cDNA clone encoding molluscan insulin-related peptide II of *Lymnaea stagnalis*, *Eur J Biochem* , 199 (1991) 699–703
26 Steiner, D F , Chan S J , Welsh, J M and Kwok, S C M , Structure and evolution of the insulin gene, *Annu Rev Genet* , 19 (1985) 463–484
27 van Minnen, J , Reichelt, D and Lodder, J C , An ultrastructural study of the neurosecretory Canopy cell of the pond snail *Lymnaea stagnalis*, with the use of the horseradish peroxidase tracer technique, *Cell Tissue Res* , 214 (1979) 453–462
28 Veldhuijzen, J P and Dogterom, G E , Incorporation of ^{14}C glucose in the polysaccharides of various body parts of the pond snail *Lymnaea stagnalis* as affected by starvation, *Neth J Zool* , 25 (1976) 125–137
29 Veldhuijzen, J P and van Beek, G , The influence of starvation and increased carbohydrate intake on the polysaccharide content of various body parts of the pond snail *Lymnaea stagnalis*, *Neth J Zool* , 26 (1976) 106–118
30 von Heijne, G , Patterns of amino acids near signal sequence cleavage sites, *Eur J Biochem* , 133 (1983) 17–21
31 Vreugdenhil, E , Geraerts, W P M , Jackson, J F and Joosse, J , The molecular basis of egg laying behavior in *Lymnaea stagnalis*, *Peptides*, 6, Suppl 3 (1985) 465–470
32 Wendelaar Bonga, S E , Ultrastructure and histochemistry of neurosecretory cells and neurohaemal areas in the pond snail *Lymnaea stagnalis*, *Z Zellforsch* , 108 (1970) 190–224