

Discovering neuropeptides in *Caenorhabditis elegans* by two dimensional liquid chromatography and mass spectrometry[☆]

Steven J. Husson^{*}, Elke Clynen, Geert Baggerman, Arnold De Loof, Liliane Schoofs

Laboratory of Developmental Physiology, Genomics and Proteomics, Zoological Institute, Katholieke Universiteit Leuven, Naamsestraat 59, B-3000 Leuven, Belgium

Received 9 July 2005
Available online 21 July 2005

Abstract

Completion of the *Caenorhabditis elegans* genome sequencing project in 1998 has provided more insight into the complexity of nematode neuropeptide signaling. Several *C. elegans* neuropeptide precursor genes, coding for approximately 250 peptides, have been predicted from the genomic database. One can, however, not deduce whether all these peptides are actually expressed, nor is it possible to predict all post-translational modifications. Using two dimensional nanoscale liquid chromatography combined with tandem mass spectrometry and database mining, we analyzed a mixed stage *C. elegans* extract. This peptidomic setup yielded 21 peptides derived from formerly predicted neuropeptide-like protein (NLP) precursors and 28 predicted FMRFamide-related peptides. In addition, we were able to sequence 11 entirely novel peptides derived from nine peptide precursors that were not predicted or identified in any way previously. Some of the identified peptides display profound sequence similarities with neuropeptides from other invertebrates, indicating that these peptides have a long evolutionary history.

© 2005 Elsevier Inc. All rights reserved.

Keywords: *Caenorhabditis elegans*; FMRFamide; FaRP; Neuropeptide; *flp*; *nlp*; Mass spectrometry; Two dimensional; Liquid chromatography; MudPit; Receptor

Neuropeptides are naturally occurring signaling molecules that interact with cell surface receptors to trigger an intracellular transduction pathway. Not only are they structurally diverse, their signaling cascades are highly variable, so there is a tremendous potential of different effects on living cells. Because of their critical signaling role, peptides are attractive for pharmaceutical use, since they can be used as therapeutics or as indirect targets via their cognate receptors or processing enzymes.

Current estimates suggest that as much as one-third of the human world population harbors one or more species of parasitic helminths [1]. In this context, the important role of neuropeptides in the biology of nematodes has captivated the interest of many parasitologists. Understanding the role of neuropeptides is unfortunately often hindered by the absence of primary sequence information and knowledge about their post-translational modifications. This sequence information has arrived very slowly, due to the huge efforts required in tissue collection and purification to ultimately isolate and functionally characterize a peptide. Although 23 FMRFamide-like peptide (*flp*) genes [2,3] and 32 neuropeptide-like protein (*nlp*) genes [4] have been predicted from the genome sequence of *Caenorhabditis elegans*, only 12 FMRFamide-related peptides (FaRPs) have been biochemically isolated and identified to date [5–10].

[☆] Abbreviations: 2D, two dimensional; nanoLC, nanoscale liquid chromatography; SCX, strong cation exchange; Q-TOF, quadrupole time-of-flight; MS/MS, tandem mass spectrometry; *nlp* (NLP), neuropeptide-like protein; *flp* (FLP), FMRFamide-like peptide; FaRP, FMRFamide-related peptide; MIP, myoinhibiting peptide; ACN, acetonitrile; TFA, trifluoroacetic acid; FA, formic acid.

^{*} Corresponding author. Fax: +32 16 32 39 02.

E-mail address: Steven.Husson@bio.kuleuven.be (S.J. Husson).

The concept for the simultaneous identification of naturally occurring peptides, derived from protein precursors in vivo within a peptide producing cell, differs from that of tryptic peptides, which in contrast are artificially derived from proteins on the laboratory bench. The first approach is called ‘peptidomics’ to clearly distinguish it from ‘proteomics.’ We recently introduced this concept of peptidomics, which aims at the simultaneous visualization and identification of all the expressed peptides with their post-translational modifications of a cell, tissue or organism [11]. In this way, the neuropeptidome of *Drosophila melanogaster* was already successfully explored and we demonstrated that a two dimensional-LC separation is more suited than a one dimensional-LC separation in order to identify the peptidome as complete as possible [12,13]. In this paper, we describe for the first time a high throughput peptidomic analysis of a nematode, i.e., *C. elegans*, by the use of specific extraction and sample preparation procedures, followed by two dimensional nanoscale liquid chromatography–quadrupole time-of-flight tandem mass spectrometry (2D-nanoLC-Q-TOF MS/MS).

Materials and methods

Animals. Wild-type *C. elegans* (variety Bristol N₂) were raised at 20 °C, under standard laboratory conditions, on conventional nematode growth media (NGM) plates, cultured with the *E. coli* strain OP-50.

Peptide extraction and sample preparation. Mixed stage worms from 10 fully grown petri dishes (diameter 90 mm), free from bacteria and dead animals by flotation on 30% sucrose, were placed in an ice-cold methanol/water/acetic acid (90:9:1, v/v/v) solution. After homogenization, sonication, and centrifugation at 500g for 15 min, the pellet was discarded. The methanol was evaporated and the remaining aqueous residue was delipidated by re-extraction with ethyl acetate and *n*-hexane. The organic solvent layer was decanted and the aqueous solution was subsequently desalted by using a SepPak C18 cartridge (Waters, Milford, MA, USA). The peptides were eluted with 50% acetonitrile (ACN) containing 0.1% trifluoroacetic acid (TFA). One-fifth of this fraction was further dried in a Speedvac concentrator, redissolved in 20 µL of 2% ACN with 0.1% formic acid (FA), and subsequently filtered through a 22 µm spinflow filter (Ultrafree-MC, Millipore, Bedford, USA) prior to analysis.

2D-nanoLC MS/MS. The two dimensional nanoscale liquid chromatography tandem mass spectrometry experiments were conducted on an Ultimate HPLC pump, a Switchos column-switching device, and a Famos autosampler (all LC-Packings, The Netherlands), coupled to a Q-TOF hybrid Quadrupole Time-Of-Flight mass spectrometer (Micromass, UK). Twenty microliters of the sample, corresponding to two fully grown NGM plates, was injected on a strong cation exchange (SCX) column (Bio-SCX, 500 µm × 15 mm, LC-Packings, The Netherlands), which was on-line with a C18 pre-column (µ-guard column MGU-30 C18, LC-Packings, The Netherlands). Sample loading was done at a flow rate of 30 µL/min with water containing 2% ACN and 0.1% FA. After loading the sample, the SCX column was switched off-line and the reversed phase pre-column was rinsed for 5 min. Next, this reversed phase trapping column was switched on-line with the nanoscale Atlantis C18 column (3 µm, 100 µm × 100 mm, Waters, USA). The peptides were separated using a 50 min gradient from 2 to 50% ACN containing 0.1% FA at a flow rate of 200 nL/min. The second fraction of

peptides was eluted from the SCX column by injection of 20 µL of a 20 mM ammonium acetate solution. The eluting peptides were again concentrated and desalted on the C18 pre-column prior to the nanoscale HPLC and MS analysis. This procedure was performed 10 times with successive concentrations of ammonium acetate (0, 20, 50, 100, 200, 400, 600, 800, 1000, and 2000 mM) that were used to elute peptide fractions from the SCX column. The 2D-LC system was directly connected with the electrospray interface of the Q-TOF mass spectrometer. The column eluent was directed through a stainless steel emitter (Proteon, Denmark) and nitrogen was used as nebulising gas. The mass spectrometer was set to automatic data-dependent MS to MS/MS switching when the intensity of the doubly and triply charged parent ions increased above 15 counts/s. The applied collision energy of the argon gas was chosen automatically between 25 and 40 eV depending on the number of charges and the mass range of the selected parent ion.

Peptide identification. The MS/MS data of all 10 SCX fractions were transformed into peak list files using the Proteinlynx software (Micromass, UK). These text files were subsequently submitted to a Mascot search on an in-house server to identify the peptides. This bioinformatics program matches the fragmentation data from the peak list files against any FASTA format protein database [14]. A home-made *C. elegans* database was constructed, containing the known FLP and NLP precursors. To identify possible unknown neuropeptide precursors, the NCBI nr database with taxonomy set on *C. elegans* was also applied. Pyroglutamic acid (Glu), carboxyterminal amidation, acetylation (aminoterminal, Lys), sulfatation (Tyr), and oxidation (Met), which are all common post-translational modifications of neuropeptides, were selected as variable modifications in the Mascot program. Individual ions with *Probability Based Mowse Scores* above the threshold required for identity or extensive homology ($p < 0.05$) were further manually explored. First, we checked whether the fragmentation spectrum shows extensive similarity to the theoretical fragment ions of the presumed peptide sequence. If this was the case, we further checked whether this peptide meets all requirements to correspond to a naturally occurring peptide. This means that the peptide must be flanked in the precursor by dibasic (or basic) cleavage sites or N-terminally flanked by the signal sequence, since this is the case for all presently known endogenous bioactive peptides. If the presumed peptide sequence carboxyterminally contains an amidation, this means that in the precursor sequence a glycine must be present as carboxyterminal amidation is always derived from a glycine. Next the presumed peptide precursor should contain a signal peptide, a common feature of all secreted proteins. Therefore, newly identified precursors were further analysed by SignalP 3.0 [15] to identify the signal sequence. The presence of a signal sequence, the presence of dibasic (or basic) cleavage sites amino- and carboxyterminally located from the peptide (which are used by the endogenously present convertases), and the presence of a glycine located aminoterminal from the cleavage place in case of carboxyterminal amidation were handled as three additional rules to identify endogenous, naturally occurring, peptides.

The few peptides with individual ion scores just below the indicated threshold, but having satisfying fragmentation spectra and meeting all requirements to correspond to an endogenous peptide, all originated from a precursor of which another peptide had already been successfully identified (a score above the threshold). This is a strong indication that these peptides, although below the threshold, do correspond to an endogenous peptide. These peptides are indicated by a footnote in Table 1 which lists all the presently identified *C. elegans* peptides.

Results

In the 2D-nanoLC experiment, different peptide fractions eluting from the strong cation exchange column in each salt step underwent a subsequent nanoscale HPLC separation on an analytical C18 column. Each

Table 1
C. elegans neuropeptides identified by 2D-nanoLC-MS/MS

Gene	Gene similarity	Peptides characterized by 2D-nanoLC-MS/MS
FMRFamide-related peptides or FaRPs		
-LRFamide family		
<i>flp-1</i>	<i>C. vulgaris</i> , <i>C. briggsae</i> , <i>P. redivivus</i> , myosuppressins	<i>SADPNFLRFamide</i> <i>AAADPNFLRFamide</i> <i>EIPGVLRamide</i> ^a <i>SEVPGVLRamide</i> <i>SYFDEKKSVPGVLRamide</i> <i>SVPGVLRamide</i> ^a <i>DFDGAMPGVLRamide</i> <i>GAMPGVLRamide</i>
<i>flp-18</i>	<i>C. briggsae</i>	
-IRFamide family		
<i>flp-5</i>	<i>C. briggsae</i>	<i>GAKFIRamide</i> <i>APEASPFIRamide</i> <i>AMDSPLIRamide</i> <i>ASPSAPLIRamide</i> ^a <i>SPSAVPLIRamide</i> <i>SAAAPLIRamide</i> <i>AADGAPLIRamide</i>
<i>flp-13</i>	<i>C. briggsae</i>	
-MRFamide family		
<i>flp-3</i>		<i>TPLGTMRamide</i> <i>EAEEPLGTMRamide</i> <i>SADDSAPFGTMRamide</i> <i>SAEPFGTMRamide</i> <i>ASEDALFGTMRamide</i> <i>NPENDTPFGTMRamide</i> <i>SPSAKWMRamide</i> <i>pQQDSEVEREMM</i>
<i>flp-22</i>		
<i>flp-6</i>		
-VRFamide family		
<i>flp-9</i>	<i>C. briggsae</i>	<i>KPSFVRamide</i> <i>ASGGMRNALVRFamide</i> <i>NGAPQPFVRFamide</i> ^a <i>SPLDEEDFAPESPLQamide</i> <i>AQTFVRFamide</i> <i>GQTFVRFamide</i> ^a <i>WANQVRFamide</i>
<i>flp-11</i>	<i>C. briggsae</i>	
<i>flp-16</i>	<i>C. briggsae</i>	
<i>flp-19</i>	<i>C. briggsae</i>	
Neuropeptide-like protein (NLP) peptides		
-MSFamide family		
<i>nlp-1</i>	<i>C. briggsae</i> , buccalin drosulfakinin-0	<i>MDANAFRMSFamide</i> <i>VNLDPNSFRMSFamide</i> <i>SAPSDFSRDIMSamide</i> <i>SSSMYDRDIMSFamide</i> ^a <i>SPVDYDRPIMAFamide</i> <i>LYLKQADFDDPRMFTSSFamide</i> ^a <i>MILPSLADLHRYTMYD</i>
<i>nlp-13</i>	<i>C. briggsae</i>	
<i>nlp-7</i>	<i>C. briggsae</i>	
-(F/M)G(L/F)amide family		
<i>nlp-6</i>	<i>C. briggsae</i> , allatostatins A	<i>APKQMVFGamide</i>
-GFxGF family		
<i>nlp-8</i>	<i>C. briggsae</i>	<i>YPYLIFPASPSGDSRRLV</i> <i>SFDRMGGTEFGLM</i> <i>ALNSLDGAGFGFE</i> <i>AFDSLAGSGFDNGFN</i>
<i>nlp-14</i>	<i>C. briggsae</i> , orcokinin	
<i>nlp-15</i>	<i>C. briggsae</i>	
-FAFA family		
<i>nlp-18</i>	<i>C. briggsae</i>	<i>SDEENLDFLE</i> <i>SPYRTFAFA</i> <i>SPYRAFAFA</i>

Table 1 (continued)

Gene	Gene similarity	Peptides characterized by 2D-nanoLC-MS/MS
<i>GGARAF- family</i>		
<i>nlp-21</i>	<i>C. briggsae</i>	pQYTSELEEDE GGARVFQGFED GGARAFLETEM TPIAEAQGAPEDVDDRRELE
<i>nlp-9</i>	<i>C. briggsae</i>	
<i>No multigene family</i>		
<i>nlp-11</i>	<i>C. briggsae</i>	SPAISPAYQFENAFGLSEALERAamide
<i>nlp-17</i>	<i>C. briggsae</i>	GSLSNMMRIamide
Newly characterized peptides		
<i>Novel FaRPs</i>		
<i>flp-24</i>	<i>C. briggsae</i>	VPSAGDMMVRFamide VPSAGDM(ox)MVRFamide VPSAGDMM(ox)VRFamide VPSAGDM(ox)M(ox)VRFamide
<i>flp-26</i>	<i>C. briggsae</i>	EFNADDLTLRFamide FNADDLTLRFamide GGAGEPLAFSPDMLSLRFamide
<i>Novel NLP peptides</i>		
<i>nlp-35</i>	<i>C. briggsae</i>	AVVSGYDNIYQVLAPRF
<i>nlp-36</i>	<i>C. briggsae</i>	SMVARQIPQTVVADH
<i>nlp-37</i>	<i>C. briggsae</i>	NNAEVVNHLKNFGALDRLGDVamide
<i>nlp-38/MIP</i>	Insect MIPs, B-type allatostatins	TPQNWNKLNLSLWamide SPAQWQRANGLWamide
<i>nlp-39</i>		EVPNFQADNVPEAGGRV
<i>nlp-40</i>	<i>C. briggsae</i>	APSAPAGLEEKL APSAPAGLEEKLR
<i>nlp-41</i>		APGLFELPSRSV

Peptides that were previously purified and sequenced are in italic face. pQ, pyroglutamic acid; ox, oxidation.

^a Peptides having Mowse scores just below the threshold required for identity or extensive homology ($p < 0.05$), but having satisfying fragmentation spectra and meeting all requirements to correspond to an endogenous peptide, as explained under Materials and methods.

nanoLC-MS analysis resulted in a complex total ion current (TIC) chromatogram, which is made up from individual mass spectra (Fig. 1). The nanoLC-MS/MS technology allows us to select and fragment multiple charged peptide ions as they elute from the column, even when co-eluting with other peptides. The amino acid sequences of the selected peptides can then be deduced from the resulting fragmentation spectra. In this way, we sequenced 60 endogenous peptides in total. Table 1 lists the sequences of all presently identified *C. elegans* peptides and indicates whether they display similarity to neuropeptides found in other invertebrates or vertebrates.

FMRFamide-related peptides or FaRPs

One of the most diverse and well-studied groups of neuropeptides in invertebrates are the FMRFamide-related peptides (FaRPs). All *C. elegans*-RFamide peptides possess a non-polar hydrophobic amino acid (L, I, M, V) in the third position from the carboxyl-terminus. We will now describe the identified peptides and place them in their respective families according to the carboxyterminal motifs.

-LRFamide family

The first discovered FMRFamide-like peptide gene of *C. elegans* was named *flp-1*. Its mRNA can be alternatively spliced [16]. The *Cv-flp-1* gene in *Caenorhabditis vulgaria* has 99.4% similarity with the *C. elegans flp-1* gene and it encodes the same putative peptides [17]. Disruption of the *flp-1* gene causes multiple behavioral defects, including uncoordination, hyperactivity, and insensitivity to high osmolarity [18]. Six of the seven -FLRFamide peptides, all contained within the putative translation products of the *flp-1* gene were already isolated from extracts of mixed stage animals in 1993 [10]. In our 2D-nanoLC-MS/MS experiment, we identified two of these formerly characterized neuropeptides, i.e., AAADPNFLRFamide and SADPNFLRFamide. The latter was initially isolated from *Panagrellus redivivus* [19].

The *flp-18* transcript can also be alternatively spliced yielding two peptide precursor proteins, one being the aminotermally extended form of the second. All five FaRPs from FLP-18 contain the carboxyterminal -PGVLRamide consensus sequence. We identified the peptides EIPGVLRamide, SEVPGVLRamide, SVPGVLRamide, DFDGAMPGVLRamide, and an aminotermally truncated form of the latter peptide

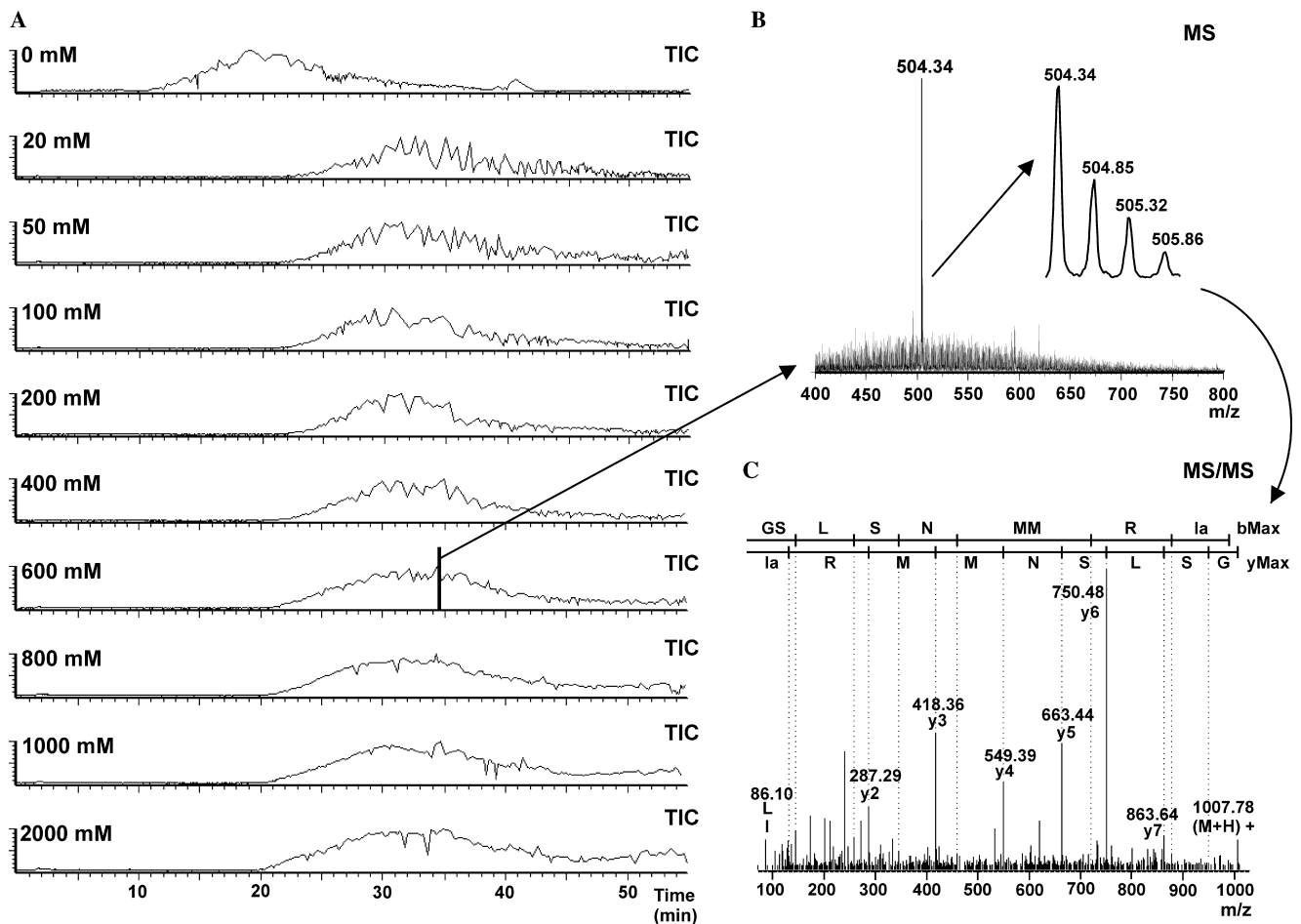


Fig. 1. Schematic overview of the 2D-nanoLC-MS/MS data. (A) Total ion current chromatograms (TIC) of all 10 SCX fractions obtained by a 2D-nanoLC separation of a *C. elegans* extract. (B) Example of one survey mass spectrum at 34.5 min of the 600 mM fraction, in which the doubly charged ion at m/z 504.34 is selected for fragmentation. (C) Fragmentation spectrum of the selected peptide, representing the sequence GSLSNMMRIamide (b-type and y-type ions are indicated in the spectrum, “a” indicates a carboxyterminal amidation).

(GAMPGVLRamide). An aminoterminally extended form of SVPGVLRamide, being SYFDEKKSVPGVLRamide, was also identified. SVPGVLRamide, the only peptide present in three copies, was previously characterized and appears to be excitatory on *Ascaris suum* somatic musculature [9].

-IRFamide family

The following two -FIRFamide containing peptides were identified in our experiment: GAKFIRFamide and APEASPFIRFamide, which are products of, respectively, *flp-5* and *flp-13*. Only APEASPFIRFamide was previously isolated and synthetic replicates of this peptide seem to induce a profound relaxation of both dorsal and ventral somatic muscle-strip preparations of *A. suum* [6]. The FLP-13 precursor contains nine putative peptides of which one displays the -FIRFamide motif. The other eight peptides have a carboxyterminal -LIRFamide sequence. Of these peptides, we confirmed the presence of AMDSPLIRFamide, ASPLIRFamide, SPSAVPLIRFamide, SAAAPLIRFamide, and

AADGAPLIRFamide. Only the latter has been isolated previously in *C. elegans* and appears to have an inhibitory effect on dorsal and ventral *A. suum* muscle strips [9]. According to the predictions of the potential FaRPs, SDRPTRAMDSPLIRFamide was also predicted as a FLP-13 derived peptide. However, we only found the short form AMDSPLIRFamide, indicating that the basic arginine at position six acts as a monobasic cleavage site.

-MRFamide family

Flp-3 encodes nine putative aminoterminally extended -GTMRFamide peptides, of which we were able to sequence six: TPLGTMRFamide, EAEPLGTMRFamide, SADDAPFGTMRFamide, SAEPFGTMRFamide, ASDALFGTMRFamide, and NPENDTPFGTMRFamide. None of these peptides were identified or biochemical characterized before, hence little is known about their function. *Flp-22* encodes three copies of the peptide SPSAKWMRFamide, which could also be observed in the present peptide profile. Some putative

FLP peptides without the -RFamide moiety were also sequenced. The presently identified sequence pQQDS-EVEREMM is for example aminoterminally preceded by the signal peptide and flanked at the carboxyl-terminus by KR in the FLP-6 precursor. This peptide has an aminoterminal pyroglutamic acid residue. The enzymes responsible for this modification are found in the secretory pathway, suggesting that the peptide is indeed secreted. A doubly charged ion which might correspond to the peptide KSAYMRFamide from FLP-6 was detected, though not selected for fragmentation, hence its sequence could not be confirmed.

-VRFamide family

Alternative splicing of *flp-11* yields three transcripts, whereby one transcript encodes three FaRPs while the others encode only two, due to a frame shift of the coding region. We were able to sequence the unique peptide NGAPQPFVRFamide derived from one transcript as well as ASGGMRNALVRFamide that is encoded by all three transcripts. We also identified another peptide derived from the *flp-11* precursor, which is not a FaRP, i.e., SPLDEEDFAPESPLQamide. This peptide is flanked on both sites by the dibasic amino acids KR, which are used as cleavage sites. The peptides GQTFVRFamide and AQTFVRFamide, products of the *flp-16* gene, and the peptides KPSFVRFamide and WANQVRFamide that originate from *flp-9* and *flp-19* respectively, were also identified.

Neuropeptide-like protein peptides

The *C. elegans* genome has been scanned for putative proteins that display structural properties of typical neuropeptide preproteins and/or are homologous with previously characterized invertebrate neuropeptides [4]. In this way, 32 *neuropeptide-like protein (nlp)* genes were identified, but none of the corresponding peptides have ever been de novo sequenced until now. In our 2D-nanoLC MS/MS experiment, we could confirm the occurrence of 21 predicted NLP peptides.

-MSFamide family

This family comprises neuropeptides with a carboxy-terminal -SFamide motif similar to *D. melanogaster* drosulfakinin-0 [20]. The NLP-1 precursor also resembles a part of the *Aplysia californica* buccalin precursor [21]. Two of the three predicted NLP-1 peptides could be identified in our experiment: MDANAFRMSFamide and VNLDPNSFRMSFamide. Six of the seven putative NLP-13 peptides have the carboxyterminal -DIMSFA-motif. We found three doubly charged ions of which the masses correspond to SSSMYDRDIMSFA-mide, SAPSDFSRDIMSFA-mide, and SPVDYDRPI-MAFA-mide, but only the latter two peptides were fully sequenced. An ion with a fragmentation spectrum

corresponding to MILPSLADLHRYTMYD was also found in our *C. elegans* extract. This peptide is flanked in the NLP-7 precursor by the dibasic cleavage site KR, but was not predicted in the bioinformatics study [4]. A peptide with the sequence QADFDDPRMFTS-SFamide was false predicted from the *nlp-7* gene, as we found an aminoterminally extended form, LYLK-QADFDD-PRMFTSSFA-mide that is aminoterminally flanked by the signal peptide, which is removed upon entrance in the secretory pathway.

-(F/M)G(L/F)amide family

NLP-5 and NLP-6 contain putative peptides with a carboxyterminal -MGLamide, -MGFA-mide or -FGFA-mide motif. We found a doubly charged ion of which the fragmentation spectrum corresponds to the sequence APKQMVFGFA-mide. This sequence is located directly carboxyterminally from the predicted signal peptide and is flanked by KR at its carboxyl-terminus in the NLP-6 precursor.

-GFxGF family

NLP-14 and NLP-15 putative peptides contain a carboxyterminal -GFxGF or -GFGF motif. We were able to identify the peptides ALNSLDGAGFGFE and AFD-SLAGSGFDNGFN from, respectively, NLP-14 and NLP-15. The predicted neuropeptides of NLP-14 all share the carboxyterminal -GFGF(E/D) sequence string with orcokinin, a myotropic neuropeptide from the crayfish *Orconectes limosus* [22]. YPYLIFPASPSSGD-SRRLV and SFDRMGTEFGFM, products from the *nlp-8* gene, were also proven to be present. The latter peptide contains a carboxyterminal motif FGLM of sequence identity with human substance P.

-FAFA family

The presently sequenced peptides SDEENLDFLE, SPYRAFAFA, and SPYRTFAFA are all products of the *nlp-18* gene, which is part of the -FAFA neuropeptide family that also contains *nlp-20*.

GGARAF- family

Nlp-9 and *nlp-21* encode putative peptides containing a aminoterminal GGARAF- or related motif, but there is no carboxyterminal conservation between these peptides. In our 2D-nanoLC-MS/MS experiment, we could confirm the existence of three peptides originating from NLP-21: pQYTSELEDE, GGARAFLETEM, and GGARVFQGFED. From the *nlp-9* gene, we were able to identify the sequence TPIAEAQGAPEDVDD-RRELE.

Other NLP peptides

Eight *nlp* genes (*nlp-4*, *nlp-10*, *nlp-11*, *nlp-12*, *nlp-16*, *nlp-17*, *nlp-19*, and *nlp-26*) could not be classified into gene families according to sequence homology. Two

abundant ions, a doubly and triply charged, were found in four successive salt steps. Their fragmentation spectra correspond to the sequence SPAISPAYQFENAFGL-SEALERAamide, which originates from the NLP-11 precursor. After transcription, translation, and processing, the *nlp-17* gene results in two predicted peptides of which we could identify GSLSNMMRIamide.

Newly characterized peptides and their putative novel peptide precursors

Besides the 28 FaRPs and 21 NLP peptides derived from predicted neuropeptide precursor genes, we discovered eleven entirely novel peptides that originate from nine different, so far unknown *C. elegans* peptide precursors (Table 2). To verify the sequences of the novel peptides, synthetic replicates were analyzed by MS/MS. Comparison of the resulting fragmentation spectra with

Table 2

Novel *C. elegans* precursors from which peptides were identified by 2D-nanoLC-MS/MS

FLP-24 3B343, C24A1.1 AAB70310
MLSSRTSSII LILAILVAIM AVAQCRNIQY DVEEMTPAEA
FRYAQWGEIP **HKEVPSAGDM MVRFGKRSI**

FLP-26 R173.4, XH649 NP 741827
MKVMFMLALLFSSSLVATSAF RLPFFFGAN EDFNSGLTKR
NYYESKPYKR **EFNADDLTLR FGRGGGAGEP LAFSPDMLSL**
RFGK

NLP-35 4K284, C33A12.2 NP 501687
MPRVSSFIVFLTFMVALAVSHAAVVSQYD NIYQVLAPRF
RRARMMSFDA EAPQYLQHL QNLKPRFRS V

NLP-36 3K579, B0464.3 NP 499088
MSVDLQKQLELADYLALAVWCIFFGVLFILSVIFNFVCI
KKDDDTALE RWGYKKNIDM KLGPFRSMV **ARQIPQTVVA**
DH

NLP-37 XC781, F48B9.4 NP 508397
MSSRISVSLLLAVVATMFFFTANVVDATPR SQGNMRYGN
SLPAYAPHVL YRFYNSRQFA PINKR**NNAEV VNHILKNFGA**
LDRLGDVGK

NLP-38/MIP 1N995, C01A2.7 NP 740928
MQLIHFIIVGLAMLISLSLAA SDDRVLGWNK AHGLWGKRSV
QEASQ**KRTP QNWNKLNSLW GKR**SASSFDD DYTENGDD
VTMLY**KRSPA QWQRANGLWG R**

NLP-39 1N53, C54C8.9 NP 493163
MKLLILFSLFAIFFGVIALD SPIPEFYSTG RSTRVPSHRH
RIPRLG**KREV PNFQADNVPE AGGRVRYVG** PPLKLVILI

NLP-40 1A489, Y74C9A.2 NP 490661
MKLVILLSFVATVAVFAAPS **APAGLEELR** ALQEQLYSLE
KENGVDVKQK EQPAAADTF LFGVPQKRMVA WQPMKRSMIN
EDSRAPLLHA IEARLAEVLR AGERLGVNPE EVLADLRARN
QFQ

NLP-41 C04H5.8 CAE17685
MLGLVCKILVVLCLSVLCLTVS**APGLFEL PSRSVRLIRS**
DPSAYDGYEN SFYRGYSDN QQFRFNSPQN W

Consensus cleavage sites and NCBI accession numbers are underlined, identified peptides are in bold face, signal peptides as predicted by SignalP 3.0 are in italic face.

the fragmentation data of the endogenous peptides from *C. elegans* confirmed the obtained sequences.

Novel FaRPs

Recently, systematic BLAST searches of nematode ESTs revealed eight novel putative *flp* sequelogs designated *flp-24* through *flp-31* [23]. Five of these (*flp-24*–*flp-28*) are present in *C. elegans*. By using our peptidomics setup, we independently were able to provide biochemical evidence for the presence of three FaRPs encoded by two of the newly identified *flp* genes. An abundant ion, whose fragmentation spectrum corresponds to VPSAGDMMVRFamide, was found in four successive salt steps (Fig. 2). Three other forms in which the methionines are oxidized were also identified. This novel FaRP from the VRFamide family is flanked by dibasic cleavage sites in the propeptide precursor which was designated FLP-24. Two peptides, EFNADDLT-LRFamide and GGAGEPLAFSPDMLSLRFamide, both ending in -LRFamide and originating from the same precursor FLP-26, were also identified by tandem mass spectrometry. An aminoterminally truncated form of the first peptide (FNADDLT-LRFamide) was also found (Fig. 3). This novel FaRP precursor gene (R173.4 or XH649) has structural features of a neuropeptide precursor gene as it encodes an aminoterminal signal peptide and three putative dibasic cleavage sites. The *Caenorhabditis briggsae* gene CBG10956 is highly similar (98.8%) to this newly annotated *C. elegans flp-26* gene and contains the putative peptides EFNADDLT-LRFamide and AGAGEPLAFSPDMLSLRFamide. The latter peptide only differs in one aminoterminal amino acid from the peptide found in *C. elegans*.

Novel neuropeptide-like peptides

Since the bioinformatics study of Nathoo et al. [4], which revealed 32 *nlp* genes, two additional *nlp* genes, *nlp-33* and *nlp-34*, have been identified by in silico analysis. In the present study, we could identify the novel peptides AVVSGYDNIYQVLAPRF, SMVARQIPQT VVADH, and NNAEVVNHILKNFGALDRLGDVamide. These potential neuropeptides share extensive sequence homology with the putative *C. briggsae* peptides AGVSGYDNIYQVLAPRF, SMVARQLPQ-TIEMD, and NNAEVVNHLLKNFGTLDRIGDVamide, respectively. The three *C. elegans* peptides originate from different precursors (4K284 or C33A12.2, 3K579 or B0464.3, and XC781 or F48B9.4), in which they are all flanked by conventional basic cleavage sites, or by a signal peptide at the aminoterminal. We propose to annotate them as neuropeptide-like protein precursors NLP-35, NLP-36, and NLP-37, respectively.

The two peptides TPQNWNKLNSLWamide and SPAQWQRANGLWamide were also found in the *C. elegans* extract and both originate from the gene 1N995 or C01A2.7. Because of the presence of a signal

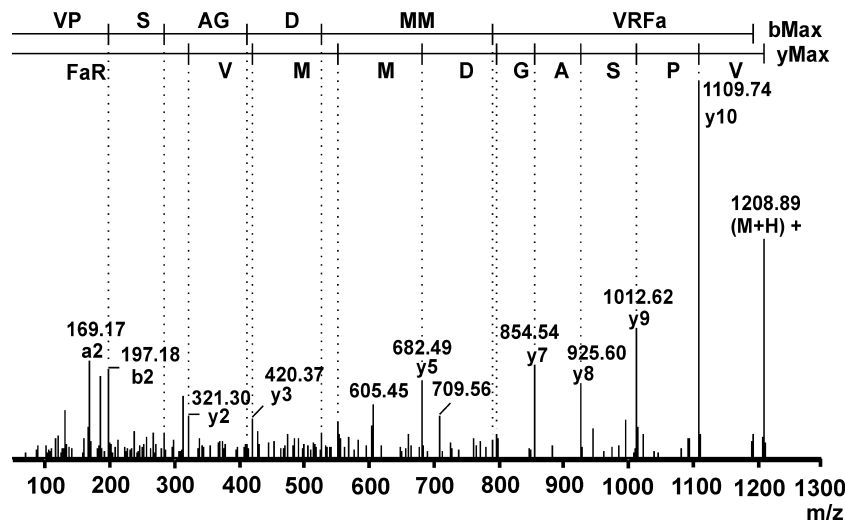


Fig. 2. Fragmentation spectrum of a novel FLP-24 peptide. The fragmentation spectrum of an ion at m/z 604.87 corresponds to the sequence VPSAGDMMVRFa. This novel FaRP originates from the precursor FLP-24 (b-type and y-type ions are indicated in the spectrum, “a” indicates a carboxyterminal amidation).

peptide and four dibasic cleavage sites (KR), this precursor most likely also represents a novel NLP precursor (NLP-38/MIP). TPQNWNKLSLWamide and SPAQWQRANGLWamide display profound similarity with neuropeptides belonging to the MIP family of insect neuropeptides. The MIPs (myoinhibiting peptides) were initially isolated and sequenced from an extract of 9000 locust brains [24], and later from CNS extracts of lepidopterans (*Manduca sexta* and *Bombyx mori*), crickets (*Gryllus bimaculatus*), and even from the stick insect *Carausius morosus* [25–28]. In crickets, MIPs inhibit juvenile hormone biosynthesis by the corpora allata and are therefore also known as B-type allatostatins. The MIP homologue in *B. mori* is known as Bom-PTSP (AWQDLNSAWamide), a prothoracicostatic peptide that inhibits ecdysteroid biosynthesis in the prothoracic glands. A MIP gene has been cloned from *D. melanogaster* [29] and two of the five predicted peptide sequences have been isolated and sequenced by means of a peptidomic experiment [12]. The *C. briggsae* gene CBG20435 has 79.6% similarity with *nlp-38* or the *C. elegans* MIP gene, and its translation product also contains TPQNWNKLSLWamide, in between dibasic cleavage places. All MIPs, including the novel *C. elegans* NLP-38/MIP peptides, share the carboxyterminal W(X)₆Wamide motif, suggesting that the biological activities exerted by this motif may be highly conserved (Table 3). The *C. elegans* MIPs also share the carboxyterminal -LWamide motif with a neuropeptide family from Hydrozoa. Metamorphosis A has the sequence pEQPGLWamide and was first isolated from the marine hydroid *Hydractinia echinata* [30]. Six -GLWamide containing peptides from *Hydra magnipapillata* were later isolated and sequenced. They are involved in metamorphosis and muscle contractions [31].

The protein precursor 1N53 or C54C8.9 has a signal peptide and two dibasic cleavage sites flanking the sequence EVPNFQADNVPEAGGRV. We confirmed the existence of this peptide in the *C. elegans* extract by tandem mass spectrometry. This precursor was designated as NLP-39.

An ion whose fragmentation spectrum corresponds to the peptide APSAPAGLEEKLR was also found in this study. The carboxyterminal extended form APSAPAGLEEKLR, which could be an intermediate form, not fully processed by carboxypeptidases, was also identified. These peptides originate from the 1A489 or Y74C9A.2 gene, which we named *nlp-40*. A highly similar putative neuropeptide APSAPGLEEKLR, which is missing the alanine in position six, is predicted to be present in *C. briggsae*.

The peptide APGLFELPSRSV was also found in the *C. elegans* extract and is located carboxyterminally from the signal peptide of the hypothetical protein C04H5.8, which therefore might be a possible neuropeptide precursor (NLP-41).

Discussion

Little is known about the cellular and molecular mechanisms whereby different neuropeptides exert their effects in *C. elegans*. Performing physiological tests to understand the role of neuropeptides is, however, often hindered by the absence of primary sequence information of the neuropeptides. One of the major advantages of using *C. elegans* as a model organism in research is the availability of its genomic sequence. Peptides can in principle be predicted from genomic sequences, but this remains even more difficult than for regular proteins,

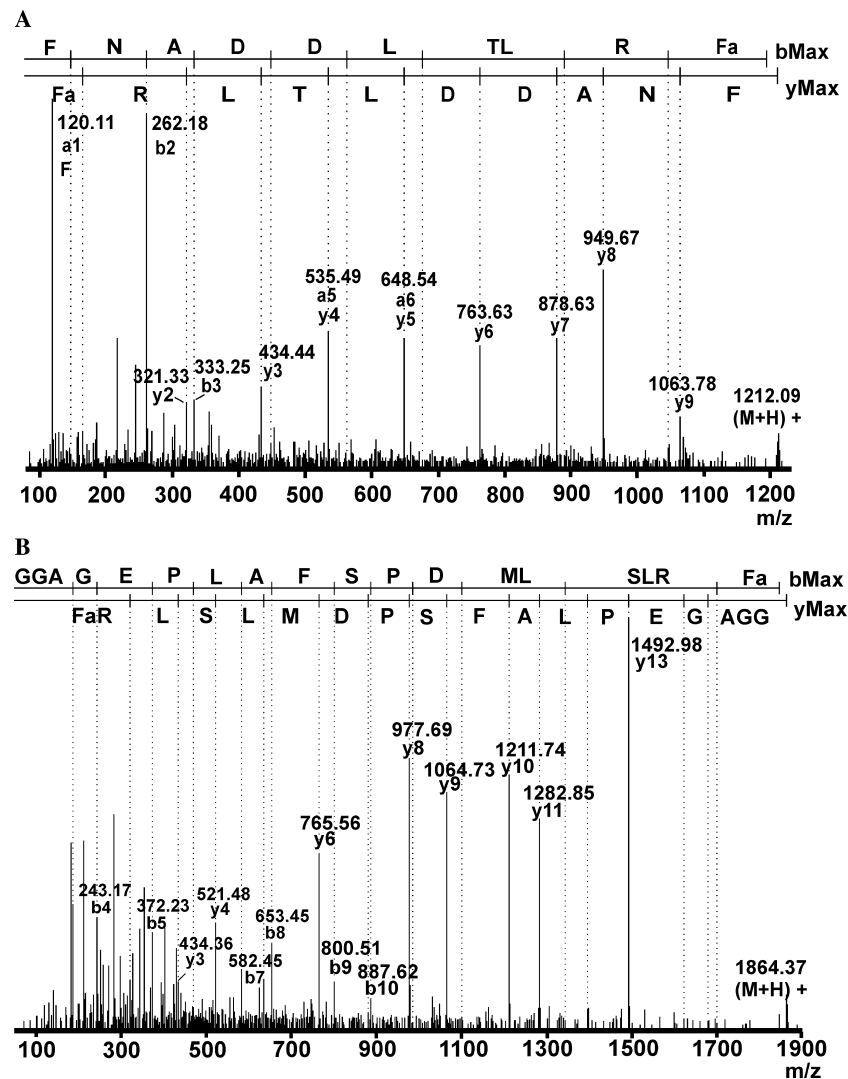


Fig. 3. Fragmentation spectrum and annotation of two novel FLP-26 peptides. (A) fragmentation spectrum of an aminoterminally truncated novel FLP-26 peptide, FNADDLTLRFamide. (B) Fragmentation spectrum of the second FaRP that originates from the novel FLP-26 precursor. The amino acid sequence was determined as GGAGEPLAFSPDMLSLRFamide (b-type and y-type ions are indicated in the spectrum, “a” indicates a carboxyterminal amidation).

where finding all exonic coding regions and splicing them together correctly in silico is already a challenge. Several search strings, such as a glycyl residue as a putative amidation site, dibasic cleavage sites (KR, KK or RK), the presence of a signal peptide and sequence homology with other known neuropeptides, can be taken into account when looking for novel neuropeptides. One can, however, not deduce whether all these peptides are actually expressed. In addition, processing of a neuropeptide precursor is regulated in time and space. Genomic sequences do not provide sufficient information to accurately predict the presence of post-translational modifications, which are often required for biological activity. Furthermore, non-conventional cleavages have been shown to exist in other invertebrates. Although the genomic sequence information has been available from 1998 and approximately 250 neuropeptides have been

predicted to be present in *C. elegans*, only 12 FaRPs from *C. elegans* have been isolated and sequenced until now, due to the large efforts needed for tissue collection and purification strategies [5–10].

By using a high throughput peptidomics approach, based on nanoscale liquid chromatography and tandem mass spectrometry, we could identify 28 predicted FaRPs, of which 21 were never sequenced before. In this study, we were also able to provide biochemical evidence for the presence of three novel FaRPs from *flp-24* and *flp-26*. Our mass spectrometry-based approach also revealed information about the post-translational processing like amidation and aminoterminal truncation of peptides, which of course could not be predicted by bioinformatics. So far, relatively few neuropeptides other than FaRPs have been isolated from helminths. However, we could confirm the presence of 21 neuropeptides

Table 3

Comparison of the primary structure of invertebrate myoinhibiting peptides (MIPs) or B-type allatostatins

Species	Designation	Peptide sequence
<i>C. morosus</i>	Cam-AST-B1	AWQDLQGGWamide
<i>C. morosus</i>	Cam-AST-B2	AWQDLNTGWamide
<i>C. morosus</i>	Cam-AST-B3	GWQDLQSSWamide
<i>C. morosus</i>	Cam-AST-B4	AWQDLQGAWamide
<i>C. morosus</i>	Cam-AST-B5	AWQDLQAGWamide
<i>C. morosus</i>	Cam-AST-B6	AWQDLGSAWamide
<i>G. bimaculatus</i>	Grb-AST-B1	GWQDLNGGWamide
<i>G. bimaculatus</i>	Grb-AST-B2	GWQDLNGGWamide
<i>G. bimaculatus</i>	Grb-AST-B3	AWQDLNGGWamide
<i>G. bimaculatus</i>	Grb-AST-B4	AWERFHGSWamide
<i>D. melanogaster</i>	Drm-MIP1	AWQSLQSSWamide
<i>D. melanogaster</i>	Drm-MIP2	AWQSMNVAWamide
<i>D. melanogaster</i>	Drm-MIP3	EAQGWKFRGAWamide
<i>D. melanogaster</i>	Drm-MIP4	EPTWNNLKGWamide
<i>D. melanogaster</i>	Drm-MIP5	DQWQKLHGGWamide
<i>Locusta migratoria</i>	Lom-MIP	AWQDLNAGWamide
<i>B. mori</i>	Bom-PTSP	AWQDLNSAWamide
<i>M. sexta</i>	Mas-MIP1	AWQDLNSAWamide
<i>M. sexta</i>	Mas-MIP2	GWQDLNSAWamide
<i>C. briggssae</i>		TPQNNWKLNSIWamide
<i>C. elegans</i>		TPQNNWKLNSIWamide
<i>C. elegans</i>		SPAQWQRANGIWamide

W and Wamide of the W(X)₆Wamide motif are highlighted.

from predicted *nlp* genes of *C. elegans*. None of them were de novo sequenced before. Eight additional peptides that originate from seven novel precursors were also identified. These precursors were so far not predicted as being putative peptide precursors. We annotated them as potential neuropeptide precursors NLP-35 till NLP-41. Some of their derived identified peptides display profound sequence similarities with insect neuropeptides like the B-type allatostatins or myoinhibiting peptides (MIPs). This indicates that the concept of neuropeptide signaling and their players is highly conserved. Most peptide receptors are members of the G-protein-coupled receptor (GPCR) family, and genes encoding predicted GPCRs are abundant in the *C. elegans* genome [32]. By means of a “reversed pharmacology approach,” it is possible to find the ligands of these “orphan” receptors (receptors of which no ligands are known) within a library of synthetic peptides. The success of such an approach is highly dependent on the correctness of the sequences of the peptides and their post-translational modifications. To date, only six *C. elegans* GPCRs were de-orphanised using this approach [33–38]. The peptide sequence information obtained by the present experiments will obviously stimulate the characterization of remaining orphan GPCRs in *C. elegans*.

Not all of the predicted neuropeptides were detected in our 2D-nanoLC-MS/MS experiments and this might be due to several reasons. First of all, many predicted neuropeptide precursors were in silico processed based on conventional cleavage sites. So we do not know whether or under which conditions these genes will be

expressed and processed. For example, three genes (*nlp-27*, *nlp-29*, and *nlp-31*), presumable all encoding antimicrobial peptides (not neuropeptides), show unique profiles of induction after fungal and bacterial infection [39]. It might also be possible that some peptide concentrations were beneath the sensitivity of our mass spectrometer, keeping in mind that we extracted whole worms, and not just its central nervous system, because this would be virtually impossible. Only ions with an intensity above a certain threshold were selected for fragmentation. Other ions which might correspond to the theoretical masses of predicted neuropeptides were not taken into account, because their sequence information could not be confirmed. Of course, there is also the possibility that some peptides were not properly extracted and not all peptides ionize with the same efficiency in the electrospray source. Nevertheless, this high throughput neuropeptide profiling study demonstrates the sensitivity and effectiveness of nanoLC combined with tandem mass spectrometry in peptide research. Since this nanoscale approach worked very well for an organism as tiny as *C. elegans*, we are convinced that virtually any organism of which the genome has been sequenced can be analyzed by this peptidomics technology.

Acknowledgments

This work was sponsored by project G.0270.04 and funded by a PhD grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). E. Clynen and G. Baggerman are postdoctoral researchers of the Fund for Scientific Research-Flanders (FWO-Vlaanderen).

References

- [1] T.A. Day, A.G. Maule, Parasitic peptides! The structure and function of neuropeptides in parasitic worms, *Peptides* 20 (1999) 999–1019.
- [2] C. Li, K. Kim, L.S. Nelson, FMRamide-related neuropeptide gene family in *Caenorhabditis elegans*, *Brain Res.* 848 (1999) 26–34.
- [3] K. Kim, C. Li, Expression and regulation of an FMRamide-related neuropeptide gene family in *Caenorhabditis elegans*, *J. Comp. Biol.* 475 (2004) 540–550.
- [4] A.N. Nathoo, R.A. Moeller, B.A. Westlund, A.C. Hart, Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species, *Proc. Natl. Acad. Sci. USA* 98 (2001) 14000–14005.
- [5] N.J. Marks, C. Shaw, A.G. Maule, J.P. Davis, D.W. Halton, P. Verhaert, T.G. Geary, D.P. Thompson, Isolation of AF2 (KHE-YLRFamide) from *Caenorhabditis elegans*: evidence for the presence of more than one FMRamide-related peptide-encoding gene, *Biochem. Biophys. Res. Commun.* 217 (1995) 845–851.
- [6] N.J. Marks, A.G. Maule, T.G. Geary, D.P. Thompson, J.P. Davis, D.W. Halton, P. Verhaert, C. Shaw, APEASPFIRFamide, a novel FMRamide-related decapeptide from *Caenorhabditis elegans*: structure and myoactivity, *Biochem. Biophys. Res. Commun.* 231 (1997) 591–595.

- [7] N.J. Marks, A.G. Maule, T.G. Geary, D.P. Thompson, C. Li, D.W. Halton, C. Shaw, KSAYMRamide (PF3/AF8) is present in the free-living nematode, *Caenorhabditis elegans*, *Biochem. Biophys. Res. Commun.* 248 (1998) 422–425.
- [8] N.J. Marks, A.G. Maule, C. Li, L.S. Nelson, D.P. Thompson, S. Alexander-Bowman, T.G. Geary, D.W. Halton, P. Verhaert, C. Shaw, Isolation, pharmacology and gene organization of KPSFVRamide: a neuropeptide from *Caenorhabditis elegans*, *Biochem. Biophys. Res. Commun.* 254 (1999) 222–230.
- [9] N.J. Marks, C. Shaw, D.W. Halton, D.P. Thompson, T.G. Geary, C. Li, A.G. Maule, Isolation and preliminary biological assessment of AADGAPLIRamide and SVPGLRamide from *Caenorhabditis elegans*, *Biochem. Biophys. Res. Commun.* 286 (2001) 1170–1176.
- [10] M.L. Rosoff, K.E. Doble, D.A. Price, C. Li, The *flp-1* propeptide is processed into multiple, highly similar FMRamide-like peptides in *Caenorhabditis elegans*, *Peptides* 14 (1993) 331–338.
- [11] E. Clynen, G. Baggerman, D. Veelaert, A. Cerstiaens, D. Van Der Horst, L. Harthoorn, R. Derua, E. Waelkens, A. De Loof, L. Schoofs, Peptidomics of the pars intercerebralis-corpus cardiacum complex of the migratory locust, *Locusta migratoria*, *Eur. J. Biochem.* 268 (2001) 1929–1939.
- [12] G. Baggerman, A. Cerstiaens, A. De Loof, L. Schoofs, Peptidomics of the larval *Drosophila melanogaster* central nervous system, *J. Biol. Chem.* 277 (2002) 40368–40374.
- [13] G. Baggerman, K. Boonen, P. Verleyen, A. De Loof, L. Schoofs, Peptidomic analysis of the larval *Drosophila melanogaster* central nervous system by 2 dimensional capillary liquid chromatography Q-ToF mass spectrometry, *J. Mass Spectrom.* 40 (2004) 250–260.
- [14] D.N. Perkins, D.J. Pappin, D.M. Creasy, J.S. Cottrell, Probability-based protein identification by searching sequence databases using mass spectrometry data, *Electrophoresis* 20 (1999) 3551–3567.
- [15] J.D. Bendtsen, H. Nielsen, G. von Heijne, S. Brunak, Improved prediction of signal peptides: SignalP 3.0, *J. Mol. Biol.* 340 (2004) 783–795.
- [16] M.L. Rosoff, T.R. Burglin, C. Li, Alternatively spliced transcripts of the *flp-1* gene encode distinct FMRamide-like peptides in *Caenorhabditis elegans*, *J. Neurosci.* 12 (1992) 2356–2361.
- [17] K. Schinkmann, C. Li, Comparison of two *Caenorhabditis* genes encoding FMRamide (Phe-Met-Arg-Phe-NH₂)-like peptides, *Brain Res. Mol. Brain Res.* 24 (1994) 238–246.
- [18] L.S. Nelson, M.L. Rosoff, C. Li, Disruption of a neuropeptide gene, *flp-1*, causes multiple behavioral defects in *Caenorhabditis elegans*, *Science* 281 (1998) 1686–1690.
- [19] T.G. Geary, D.A. Price, J.W. Bowman, C.A. Winterrowd, C.D. Mackenzie, R.D. Garrison, J.F. Williams, A.R. Friedman, Two FMRamide-like peptides from the free-living nematode *Panagrellus redivivus*, *Peptides* 13 (1992) 209–214.
- [20] R. Nichols, S.A. Schneuwly, J.E. Dixon, Identification and characterization of a *Drosophila* homologue to the vertebrate neuropeptide cholecystokinin, *J. Biol. Chem.* 263 (1988) 12167–12170.
- [21] E.C. Cropper, M.W. Miller, R. Tenenbaum, M.A. Kolks, I. Kupfermann, K.R. Weiss, Structure and action of buccalin: a modulatory neuropeptide localized to an identified small cardioactive peptide-containing cholinergic motor neuron of *Aplysia californica*, *Proc. Natl. Acad. Sci. USA* 85 (1988) 6177–6181.
- [22] J. Strangier, C. Hilbich, S. Burdzik, R. Keller, Orcokinin: a novel myotropic peptide from the nervous system of the crayfish, *Orconectes limosus*, *Peptides* 13 (1992) 859–864.
- [23] P. McVeigh, S. Leech, J.R. Mair, N.J. Marks, T.G. Geary, A.G. Maule, Analysis of FMRamide-like peptide (FLP) diversity in phylum Nematoda, *Int. J. Parasitol.* (2005), in press.
- [24] L. Schoofs, G.M. Holman, T.K. Hayes, R.J. Nachman, A. De Loof, Isolation, identification and synthesis of locustamyoinhibiting peptide (LOM-MIP), a novel biologically active neuropeptide from *Locusta migratoria*, *Regul. Pept.* 36 (1991) 111–119.
- [25] M.B. Blackburn, R.M. Wagner, J.P. Kochansky, D.J. Harrison, P. Thomas-Laemont, A.K. Raina, The identification of two myoinhibitory peptides, with sequence similarities to the galanins, isolated from the ventral nerve cord of *Manduca sexta*, *Regul. Pept.* 57 (1995) 213–219.
- [26] Y.J. Hua, Y. Tanaka, K. Nakamura, M. Sakakibara, S. Nagata, H. Kataoka, Identification of a prothoracicostatic peptide in the larval brain of the silkworm *Bombyx mori*, *J. Biol. Chem.* 274 (1999) 31169–31173.
- [27] M.W. Lorenz, R. Kellner, K.H. Hoffmann, A family of neuropeptides that inhibit juvenile hormone biosynthesis in the cricket *Gryllus bimaculatus*, *J. Biol. Chem.* 270 (1995) 21103–21108.
- [28] M.W. Lorenz, R. Kellner, K.H. Hoffmann, G. Gäde, Identification of multiple peptides homologous to cockroach and cricket allatostatins in the stick insect *Carausius morosus*, *Insect Biochem. Mol. Biol.* 30 (2000) 711–718.
- [29] M. Williamson, C. Lenz, A.M. Winther, D.R. Nassel, C.J. Grimmelikhuijzen, M.E. Winther, Molecular cloning, genomic organization, and expression of a B-type (cricket-type) allatostatin preprohormone from *Drosophila melanogaster*, *Biochem. Biophys. Res. Commun.* 281 (2001) 544–550.
- [30] T. Leitz, K. Morand, M. Mann, Metamorphosin A: a novel peptide controlling development of the lower metazoan *Hydra-ctinia echinata* (Coelenterata, Hydrozoa), *Dev. Biol.* 163 (1994) 440–446.
- [31] T. Takahashi, Y. Muneoka, J. Lohmann, M.S. Lopez de Haro, G. Solleder, T.C. Bosch, C.N. David, H.R. Bode, O. Koizumi, H. Shimizu, M. Hatta, T. Fujisawa, T. Sugiyama, Systematic isolation of peptide signal molecules regulating development in hydra: LWamide and PW families, *Proc. Natl. Acad. Sci. USA* 94 (1997) 1241–1246.
- [32] C.I. Bargmann, Neurobiology of the *Caenorhabditis elegans* genome, *Science* 282 (1998) 2028–2033.
- [33] T.M. Kubiak, M.J. Larsen, M.R. Zantello, J.W. Bowman, S.C. Nulf, D.E. Lowery, Functional annotation of the putative orphan *Caenorhabditis elegans* G-protein-coupled receptor C10C6.2 as a FLP15 peptide receptor, *J. Biol. Chem.* 278 (2003) 42115–42120.
- [34] T.M. Kubiak, M.J. Larsen, S.C. Nulf, M.R. Zantello, K.J. Burton, J.W. Bowman, T. Modric, D.E. Lowery, Differential activation of “social” and “solitary” variants of the *Caenorhabditis elegans* G protein-coupled receptor NPR-1 by its cognate ligand AF9, *J. Biol. Chem.* 278 (2003) 33724–33729.
- [35] C. Rogers, V. Reale, K. Kim, H. Chatwin, C. Li, P. Evans, M. De Bono, Inhibition of *Caenorhabditis elegans* social feeding by FMRamide-related peptide activation of NPR-1, *Nat. Neurosci.* 6 (2003) 1178–1185.
- [36] I. Mertens, A. Vandingenen, T. Meeusen, T. Janssen, W. Luyten, R.J. Nachman, A. De Loof, L. Schoofs, Functional characterization of the putative orphan neuropeptide G-protein coupled receptor C26F1.6 in *Caenorhabditis elegans*, *FEBS Lett.* 573 (2004) 55–60.
- [37] I. Mertens, A. Vandingenen, E. Clynen, R.J. Nachman, A. De Loof, L. Schoofs, Characterization of an RFamide-related peptide orphan GPCR in *C. elegans*, *Ann. NY Acad. Sci.* 1040 (2005) 410–412.
- [38] I. Mertens, T. Meeusen, T. Janssen, R.J. Nachman, L. Schoofs, Molecular characterization of two G protein-coupled receptor splice variants as FLP2 receptors in *Caenorhabditis elegans*, *Biochem. Biophys. Res. Commun.* 330 (2005) 967–974.
- [39] C. Couillault, N. Pujol, J. Reboul, L. Sabatier, J.F. Guichou, Y. Kohara, J.J. Ewbank, TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM, *Nat. Immunol.* 5 (2004) 488–494.