ELSEVIER

Progress in Neurobiology 82 (2007) 33–55

progress in Neurobiology

www.elsevier.com/locate/pneurobio

Neuropeptidergic signaling in the nematode *Caenorhabditis elegans*

Steven J. Husson *, Inge Mertens, Tom Janssen, Marleen Lindemans, Liliane Schoofs

Functional Genomics and Proteomics Unit, Department of Biology, Katholieke Universiteit Leuven, Naamsestraat 59, B-3000 Leuven, Belgium

Received 31 August 2006; received in revised form 14 December 2006; accepted 29 January 2007

Abstract

The nematode *Caenorhabditis elegans* joins the menagerie of behavioral model systems next to the fruit fly *Drosophila melanogaster*, the marine snail *Aplysia californica* and the mouse. In contrast to *Aplysia*, which contains 20,000 neurons having cell bodies of hundreds of microns in diameter, *C. elegans* harbors only 302 tiny neurons from which the cell lineage is completely described, as is the case for all the other somatic cells. As such, this nervous system appears at first sight incommensurable with those of higher organisms, although genome-wide comparison of predicted *C. elegans* genes with their counterparts in vertebrates revealed many parallels. Together with its short lifespan and ease of cultivation, suitability for high-throughput genetic screenings and genome-wide RNA interference approaches, access to an advanced genetic toolkit and cell-ablation techniques, it seems that this tiny transparent organism of only 1 mm in length has nothing to hide. Recently, highly exciting developments have occurred within the field of neuropeptidergic signaling in *C. elegans*, not only because of the availability of a sequenced genome since 1998, but especially because of state of the art post genomic technologies, that allow for molecular characterization of the signaling molecules. Here, we will focus on endogenous, bioactive (neuro)peptides and mainly discuss biosynthesis, peptide sequence information, localization and G-protein coupled receptors of the three major peptide families in *C. elegans*.

Keywords: Nematode; *Caenorhabditis elegans*; Neuropeptide; Insulin; FMRFamide-like peptide; *flp*; Neuropeptide-like protein gene; *nlp*; G-protein coupled receptor

Contents

Abbreviations: cGMP, 3′,5′-cyclic guanosine monophosphate; 5-HT, serotonin; AC, adenylyl cyclase; ACE, angiotensin-converting enzyme; ACh, acetylcholine; CAPS, Ca^{2+}-dependent activator protein for secretion; CELPC2, *C. elegans* proprotein convertase 2; CHO, Chinese hamster ovary; CPE, carboxypeptidase E; CNC, caenacins; DAG, diacylglycerol; DCV, dense core vesicle; FaRP, FMRFamide-related peptide; *flp*, FMRFamide-like peptide gene; GABA, γ-aminobutyric acid; GC, guanylate cyclase; GIRK, G-protein-regulated inward-rectifier K^+ channel; GPCR, G-protein coupled receptor; HEK, human embryonic kidney; HPLC, high performance liquid chromatography; IP_3, 1,4,5-trisphosphate; *ins*, insulin-like peptide gene; KPC, kex2/subtilisin-like proprotein convertase; PAL, peptidyl hydroxyglycine α-amidating lyase; PAM, peptidylglycine α-amidating monooxygenase; PC, proprotein convertase; PHM, peptidylglycine α-hydroxylating monooxygenase; PIP_2, phosphatidylinositol 4,5-bisphosphate; RIA, radioimmunoassay; RNAi, RNA interference; SV, synaptic vesicle; NEP, neprilysin; *nlp*, neuropeptide-like protein gene

* Corresponding author. Tel.: +32 16324260; fax: +32 16323902.
E-mail address: steven.husson@bio.kuleuven.be (S.J. Husson).

0301-0082/$ – see front matter
doi:10.1016/j.pneurobio.2007.01.006

1. Introduction

Nematodes include major parasites of livestock, plants and humans in addition to free-living species such as the model organism *Caenorhabditis elegans*. The nervous system is exceptionally well defined in *C. elegans* and contains about one-third of all somatic cells. These 302 neuronal cells with 5000 synapses probably dominate a significant portion of gene circuits involved in regulatory behaviors vital to the worm's survival. Obviously, worms display a significantly less complex behavior than vertebrates. Despite the fact that *C. elegans* has a quite simple anatomy of only 959 somatic cells, it displays diverse behaviors like chemotaxis, thermotaxis, oxygen sensing, osmotic avoidance, associative learning, etc. Genome-wide comparison of predicted *C. elegans* genes with their counterparts in vertebrate nervous systems revealed many striking parallels (Bargmann, 1998; The *C. elegans* Sequencing Consortium, 1998). The major neurotransmitter biosynthetic machinery, release mechanisms and receptors are conserved. Like all other metazoan organisms, *C. elegans* uses small molecule neurotransmitters such as acetylcholine (ACh), γ-aminobutyric acid (GABA) and nitric oxide (NO); excitatory amino acids such as glutamate; and biogenic amines such as octopamine, tyramine, serotonin (5-HT) and dopamine (Brownlee and Fairweather, 1999) which are packaged into synaptic vesicles (Gasnier, 2000) and are released by exocytosis (Weimer and Jorgensen, 2003; Scalettar, 2006). In addition to these small molecule neurotransmitters, the *C. elegans* genome encodes a wide variety of bioactive peptides, which can be subdivided into three major families according to their conserved motifs. The best studied neuropeptide group in nematodes is the FMRFamide-like peptide (*flp*) gene family. A second family encloses the *ins* genes which encode insulin-like peptides, and finally, peptides without sequence resemblance with FMRFamide or insulin are derived from the so-called neuropeptide-like protein (*nlp*) genes.

Furthermore, heterotrimeric GTP-binding protein (G-protein) coupled second messenger pathways are highly conserved

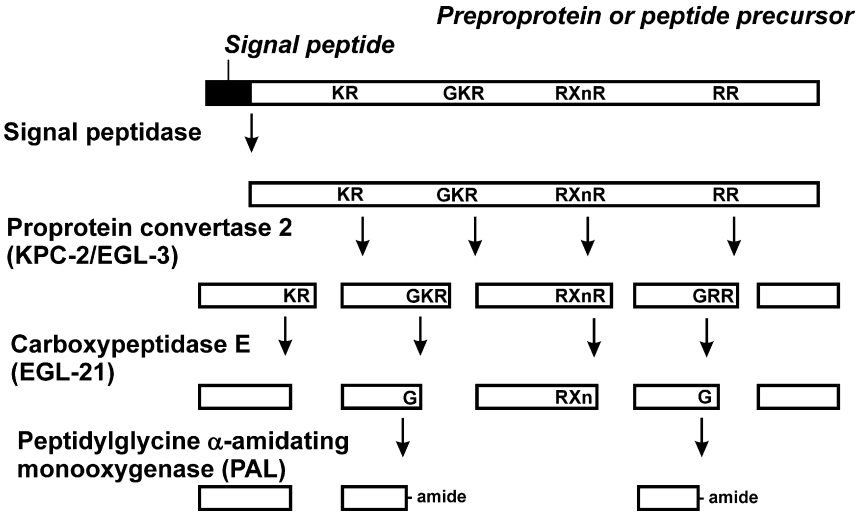

Fig. 1. Schematic representation of the neuropeptide processing pathway. A typical proprotein (neuro)peptide precursor contains an aminoterminally located signal peptide sequence that drives translocation of the precursor into the secretory pathway. The signal peptide is cleaved off by the specific action of a signal peptidase in the lumen to the endoplasmatic reticulum. Next, proprotein convertases (PCs) cleave the remaining part of the precursor at specific cleavage sites composed of basic amino acids like lysine-arginine (KR), arginine-arginine (RR) or arginine-X*n*-arginine (RX*n*R) where *n* is 2, 4, 6 or 8. Carboxypeptidases trim the carboxyterminal basic amino acids from the intermediate peptides after PC-cleavage. Finally, the carboxyterminal glycine residue, if present, is converted into an amide by peptidyl glycine α-amidating monooxygenase. This post-translational modification is a common feature for all secreted, bioactive peptides.

between *C. elegans* and mammals. Over 1100 G-protein coupled receptors (GPCRs) of which 50–60 are predicted to encode peptide receptors are contained within the genomic sequence, counting for about 5% of all *C. elegans* genes (Bargmann, 1998). It is clear that a rich diversity of signaling molecules and receptors underlies the seeming simplicity of the nematode nervous system and the apparent limited behavioral repertoire. Using a wide array of post genomic technologies and exploiting the unique features of this model organism to analyse genes from their sequence to their role in neuronal signaling, a lot of information has become available on peptidergic signal transduction in *C. elegans*. Here, recent progress on the identification and isolation of neuropeptides, as well as their biosynthesis, localization, functions and their G-protein coupled receptors, is explored and discussed.

2. From precursor to fully processed bioactive peptide

2.1. Peptide biosynthesis

All endogenous bioactive peptides are derived from large proprotein peptide precursors or prohormone genes that encode for a single or multiple neuropeptides. The proprotein peptide precursor undergoes extensive posttranslational processing prior to producing final neuropeptides (Fig. 1). A typical preproprotein or peptide precursor contains an aminoterminal signal peptide, which is cleaved off upon entry into the secretory pathway by a signal peptidase. Subsequently, proprotein convertases (PCs) cleave the remaining part of the precursor at specific cleavage sites. These endopeptidases typically cleave at pairs of basic amino acids, usually KR and RR, while RK and KK are found in lower frequency (Veenstra, 2000; Fricker, 2005). On occasion, the two basic residues are separated from each other by 2, 4, 6 or 8 other residues and were earlier described as "monobasic" cleavage places (Veenstra, 2000). Seven proprotein convertases (furin, PC1/3, PC2, PC4, PC5/6, PC7/8/LPC and PACE4) displaying sequence similarity with Kex2p from yeast (Julius et al., 1984), which is the prototype member of the PC family, have been described in mammals (Rouille et al., 1995; Beinfeld, 1998; Canaff et al., 1999). PC2 and PC1/3 specifically recognize substrates composed of dibasic residues, reflecting their role in the processing of neuropeptide precursors (Rouille et al., 1995; Beinfeld, 1998; Canaff et al., 1999). PC2 is the only member of this family whose activation requires intracellular interaction with a helper protein, the neuroendocrine protein 7B2 (Mbikay et al., 2001). 7B2 is biosynthesized as a precursor protein that is cleaved into an aminoterminal fragment and a carboxyterminal peptide. The aminoterminal fragment, which contains a proline-rich sequence, functions as a specific chaperone for the proteolytic maturation and activation of PC2 (Muller et al., 1999). The carboxyterminal peptide can inhibit PC2 in vitro and may contribute to keep the enzyme transiently inactive in vivo. Interestingly, unlike PC2-null mice, which are viable (Furuta et al., 1997), 7B2-null mutants die early in life from Cushing's disease, suggesting a possible involvement of 7B2 in secretory granule formation and in secretion regulation (Westphal et al., 1999).

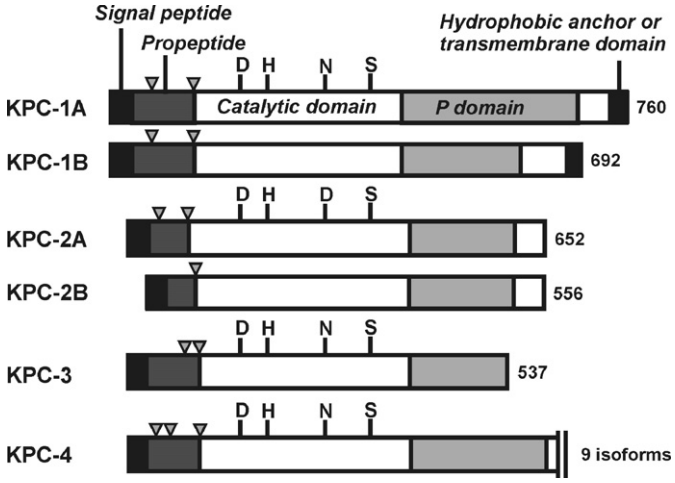

Fig. 2. Representation of the structures of the 4 *kpc* gene products of *C. elegans*. Only the KPC designation has been used. Numbers at the end of the protein structures indicate the length of the products. All four members of the kex2/subtilisin-like proprotein convertase family are transported into the secretory pathway via the recognition of a signal peptide. The enzymes are synthesized as inactive precursor proteins or zymogens and the aminoterminal propeptide has to be removed upon activation. Cleavage sites for this autocatalytical event are indicated with inverted triangles. The conserved subtilisin-like catalytical domain and the three conserved amino acids aspartic acid (D), histidine (H) and serine (S) are absolutely required for enzymatic activity. A conserved asparagine (N) residue, which stabilizes an oxyanion intermediate, is substituted for an aspartic acid (D) in KPC-2/EGL-3. Next to the catalytic domain is a region that is solely conserved within all members of the KPC family, and has been termed the P-domain. The carboxyterminal region is unique for each enzyme and appears to be the least conserved domain. The nine isoforms of the KPC-4 protein all differ in their carboxytermini and are not shown.

A survey of the *C. elegans* genome sequence has revealed the existence of four genes that encode members of the kex2/subtilisin-like proprotein convertase (KPC) family (Gomez-Saladin et al., 1994, 1997; Thacker et al., 1995; The *C. elegans* Sequencing Consortium, 1998; Thacker and Rose, 2000; Kass et al., 2001) (Fig. 2). *kpc-1* was originally named CelfurPC, based on its sequence similarity with that of human furin (Gomez-Saladin et al., 1997). The two *kpc-1* isoforms that can be generated by alternative splicing are most similar to mammalian furin, PC5/6 and PACE4. At present, however, a possible function for *C. elegans kpc-1* has not yet been described. The *kpc-2/egl-3* gene encodes a PC2-like convertase that was originally named CELPC2, based on sequence analysis of cDNA clones that were isolated by degenerate PCR analysis (Gomez-Saladin et al., 1994). As the name implies, the gene product displays sequence similarities with the mammalian neuroendocrine convertase PC2. Because different alleles of this *C. elegans* gene were initially identified in a mutant screen for deficiency in egg-laying, the name *egl-3* was introduced (Trent et al., 1983). Evidence that both genes, *egl-3* and *kpc-2*, are identical was provided by the rescue of several *egl-3* alleles with a cosmid clone that carries the CELPC2 gene (*kpc-2*) (Kass et al., 2001). The sequence similarity between EGL-3 and mammalian PC2 as well as the restricted expression of *egl-3* in neuronal structures such as the nerve ring, ventral cord and ganglia in the tail (Kass et al., 2001)

suggests that KPC-2/EGL-3 is required for the processing of neuropeptide precursors in *C. elegans*. This has recently been demonstrated by a mass spectrometric peptide profiling study which revealed a drastic reduction of peptides in different *egl-3* mutant strains, indicating that this enzyme is the major proprotein convertase needed for the proper processing of both FLP and NLP precursors (Husson et al., 2006). Mutations in *kpc-3/aex*-5, which were first identified in a genetic screen for specific defecation phenotypes (Thomas, 1990), can also affect the peptide profile, although to a lesser extent as was the case for *kpc-2/egl-3* (Husson et al., 2006). As *kpc-3/aex-5* displays limited sequence conservation with other members of the PC family, it is difficult to predict a possible role in neuropeptide processing. Which part KPC-3/AEX-5 plays in the defecation process is not yet known. *kpc-4/bli-4* represents one of the more complex *C. elegans* genes, encoding at least nine isoforms, named A to I (Thacker et al., 1999). The *kpc-4/bli-4* locus was originally identified by a single viable mutation, *e937*, resulting in a blistered adult cuticle (Brenner, 1974). Other mutations in *kpc-4/bli-4* result in an arrest of nematode development in late embryogenesis.

The *C. elegans* genome also contains a gene that encodes a protein that resembles the vertebrate 7B2 neuroendocrine helper protein (Lindberg et al., 1998). The *C. elegans* 7B2 (*sbt-1*) has a proline-rich sequence motif in the aminoterminal region as well as a highly conserved carboxyterminal peptide. The aminoterminal domain of *C. elegans* 7B2 facilitates the activation of vertebrate proPC2 as analyzed by in vitro assays. Moreover, a significant inhibitory activity against recombinant vertebrate PC2 could also be observed, indicating that both functional domains within 7B2 have been conserved between the *C. elegans* and the vertebrate orthologous proteins. A direct involvement of *C. elegans* 7B2 with processing of FLP and NLP precursors has recently been demonstrated by mass spectrometry (Husson et al., manuscript submitted).

After cleavage from the preproprotein or peptide precursor, the resulting peptide-processing intermediates contain carboxyterminal basic amino acids, which are further processed by a specific carboxypeptidase. In the screen for egg-laying defective mutants of *C. elegans*, strains with mutations in *egl-21* were isolated (Trent et al., 1983). Animals with mutated *egl-21* genes also displayed an impaired defecation cycle and showed uncoordinated locomotion, a phenotype similar to the one observed for *egl-3* PC2 mutants (Trent et al., 1983; Kass et al., 2001). Therefore, the genome sequence in the *egl-21* region was scanned for genes that could play a role in neuropeptide processing or secretion (Jacob and Kaplan, 2003). This revealed a gene (F01D4.4) having sequence similarity with vertebrate carboxypeptidase E. As a decreased immunoreactivity for anti-RFamide antibodies could be observed in *C. elegans* strains with mutated *egl-21* genes, a role for CPE/EGL-21 in the processing of FMRFamide-like neuropeptide precursors was suggested. Mass spectrometric analysis of *egl-21* strains strengthens this hypothesis since only extended intermediate peptides with the basic amino acids still attached at the carboxyterminus could be sequenced (Husson et al., manuscript submitted). In addition to the proprotein convertase and carboxypeptidase processing steps, most endogenous peptides require further posttranslational events such as carboxyterminal amidation. In order to become amidated, the peptide must have a carboxyterminal glycine residue that is transformed into an amide by two distinct enzymes, peptidylglycine α-hydroxylating monooxygenase (PHM) and peptidyl hydroxyglycine α-amidating lyase (PAL). In higher organisms, these activities are contained within a multifunctional enzyme, peptidylglycine α-amidating monooxygenase (PAM) (Eipper et al., 1993; Prigge et al., 2000). In invertebrates such as the fruit fly *Drosophila melanogaster*, however, the two enzymatic activities are contained within two separate enzymes, originating from two genes (Kolhekar et al., 1997). Even in cnidarians, the lowest animal group having a nervous system, PHM appears to be present. This monooxygenase has been cloned from the sea anemone *Calliactis parasitica* (Hauser et al., 1997). Recently, PHM from the human parasite flatworm *Schistosoma mansoni* has been cloned and characterized (Mair et al., 2004), while putative amidation enzymes have only been tentatively identified from the *C. elegans* genome. This post-translational modification (conversion of a glycine into an amide) is a common feature of secretory, bioactive peptides and is believed to improve peptide stability and to be involved in receptor recognition.

2.2. Transport and exocytosis

Neurons basically contain two types of neurosecretory vesicles. Synaptic vesicles (SVs) store and release the classic small molecule neurotransmitters at presynaptic sites, whereas dense core vesicles (DCVs) store, transport and release bioactive peptides at multiple sites. Prior to exocytosis, DCVs bud from the *trans*-Golgi network and are subject to microtubule-based fast neuronal transport. This long-range DCV transport is mediated by motor proteins of the kinesin and dynein superfamilies (Goldstein and Yang, 2000). A large number of genes encoding microtubule-based molecular motors have been identified in *C. elegans*, with 22 kinesin genes and at least two dynein genes (Koushika and Nonet, 2000).

The regulated exocytosis of SVs and DCVs is based on conserved mechanisms showing some important differences. Many aspects of vesicle exocytosis have been elucidated during the past years, showing that SVs and DCVs employ a very similar set of proteins for the regulation and execution of their Ca^{2+}-triggered fusion with the plasma membrane. Although, there has been a major surge in the understanding of mechanisms underlying cargo release from these vesicles, many controversies in SV and DCV exocytosis are still subject of intense debate and thoroughly reviewed elsewhere (Weimer and Jorgensen, 2003; Rutter and Tsuboi, 2004; Sudhof, 2005).

2.3. Peptide degradation

After exerting their physiological effects at the synapses or neuromuscular junctions, peptides will undergo extracellular

degradation via specific proteolytic enzymes in order to turn off the neuropeptide signals. Incubation of the synthetic peptide KNEFIRFamide with locomotory muscle membranes from *Ascaris suum* and analysis of the degradation products revealed the presence of endopeptidase, aminopeptidase and deamidase activity (Sajid et al., 1996). This indicates that nematodes indeed possess an effective mechanism for inactivating peptides and hereby control their signaling properties.

Neprilysin (NEP, neutral endopeptidase) and angiotensin-converting enzyme (ACE) are two prototypical mammalian zinc metallopeptidases involved in the metabolism of neuropeptides. It has been shown that ACE from the housefly *Musca domestica* can hydrolyse a variety of insect neuropeptides by cleaving dipeptide-amides from the carboxyterminus (Lamango et al., 1997). No ACE-like activity has been found in nematode tissues although the *C. elegans* genome is predicted to display one ACE-like gene (*acn-1*) (Isaac et al., 2000; Coates et al., 2000). ACN-1 lacks the metallopeptidase active site, but seems to have an essential role in moulting and morphogenesis (Brooks et al., 2003). Survey of the *C. elegans* genome revealed the presence of 22 NEP-like genes (Coates et al., 2000; Turner et al., 2001). Only one (*nep-1*) has been studied thoroughly and appears to be involved in locomotion and pharyngeal pumping (Spanier et al., 2005). Although their neuronal system is composed of only 302 neurons and worms appear (at first sight) to have a limited behavioral repertoire, *C. elegans* displays a very complex peptide biosynthesis machinery. Not only the transcription and translation of neuropeptide genes but also the proteolytic processing of the peptide precursors, transport and exocytosis of the resulting bioactive peptides and finally their metabolic breakdown to terminate the signaling pathways are complicated matters that have to be highly regulated, spatially and temporally.

3. *C. elegans* neuropeptides: getting the sequence information

3.1. Methodology: a short overview

Understanding the role of neuropeptides has unfortunately often been hindered by the absence of primary sequence information and knowledge about their post-translational modifications. Peptide sequences have arrived very slowly, due to laborious efforts required for tissue collection and the need of multiple chromatographic separations to ultimately purify one active compound. The molecular mass of the peptide (as measured by mass spectrometry), combined with the amino acid sequence (determined by Edman degradation), allowed biochemical characterization of the bioactive fractions. Since the efficacy of the Edman technique is mostly limited by the sensitivity of the sequencer and its requirement of highly pure samples, this classical approach revealed the characterization of only 12 FMRFamide-related peptides from *C. elegans* (Rosoff et al., 1993; Marks et al., 1995, 1997, 1998, 1999a, 2001).

With the progress of tandem mass spectrometry, where peptide sequences become available from fragmentation spectra, an extremely powerful tool with unequalled sensitivity can be applied to biological extracts containing multiple components. A few years ago the concept of peptidomics (Clynen et al., 2001; Schulz-Knappe et al., 2001), which aims at the simultaneous visualization and identification of all the expressed peptides with their post-translational modifications of a cell, tissue or organism by mass spectrometry (Clynen et al., 2003; Baggerman et al., 2004; Fricker et al., 2006), was introduced. Doing so, the neuropeptidome of many invertebrates, including the nematodes *C. elegans* and *A. suum*, have been explored (Yew et al., 2003, 2005; Husson et al., 2005, 2006).

Completion of the *C. elegans* genome sequencing project in 1998 (The *C. elegans* Sequencing Consortium, 1998) has provided more insight into the complexity of nematode neuropeptide signaling. Peptides can in principle be predicted from genomic sequences, but this remains even more difficult than for regular proteins, where finding all exonic coding regions and splicing them together correctly in silico is already a challenge. Several search strings such as a glycine as a putative amide donor, 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. The extraordinary amount of 109 neuropeptide genes have been identified to date by bioinformatics and peptidomics strategies (Li et al., 1999a; Nathoo et al., 2001; Pierce et al., 2001; Husson et al., 2005; McVeigh et al., 2005b). According to their conserved motifs and sequence similarities, the resulting peptides can be subdivided into three main families.

3.2. FMRFamide-related peptides (FaRPs)

Based on carboxyterminal sequence similarity with the cardioexcitatory peptide FMRFamide, which was first discovered from a bivalve mollusc in 1977 (Price and Greenberg, 1977a,b), the FMRFamide-like peptide (*flp*) gene family was discovered in *C. elegans*. These FMRFamide related peptides or FaRPs all contain variation on the tetrapeptide motif X/Y-X-R-F-amide, where X is a non-polar hydrophobic (L, I, M or V) residue and Y is aromatic. This multifunctional group of structurally related neuropeptides is the largest well-studied peptide family in insects (Orchard et al., 2001) and also appears to be highly represented in nematodes (McVeigh et al., 2005b, 2006). With the identification of the first FaRPs-encoding gene, *flp-1*, from *C. elegans* (Rosoff et al., 1992), a start shot for the identification of additional members of this well-studied neuropeptide family was given. Radioimmunometrical screening of *C. elegans* extracts, however, yielded the biochemical evidence of only 12 FMRFamide-related peptides (FaRPs) that were purified by reversed phase high performance liquid chromatography (RP-HPLC) and subjected to Edman degradation analysis or gas-phase sequencing (Rosoff et al., 1993; Marks et al., 1995, 1997, 1998, 1999a, 2001).

Table 1
FMRFamide-like peptide (*flp*) genes and derived peptide sequences

Gene		Peptide sequence
LRFa family		
flp-1		**SADPNFLRFa**
		SQPNFLRFa
		ASGDPNFLRFa
		SDPNFLRFa
		AAADPNFLRFa
		(K)PNFLRFa
		AGSDPNFLRFa
flp-14	4x	**KHEYLRFa**
flp-15		**GGPQGPLRFa**
		RGPSGPLRFa
flp-18		**(DFD)GAMPGVLRFa**
		EMPGVLRFa
	3x	**(SYFDEKK)SVPGVLRFa**
		EIPGVLRFa
		SEVPGVLRFa
		DVPGVLRFa
flp-21		GLGPRPLRFa
flp-23		TKFQDFLRFa
flp-26		**(E)FNADDLTLRFa**
		GGAGEPLAFSPDMLSLRFa
IRFa family		
flp-2		**SPREPIRFa**
		LRGEPIRFa
flp-4		PTFIRFa
		ASPSFIRFa
flp-5		**GAKFIRFa**
		AGAKFIRFa
		APKPKFIRFa
flp-8	3x	**KNEFIRFa**
flp-10	(a)	QPKARSGYIRFa
flp-12		RNKFEFIRFa
flp-13		**AMDSPFIRFa**
		AADGAPFIRFa
		APEASPFIRFa
		ASPSAPFIRFa
		SPSAVPFIRFa
		ASSAPFIRFa
		SAAAPLIRFa
flp-17		KSQYIRFa
flp-25		**ASYDYIRFa**
MRFa family		
flp-3		SPLGTMRFa
		TPLGTMRFa
		EAEEPLGTMRFa
		NPLGTMRFa
		ASEDALFGTMRFa
		EDGNAPFGTMRFa
		SAEPFGTMRFa
		SADDSAPFGTMRFa
		NPENDTPFGTMRFa
flp-6	6x	**KSAYMRFa**
flp-20	2x	AMMRFa
flp-22	3x	**SPSAKWMRFa**
flp-27		GLGGRMRFa
flp-28		**APNRVLMRFa**

Table 1 (*Continued*)

Gene		Peptide sequence
VRFa family		
flp-7	3x	**SPMQRSSMVRFa**
	2x	**TPMQRSSMVRFa**
		SPMERSAMVRFa
		SPMDRSKMVRFa
flp-9	2x	**KPSFVRFa**
flp-11		**ASGGMRNALVRFa**
		AMRNALVRFa
		NGAPQPFVRFa
flp-16	2x	**AQTFVRFa**
		GQTFVRFa
flp-17	2x	KSAFVRFa
flp-19		**WANQVRFa**
		ASWASSVRFa
flp-24	(b)	**VPSAGDM(ox)M(ox)VRFa**
flp-25		DYDFVRFa
flp-32		AMRNSLVRFa
PRFa family		
flp-33		**APLEGFEDMSGFLRTIDGIQKPRFa**

For all *flp* genes, excluding *flp-32*, ESTs, ORFeomes, cDNAs or the derived peptides have been isolated. The number of copies of a peptide is indicated between brackets when more than one. Peptides with and without residues in brackets have been isolated. Only peptides with the -RFamide motif are included. Peptides in bold have been biochemically identified by Edmann degradation, MALDI-TOF MS or Q-TOF MS/MS. (a) The aminoterminal glutamine, Q, most likely will be transformed into a pyroglutamic acid although no biochemical evidence is present. (b) Four forms of the FLP-24 peptide have been characterized; without oxidations, with two oxidations or with one oxidation at each of the two methionins indicated. a, Carboxyterminal amidation.

Through conventional screens of cDNA libraries and bioinformatic programs to search the genome, the group of Li was able to find 24 FMRFamide-like peptide (*flp*) genes (*flp-1* to *flp-23* and *flp-28*) encoding peptides with the common -RFamide motif at their carboxyterminus (Li et al., 1999a,b; Kim and Li, 2004). McVeigh and co-workers recently reported the identification of five additional *flp* genes, *flp-24* to *flp-27* and *flp-32*, by EST data mining and BLAST analysis (McVeigh et al., 2005b, 2006), while peptides from *flp-24* and *flp-26* were independently sequenced by mass spectrometry at that time (Husson et al., 2005) (Table 1). Doing so, we used a high throughput peptidomics setup combining two-dimensional nanoscale liquid chromatography (2D-nanoLC) and quadrupole time-of-flight tandem mass spectrometry (Q-TOF MS/MS) to biochemically characterize 31 FaRPs (Husson et al., 2005). This list was recently extended by combining HPLC separation with mass spectrometric detection of naturally occurring peptides using a matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometer (Husson et al., 2006). Quite recently, another novel FaRP from a newly identified precursor gene, which we would like to annotate as *flp-33*, has been found by mass spectrometry. These data confirmed many of the previously predicted peptide sequences and hence the processing at basic and dibasic cleavage sites. In addition, a few truncated and extended

peptides were sequenced due to incomplete processing or the in vivo existence of longer/shorter functional peptides. Although, SDRPTRAMDSPLIRFamide was predicted as a FLP-13 derived peptide, for example, only the short form AMDSPLIRFamide was sequenced, indicating that the basic arginine at position six acts as a monobasic cleavage site. As a second example, the truncated FLP-18 peptide GAMPGVLR-Famide was identified, in addition to the initially predicted aminoterminally extended DFDGAMPGVLRFamide. The major advantage of biochemical characterization techniques like mass spectrometry-based peptidomics is that we can now deduce which peptides are actually expressed and processed and indicate the presence of post-translational modifications, if present. The identification of additional peptides using biochemical techniques can moreover lead to better bioinformatic predictions of novel peptide precursors. The peptidomics work on *D. melanogaster* (Baggerman et al., 2002, 2005), for example, has certainly yielded more accurate predictions of novel neuropeptide genes in the fruit fly. Using search strings such as the presence of a glycine residue as a putative amidation site, dibasic (or basic) cleavage sites, the presence of a signal peptide, specific peptide motifs and sequence homology with other known neuropeptides, 76 additional neuropeptide genes have recently been predicted in addition to the 43 annotated ones in *D. melanogaster* (Liu et al., 2006a). Even metazoan-wide bioinformatic searches based on conserved motifs deduced from vertebrates as well as invertebrate peptides can lead to the discovery of putative peptides (Liu et al., 2006b). Many software tools are trained using biochemical sequence data. The ''ProP'' software tool predicts cleavage sites of peptide precursors based on experimental results (Duckert et al., 2004), as is also the case for the ''NeuroPred'' web application (Southey et al., 2006). Screening of the *C. elegans* genome using structural peptide motifs has also lead to the discovery of an enormous amount of novel neuropeptides.

3.3. Neuropeptide-like protein (nlp) genes

Nathoo and co-workers introduced the second neuropeptide gene family in *C. elegans* using bioinformatic programs like BLAST and PATTERNFIND to search its genomic sequence (Table 2). Thirty-two so-called neuropeptide-like protein (*nlp*) genes were identified, based on structural criteria (Nathoo et al., 2001). These *nlp* genes define at least 11 families of putative neuropeptides with unique motifs such as FAFA, GGXYamide, MRXamide, LQFamide, LXDXamide and GGARAF or having sequence homology with other invertebrate neuropeptides like allatostatin, myomodulin, buccalin/drosulfakinin, orcokinin, etc. Using microarray analysis, expression of a novel gene, *nlp-33*, was induced in response to fungal or bacterial challenge, as was the case for the predicted *nlp-29* and *nlp-31* genes (Couillault et al., 2004). Concurrent with the biochemical detection of the FaRPs by mass spectrometry, we isolated the first NLP peptides from *C. elegans*, which allowed the identification of the novel genes *nlp-35* to *nlp-41* (Husson et al., 2005).

Table 2
Neuropeptide-like protein (*nlp*) genes and derived peptide sequences

Gene		Peptide sequence
nlp-1	x3	**MDANAFRMSFa**
		MDPNAFRMSFa
		VNLDPNSFRMSFa
nlp-2		SIALGRSGFRPa
		SMAMGRLGLRPa
	x3	SMAYGRQGFRPa
nlp-3		AINPFLDSMa
		AVNPFLDSIa
		YFDSLAGQSLa
nlp-4		SLILFVILLVAFAAARPVSEEVDRV
		DYDPRTEAPRRLPADDDEVDGEDRV
		DYDPRTDAPIRVPVDPEAEGEDRV
nlp-5		SVSQLNQYAGFDTLGGMGLa
		ALSTFDSLGGMGLa
		ALQHFSSLDTLGGMGFa
nlp-6		(MA)**APKQMVFGFa**
		YKPRSFAMGFa
		AAMRSFNMGFa
		LIMGLa
nlp-7	(a)	QADFDDPRMFTSSFa
		SMDDLDDPRLMTMSFa
		MILPSLADLHRYTMYD
		LYLKQADFDDPRMFTSSFa
nlp-8		**AFDRFDNSGVFSFGA**
		AFDRMDNSDFFGA
		SFDRMGGTEFGLM
		YPYLIFPASPSSGDSRRLV
nlp-9		GGARAFYGFYNAGNS
		GGGRAFNHNANLFRFD
		GGGRAFAGSWSPYLE
		TPIAEAQGAPEDVDDRRELE
nlp-10		AIPFNGGMYa
		STMPFSGGMYa
		AAIPFSGGMYa
		GAMPFSGGMYa
nlp-11		**HISPSYDVEIDAGNMRNLLDIa**
		SAPMASDYGNQFQMYNRLIDAa
		SPAISPAYQFENAFGLSEALERAa
nlp-12	x2	**DYRPLQFa**
		DGYRPLQFa
nlp-13		**NDFSRDIMSFa**
		SGNTADLYDRRIMAFa
	(a)	QPSYDRDIMSFa
		SAPSDFSRDIMSFa
		SSSMYDRDIMSFa
		SPVDYDRPIMAFa
		AEDYERQIMAFa
nlp-14	x2	ALDGLDGSGFGFD
	x5	**ALNSLDGAGFGFE**
	x3	ALDGLDGAGFGFD
		ALNSLDGQGFGFE
	x3	ALNSLDGNGFGFD
nlp-15		**AFDSLAGSGFDNGFN**
	x2	AFDSLAGSGFGAFN
		AFDSLAGSGFSGFD
		AFDSLAGQGFTGFE
		AFDTVSTSGFDDFKL

Table 2 (*Continued*)

Gene		Peptide sequence
nlp-16		STEHHRV
		SEGHPHE
		ATHSPEGHIVAKDDHHGHE
		SSDSHHGHQ
		SVDEHHGHQ
		NAEDHHEHQ
		SEHVEHQAEMHEHQ
		STQEVSGHPEHHLV
nlp-17		**GSLSNMMRIa**
	(a)	QQEYVQFPNEGVVPCESCNLGTLMRIa
nlp-18		**SPYRAFAFA**
		ARYGFA
		SPYRTFAFA
		ASPYGFAFA
		SDEENLDFLE
nlp-19		IGLRLPNFLRF
		IGLRLPNML
		MGMRLPNIFLRNE
nlp-20		FAFAFA
		SGPQAHEGAGMRFAFA
		APKEFARFARASFA
nlp-21		GGARAMLH
		GGARAFSADVGDDY
		GGARAFYDE
		GGARAFLTEM
		GGARVFQGFEDE
		GGARAFMMD
		GGGRAFGDMM
		GGARAFVENS
		GGGRSFPVKPGRLDD
		pQYTSELEEDE
nlp-22		SIAIGRAGFRPa
nlp-23		LYISRQGFRPA
		SMAIGRAGMRPa
		AFAAGWNRa
nlp-24	(a)	QWGGGPYGGYGP
		GYGGGYGGa
		YGGYGa
		FTGPYGGYGa
		GPYGYGa
		GPYGGGGLVGALLa
nlp-25		**pQWGGGYGNPYGGYa**
		GGGYGGGYGGGFGAQQAYNVQNAA
nlp-26	(a)	QFGFGGQQSFGa
		GGQFGGMQ
		GGFNGN
		GGFGQQSQFGa
	x2	GGNQFGa
		GGSQFNa
		GGFGFa
nlp-27	(a)	QWGYGGMPYGGYGGMGGYGMGGYGMGY
		MWGSPYGGYGGYGGYGGWa
nlp-28		QWGYGGYa
		GYGGYGGYa
	x2	GMYGGYa
		GMYGGWa
nlp-29	(a)	QWGYGGYa
		GYGGYGGYa
	x3	GMYGGYa
		GMYGGWa
nlp-30	(a)	QWGYGGYa
		GYGGYGGYa
		GYGGYa
		GMWa
		PYGGYGWa
nlp-31	(a)	QWGYGGYa
	x2	GYGGYGGYa
		GYGGYa
		GMYGGYa
		PYGGYGWa
nlp-32		GFYGGa
		GGWa
		GGa
		GYGa
		GGGWGa
		GGGWa
		GGGa
		FGYGGa
		GWa
nlp-33	(a)	QWGYGGPYGGYGGGYGGGPWGYGGGW
		HWGGYGGGPWGGYGGGPWGGYY
nlp-34		PYGYGGYGGW
		PYGYGWa
nlp-35		**AVVSGYDNIYQVLAPRF**
nlp-36		DDDVTALERWGY
		NIDMKLGPH
		SMVARQIPQTVVADH
nlp-37		**NNAEVVNHILKNFGALDRLGDVa**
nlp-38		**(ASDDR)VLGWNKAHGLWa**
		TPQNWNKLNSLWa
		SPAQWQRANGLWa
nlp-39		**EVPNFQADNVPEAGGRV**
nlp-40		**APSAPAGLEEKL(R)**
		MVAWQPM
nlp-41		**APGLFELPSRSV**
nlp-42		SALLQPENNPEWNQLGWAWa
		NPDWQDLGFAWa
nlp-43	x2	KQFYAWAa
nlp-44		APHPSSALLVPYPRVa
		LYMARVa
		AFFYTPRIa
nlp-45		RNLLVGRYGFRIa

The number of copies of a peptide is indicated when more than one. Peptides with and without residues in parentheses have been isolated. (a), The amino-terminal glutamine, Q, most likely will be transformed into a pyroglutamic acid although no biochemical evidence is present. Peptides in bold have been biochemically identified by Edman degradation, MALDI-TOF MS or Q-TOF MS/MS. a, carboxyterminal amidation.

Taken all together, the number of identified neuropeptide genes in *C. elegans* is likely to be an underestimate of the total number since novel peptides having striking sequence similarity with known invertebrate peptides are still being identified. For example, peptides derived from the novel *C.*

C45G9.13 *nlp-43*

*MSLAQSTFYL LFVAFLAVVI AVTA*DKQQSA YDLETPISAY
KRFYSWEDAK RAASSEEGMR NKR**KQFYAWA** GKRSSAPVHY
FEDSIAQEAG PSMEKR**KQFY AWAG**K

Y23B4A.2 *nlp-44*

*MLLWIVATLL IFSLPVSTA*L DYNDFSLQRI AR**APHPSSAL**
LVPYPRVGKR SNILNNNSES QNSVQKR**LYM ARVG**KR**AFFY**
TPRIGK

T01B6.4 *nlp-45*

*MRAVTILSLL GAIIVVNA*SY RATESPEQWR QRFLKSQSED
PSWLRFPTPP SRLPPLYVET EKQIRDRALA YREKYEKEHP
RPTVAADSDI DVEKR**RNLLV GRYGFRIG**KR FVGDQNDSPS
SSAVQMLPVQ VMYDDVLQ

Fig. 3. Novel NLP precursors from *C. elegans*. Amino acid sequences of three putative novel NLP precursors with their conventional protein name and novel *nlp* designation are represented. Consensus cleavage sites consisting of basic or dibasic residues are underlined. Putative peptides are in bold face and signal peptides as predicted by SignalP 3.0 (Bendtsen et al., 2004) are in italic face.

elegans gene *nlp-42* show sequence similarity with the myoinhibiting peptide (MIP) family of insect neuropeptides, as was the case for the MIP-like peptides from NLP-38. The MIPs were initially isolated and sequenced from an extract of 9000 locust brains (Schoofs et al., 1991), and later from CNS extracts of lepidopterans (*Manduca sexta* and *Bombyx mori*), crickets (*Gryllus bimaculatus*) and even from the stick insect *Carausius morosus* (Blackburn et al., 1995; Lorenz et al., 1995, 2000; Hua et al., 1999). They all share the carboxyterminal $W(X)_6$Wamide motif, suggesting that the biological activities exerted by this motif may be highly conserved. By using BLAST searches, we were able to find three additional neuropeptide genes (Fig. 3). These novel *nlp* genes display all common features of a typical neuropeptide precursor gene, for example, the presence of dibasic cleavage sites and the existence of a signal peptide, as analyzed by SignalP 3.0 (Bendtsen et al., 2004). The derived peptides from C45G9.13, which we would like to annotate as *nlp-43*, display sequence similarity with the insect leucokinins. These peptides contain a highly conserved carboxyterminal $F(X)_2$WA/Gamide sequence and were first isolated and identified from heads of the cockroach *Leucophaea maderae* where they are involved in osmoregulation and diuresis in insect Malpighian tubules (Hayes et al., 1989; Beyenbach, 2003). Since this novel *C. elegans* gene appears to be expressed in the intestine and the nervous system, especially in head neurons (McKay et al., 2004), it is likely that the derived leucokinin-like peptides of *C. elegans* are also involved in osmoregulation and fluid secretion. The next putative neuropeptide gene, Y23B4A.2, that we suggest to name *nlp-44*, encodes three peptides displaying sequence similarity with the cardioactive or cardioacceleratory peptides (CAPs). These peptides can be grouped in the widely occurring periviscerokinin (PVK) family of insect neuropeptides of which most members contain a carboxyterminal PRVamide or PRLamide motif (Wegener et al., 2002). Finally, T01B6.4 or *nlp-45* encodes one putative novel neuropeptide, RNLLVGRYGFRIamide. Taken all together, it is clear that the concept of neuropeptide signaling or at least their players is highly conserved among invertebrates and that the neuropeptide gene families in *C. elegans* are still expanding.

3.4. Insulin-like peptides

Insulin and related peptides were originally discovered in vertebrates, but also appear to be present in mollusks and insects where they are key regulators of growth and metabolism. Based on protein sequences issued from the 2.2 Mb nucleotide sequences from chomosome III (Wilson et al., 1994), 10 insulin homologues (*ins-2* through *ins*-7, *ins-11* and *ins-21* through *ins-23*) were identified in *C. elegans* (Duret et al., 1998). Later, the sequence characterization and gene structure of a novel insulin-like growth factor-1-like protein (IGF-1 or INS-1) was identified (Gregoire et al., 1998). Using a comprehensive bioinformatics study with multiple sequence searching tools, 37 insulin-like genes (*ins-1* through *ins-37*) were revealed, including the previously described members of the super family (Pierce et al., 2001). Later, a novel insulin-like gene, *daf-28*, was reported to be critical for the dauer signaling pathways (Li et al., 2003). Finally, an additional insulin-like peptide gene, *ins-38*, appeared to be present in *C. elegans* (Steve Doberstein and Kim Ferguson, personal communication), while a novel insulin-like gene, *ins-39*, has been found by BLAST searches (Dhaval Patel and David Gems, personal communication). All insulin-like peptide sequences are listed in Table 3.

Most insulin-like peptides of *C. elegans* are characterized by a signal peptide, a B-chain and an A-chain. Only two members, INS-1 and INS-18, are predicted to have a connecting C-peptide between the B and A chain, as is the case for the insulins of most other animals. Insulin precursors are processed by cleavage of the signal peptide and excision of the C domain, forming the two distinct B and A domains, which are joined by disulfide bonds. However, some *C. elegans* INS precursors display an additional aminoterminal F-peptide or small carboxyterminal domains.

The *C. elegans* insulin-like genes can be subdivided into three major classes, α, β and γ, on the basis of predicted arrangements of disulfide bonds (Fig. 4). Type γ insulin-like peptides have the arrangement of three canonical disulfide bonds, including one intrachain bond in the A chain. Eleven *ins* genes (*ins-11* through *ins-19*, *ins-32* and *ins-37*) display these types of disulfide bond arrangements, identical to those found in vertebrate insulins. Type β insulin peptides, including INS-1 through INS-10 and INS-38, contain the same three canonical bonds as well as an additional interchain disulfide bond. The intra A-chain disulfide link missing in the α-class peptides is substituted by the interaction of aromatic acid side chains in appropriate positions. INS-31 is a unique member of the insulin-like peptide family having three repeats of (B-A) chains. Although, cDNA sequences of *ins-1* through *ins-31* and *ins*-33 have been established, no INS peptides from *C. elegans* have ever been biochemically isolated and characterized. Therefore, it remains difficult to adequately predict the actual INS peptide sequences, which involves

Table 3
Insulin-like sequences

Gene name	Sequence
ins-1	*MYWFRQVYRPSFFFGFLAILLLSSPTPSDA*SIRLCGSRLTTTLLAVCRNQLCTGLTAFKRSADQSYAPTTRDLFHIHHQQKRGGIATECCEKRCSFAYLKTFCCNQDDN
ins-2	*MNAIIFCLLFTTVTA*TYEVFGKGIEHRNEHLIINQLDIIPVESTPTPNRASRVQKRLCGRRLILFMLATCGECDTDSSEDLSHICCIKQCDVQDIIRVCCPNSFRK
ins-3	*MKLSVVLALFIIFQLGAA*SLMRNWMFDFEKELEHDYDDSEIGFHNIHSLMARSRRGDKVKICGTKVLKMVMVMCGGECSSTNENIATECCEKMCTMEDITTKCCPSR
ins-4	*MFSFFTYFLLSALLLSASC*RQPSMDTSKADRILREIEMETELENQLSRARRVPAGEVRACGRRLLLFVWSTCGEPCTPQEDMDIATVCCTTQCTPSYIKQACCPEK
ins-5	*MHSIVALMLIGTILPIAA*LHQKHQGFILSSSDSTGNQPMDAISRADRHTNYRSCALRLIPHVWSVCGDAANHKTESMSLKNVAPLIAAPITSKKSAAHLTNWIPE
ins-6	*MNSVFTIIFVLCALQVAAS*FRQSFGPSMSEESASMQLLRELQHNMMESAHRPMPRARRVPAPGETRACGRKLISLVMAVCGDLCNPQEGKDIATECCGNQCSDDYIRSACCP
ins-7	*MPPIILVFFLVLIPASQQ*YPFSLESLNDQIINEEVIEYMLENSIRSSRTRRVPDEKKIYRCGRRIHSYVFAVCGKACESNTEVNIASKCCREECTDDFIRKQCCP
ins-8	*MSPIILIFFLVFIPFSQQ*HTSLEESLNDRIISEEVVEMLSEKEIRPSRVRRVPEQKNKLCGKQVLSYVMALCEKACDSNTKVDIATKCCRDACSDEFIRHQCCP
ins-9	*MIVTLIVFLVIGLQMAHLSQVSG*NNENGFLNPFDLSQWSEEILHRQYHHHHHHHHHGNRARRTLETEKIYRCGRKLYTDVLSACNGPCEPGTEQDLSKLCCGNQCTFVEIRKACCADKL
ins-10	*MSLHFSTIQKTILLISFLLLVTLAPRTSA*AFPFQICVKKMEKMCRIINPEQCAQVNKITEIGALTDCCTGLCSWEEIRISCCSVL
ins-11	*MSSYRQTLFILILIVIILFVNEGQG*APHHDKRHTACVLKIFKALNVMCNHEGDADVLRRTASDCCRESCSLTEMLASCTLTSSEESTRDI
ins-12	*MQSNITASLFIALLIFGVISA*APSHEKTHKKCSDKLYLAMKSLCSYRGYSEFLRNSATKCCQDNCEISEMMALCVVAPNFDDDLLH
ins-13	*MKLLHIFIIFLLFQSCSN*KMCQYSKKKYKICGVRALKHMKVYCTRGMTRDYGKLLVTCCSKGCNAIDIQRICL
ins-14	*MLTHLKFLLLVSLFINFAVS*SEDIKCDAKFISRITKLCIHGITEDKLVRLLTRCCTSHCSKAHLKMFCTLKPHEEEPHHEI
ins-15	*MKLLPLIVVFALLAVISESYS*GNDFQPRDNKHHSYRSCGESLSRRVAFLCNGGAIQTEILRALDCCSTGCTDKQIFSWCDFRKLTRKEKQNYSGLIFRNLNRTQIVSFLLCHVE
ins-16	*MQSLPILACLLTLSVFAPEIHG*RELKRCSVKLFDILSVICGTESDAEILQKVAVKCCQEQCGFEEMCQHANLKIDKI
ins-17 (a) *Ceinsulin-2*	*MFSTRGVLLLLSLMAAVAA*FGLFSRPAPITRDTIRPPRAKHGSLKLCPPGGASFLDAFNLICPMRRRRRSVSENYNDGGGSLLGRTMNMCCETGCEFTDIFAICNPFG
ins-18 (b) *Ceinsulin-1*	*MVHRLFIVLIAIILVAKSTA*ISLQQADGRMKMCPPGGSTFTMAWSMSCSMRRRKRDVGRYFEKRALIAPSIRQLQTICCQVGCNVEDLLAYCAPI
ins-19	*MIFYLTTYLVTMSPLFLILLLLVSTTYP*YIIDSSESYEVLMLFGYKRTCGRRLMNRINRVCVKDIDPADIDPKIKLSEHCCIKGCTDGWIKKHICSEEVLNFGFFEN
ins-20 (c)	MDKPSYLSSKEAWKMLNELLKEPKHHHHHHHRHKGYCGVKAVKKLKQICPDLCSNVDDNLLMEMCSKNLTDDDILQRCCPE
ins-21	*MKTYSFFVLFIVFIFFISSSKSHS*KKHVRFLCATKAVKHIRKVCPDMCLTGEEVEVNEFCKMGYSDSQIKYICCPE
ins-22	*MHTTTILICFFIFLVQVST*MDAHTDKYVRTLCGKTAIRNIANLCPPKPEMKGICSTGEYPSITEYCSMGFSDSQIKFMCCDNQ
ins-23	*MFVLLIILSIILAQVTDA*HSELHVRRVCGTAIIKNIMRLCPGVPACENGEVPSPTEYCSMGYSDSQVKYLCCPTSQ
ins-24	*MRSPTLFLLLLLVPLALC*HVFSEPADLELKSYQALEKSLKEMGLIRANQGPQKACGRSMMMKVQKLCAGGCTIQNDDLTIKSCSTGYTDAGFISACCPSGFVF
ins-25	*MLFKIIILFFLLLQLSEA*KPEAQRRCGRYLIRFLGELCNGPCSGVSSVDIATIACATAVPIEDLKNMCCPNL
ins-26	*MRALVAILCLMALCHA*AMLDELEMQKEVQEFHHMNGMLQEFMNKGLIGNHHHGTKAGLTCGMNIIERVDKLCNGQCTRNYDALVIKSCHRGVSDMEFMVACCPTMKLFIH
ins-27	*MKFFRLILLCALVLTTMAFLAPSTAA*KRRCGRRLIPYVYSICGGPCENGDIIIEHCFSGTTPTIAEVQKACCPELSEDPTFSS
ins-28	*MMRSFFVLLALLAIVTSTA*SPTCGRALLHRIQSVCGLCTIDAHHELIAIACSRGLGDKEIIEMCCPI
ins-29	*MFCKFVFLIFLLISLSVATA*DFGAQRRCGRHLVNFLEGLCGGPCSEAPTVELASWACSSAVSIQDLEKLCCPSNLA
ins-30	*MSSHALVLFLLLFLLPVALG*HFLSKPAPDPRITFNRKLAETLKELQDMGLIQAPREPVVAAQGAKKTCGRSLLIKIQQLCHGICTVHADDLHETACMKGLTDSQLINSCCPPIPQTPFVF
ins-31	*MKMPLILLLLVAAAS*AFVHHFDHSMFARPEKTCGGLLIRRVDRICPNLNYTYKIEWELMDNCCEVVCEDQWIKETFCRAPRFNFFGPSFKA-LERSCGPKLFTRVKTVCGEDINVDNKVKISDHCCTPEGGCTDDWIKENVCKQTRFNFFRQFLDSPQRSCGPQLFKRVNTLCNENINVENNVSVSKSCCESAAGCTDDWIKKNVCT-QHKPFVFRPGFY
ins-32	*MTSILLILLLVITVTG*MFQELSDLQNLHRFLEGLQGSSSLAVKSRSRRELICGRRLSKTVTNLCVEMNPQKEEDIATKCCKNKGCSREYIKSIMCPDE
ins-33	*MANTCLILLLLLVIFVTVGFS*MPRIFRASENGVNSSDEVSEELSYSPEEAMDLVKQVIKVREQRRHRRHRHHGQKHCGTKIVRKLQMLCPKMCTISDDTLLTEMCSHSLFDDEIQLRCCPKEDE
ins-34	*MLHHKTLIIALLLTLFISGIDS*LPFRKHNNHRHLKNQKAQQLKEEATEAPTPAPTTTKAPSGSATTTTTVKTTAAPLAQVNPQCLRRLTLLARGVCRQPCQPSDKPKTSAQQLLQL-ACSARRPTNEQIISYCCPEKSG
ins-35	*MKQIFLVILAACLLAIILA*SPTGKHHKMDENAFGINNRHCQRALKVYSFAICGAICQNYEKILMEGCGSTVMLTMQRTKLICCPEPVDSDELFN
ins-36	*MNIGKCSIIFLLFCVFGSILS*RAIRKRHPEGKLVIRDCKRYLIMYSRTICKEKCEKFDERNDITFSINLQFIFTDLLVEGCHSNQTLSNERTRELCCPNAGSN
ins-37	*MAAFLPIALSIAMLTVLTNA*NPIHPVPNAAFLPYRSCGSHLVHRAFEACSGKKDRSSDVDLWKMCCKDECTDLDIKESLCKYASQGYGVK-FEEEAEEIDMVSFAAEGFKKSCGHDIVVKTVNVPTKISLKQCARTRGWRRRPRRIPDKFNRVQNYACINCVKIK
ins-38	*MNLFLLVCIAFAIITVTS*FTPDEKSQRSHVFSYKKHCGRRIVSLVQACDRIDHDLSIDCCTQNCSSEFVKKIMCPSKL
ins-39	*MNTFFFLAVLLVFCSA*EQMTAKKFSTKTSSPIPELQEVFATVAADEFPFHKANTQPLAIYLNISTPQDCIHKIFRMTISFCSQVECQNMEAMQKICNTTTPTIKHVGELCCPEFFEQVKDDFVTLL
daf-28	*MNCKLIAIFAVLVLSVSAHLGAQAAA*ANFKAEGPLSRAVRVPGVAVRACGRRLVPYVWSVCGDACEPQEGIDIATQCCTYQCTAEYIQTACCPRLLL

Signal sequences as predicted by SignalP 3.0 are indicated in italic. (a) Another name for *ins-17* is *Ceinsulin-2* or *ins-2*, which are EMBL gene names that were not approved by the *Caenorhabditis* Genetics Centre (CGC) as indicated on wormbase. (b) The same holds true for *ins-18* and the EMBL name *Ceinsulin-1*. (c) INS-20 is not predicted to contain a signal peptide.

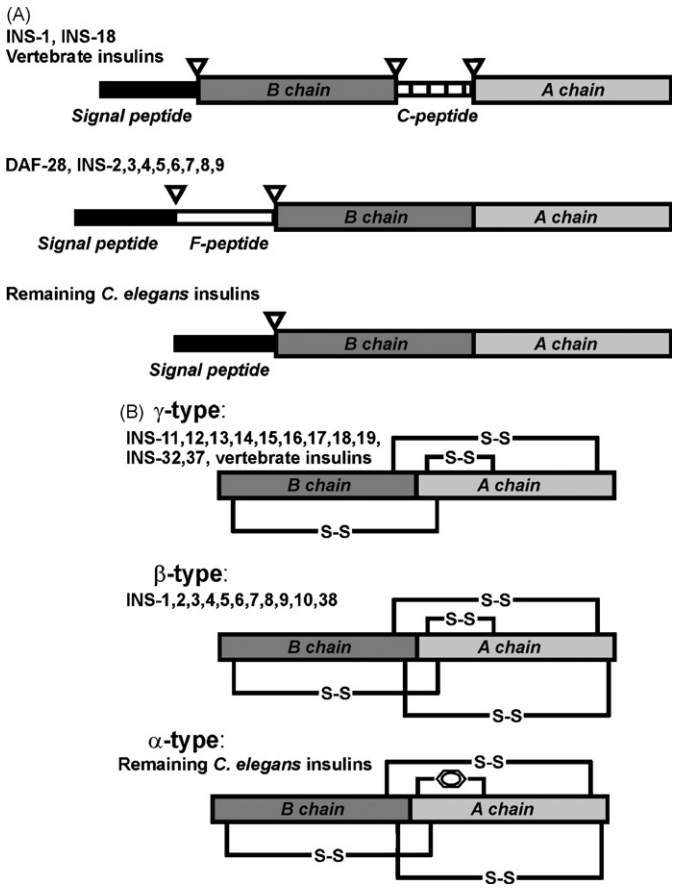

Fig. 4. Gene structure and domain organisation of *C. elegans* insulin-like peptides. (A) Domain structure of the insulin-like peptides (INS) from *C. elegans*. All prepro-insulin super family proteins contain at least a signal peptide, B chain and A chain. Only for INS-1 and INS-18, an additional C-peptide in between the B and A chains is present, as is the case for the vertebrate insulins. Other *C. elegans* INS peptides are characterized by the absence of a C-peptide. An additional F-peptide appears to be present in DAF-28 and INS-2 through INS-9. Predicted cleavage sites for the proteolytic processing of the preproinsulins are indicated by inverted triangles. No information regarding the domain structure for the novel insulins INS-38 and INS-39 is available. (B) Mature insulin-like peptides of *C. elegans* can be classified into three major types (α, β and γ) on the basis of the predicted organisation of disulfide bonds. Type-γ insulins have the arrangement of three disulfide bonds as found in vertebrates while type-α and β insulins contain an additional interchain disulfide bond. Type-α insulins lack the common intrachain bond in the A chain, which is substituted by the interaction of aromatic amino acid side chains in the case of INS-21, INS-22 and INS-23. INS-31 constitutes its own additional class with three repeats of the B and A peptide chains. Classes of *C. elegans* INS peptides are based on similarity of gene structures, as adapted from Duret et al. (1998) and Pierce et al. (2001).

prediction of putative cleavage sites and possible post-translational modifications.

4. Cellular localization and expression patterns

4.1. Immunocytochemistry

The first localization studies of FaRPs in *C. elegans* were based on whole mount immunocytochemical techniques using antibodies that recognize the -RFamide moiety of the FLP peptides (Atkinson et al., 1988; Li and Chalfie, 1990; Schinkmann and Li, 1992). About 25–30 neurons, or roughly 10% of the neurons in the animal, consistently stained in the hermaphrodite adult animal, in addition to four cells in the somatic gonad (Schinkmann and Li, 1992). This number, however, appeared to be an underestimate of the number of FaRP-expressing neurons. More than 50% of the neurons in *C. elegans* have been reported to be FaRP-immunoreactive, using different antisera (Li et al., 1999b). These data appeared to be closer to the actual figure, which are in agreement with subsequent reporter construct studies. Immunocytochemistry, however, is based on a non-competitive antibody-antigen interaction and as such there is a realistic potential for non-specific interaction of the antisera employed. Since most antibodies against various FLP peptides do not distinguish among the different FaRPs (and therefore suffer from cross-reactivity with other FLP peptides), the use of reporter constructs is a more elegant method for studying the expression pattern of particular neuropeptide genes. Designing specific antibodies against the numerous INS or NLP peptides, which also display very similar sequences, is a tremendous amount of work and, consequently, no histochemical experiments regarding these peptides have been performed so far.

4.2. Reporter constructs

Expression of the most studied neuropeptide gene *flp-1* has been demonstrated using *flp-1-lacZ* reporter constructs in transgenic animals (Nelson et al., 1998). This way, the expression pattern of one single peptide precursor gene could relatively easily be monitored. Microinjection of transcriptional fusions of the promoter region of the neuropeptide gene of interest with a reporter gene like *lacZ* or the *green fluorescent protein gene* (*gfp*) is a relatively simple technique to study expression profiles in *C. elegans* (Chalfie et al., 1994). Doing so, the cellular localization of 15 insulin-like genes (Pierce et al., 2001; Li et al., 2003), 19 *flp* genes (Kim and Li, 2004) and 26 *nlp* genes (Nathoo et al., 2001) have been determined using GFP fusion constructs. Injection of promoter::*gfp* reporter constructs of *flp-1* to *flp-23* resulted in a total number of 160 *flp* expressing neurons, corresponding to just over 50% of the neurons in *C. elegans*. GFP expression could be detected in all neuronal cell types, including sensory neurons, interneurons, motor neurons and pharyngeal neurons. Six *flp* genes were also expressed in non-neuronal cells such as head muscle (*flp-2* and *flp-11*), pharyngeal muscle (*flp-5* and *flp-15*), vulva cells (*flp-10*) and uterine cells (*flp-11* and *flp-2*). Although each *flp* gene has a unique expression pattern, many cells display the expression of more than one *flp* gene, resulting in a considerable overlap among the *flp* expression profiles (Kim and Li, 2004).

Although very elegant and useful, there are some limitations associated with deriving expression patterns using reporter constructs. It is always hard to identify the promoter region of a gene. Since the upstream region of a few kb is mostly used, additional upstream regulatory or enhancer regions can be lost. Another drawback is the detection sensitivity of the signal obtained. Transgenic animals carrying P*flp9*::*gfp*, P*flp14*::*gfp*, P*flp16*::*gfp* or P*flp23*::*gfp*, for example, did not show any

fluorescence signal (Kim and Li, 2004), while all peptides encoded by *flp-9*, *flp-14* and *flp-16* have been sequenced (Husson et al., 2005). The FLP-14-derived peptide KHEYLR-Famide, which could be isolated from *C. elegans* (Marks et al., 1995), *A. suum* (Cowden and Stretton, 1993) and *H. contortus* (Marks et al., 1999b), appears to be a potently bioactive peptide (Cowden and Stretton, 1993; Geary et al., 1999a; Thompson et al., 2003). Moreover, the peptidomics work by the Stretton group (Yew et al., 2003, 2005) in *Ascaris* suggests that more peptides are expressed in each neuron or group of neurons than has been previously thought.

Of the 32 P*nlp*::*gfp* constructs analyzed, many transgenic lines had complex neuronal expression patterns, as was the case for the *flp* genes, while for six *nlp* genes (*nlp-4, nlp-17, nlp-22, nlp-25, nlp-28* and *nlp-32*), no detectable fluorescent signals could be observed (Nathoo et al., 2001). Non-neuronal expression, e.g. in somatic gonad tissues and secretory cells, was also observed for nine *nlp* genes, suggesting an additional neuroendocrine role. Family members of the YGGWamide subgroup, including *nlp-24, nlp-27, nlp-29, nlp-30* and *nlp-31*, displayed predominant hypodermal expression in comparison to the other *nlp* genes. Later, three genes (*nlp-27*, *nlp-29* and *nlp-31*) were found to show unique profiles of induction after fungal and bacterial infection by microarray analysis, suggesting a role as antimicrobial peptide genes instead of genes encoding neuropeptides (Couillault et al., 2004).

Most of the *ins* genes analyzed are primarily expressed in subsets of neurons throughout most of the life cycle (Pierce et al., 2001). They appear to be present in at least some amphid sensory neurons, with exception of *ins-11*. Most *ins* genes are expressed in distinct but overlapping sets of neurons, although several genes also appear to be transcribed in non-neuronal tissues, including the hypodermis, vulva and pharynx.

5. Unraveling neuropeptide functions

5.1. G-protein coupled receptors (GPCRs) and G-protein signaling

The ability to fully unravel the neuropeptidergic signaling system in *C. elegans* is not only constrained by the worm's size, but also by a limited understanding of endogenous targets, i.e. the peptide receptors. The modulatory effect of neuropeptides on neurotransmission is first initiated by binding on their appropriate G-protein coupled receptors (GPCRs) (Fig. 5). Over 1100 GPCRs are predicted to be present in *C. elegans*, 50 of which were suggested as candidate neuropeptide receptors (Bargmann, 1998). These receptors typically display seven transmembrane-spanning structures and signal via heterotrimeric G-proteins in which a GDP-bound Gα subunit is tightly bound to the obligate heterodimer of Gβγ (McCudden et al., 2005). *C. elegans* contains 2 Gβ, 2 Gγ and 21 Gα genes among which homologues of each of the 4 mammalian classes of Gα genes, $G\alpha_i$/$G\alpha_o$, $G\alpha_s$, $G\alpha_q$ and $G\alpha_{12}$, are represented (Jansen et al., 1999; Cuppen et al., 2003). In addition, *C. elegans* and other nematodes also contain nematode-specific Gα genes ($G\alpha_{ns}$), which are hypothesized to increase the functional complexity of individual chemosensory neurons (O'Halloran et al., 2006).

Upon ligand activation, the agonist-bound GPCR acts as a guanine nucleotide exchange factor and thus promotes the release of bound GDP by Gα. Binding of the nucleotide-free Gα subunit with GTP leads to conformational changes, which result in the dissociation of Gβγ. The exchange of GDP by GTP is therefore the first reaction that can be monitored in an assay in order to detect the appropriate ligands for GPCRs. Gβγ can directly stimulate phospholipase C (PLC), as is the

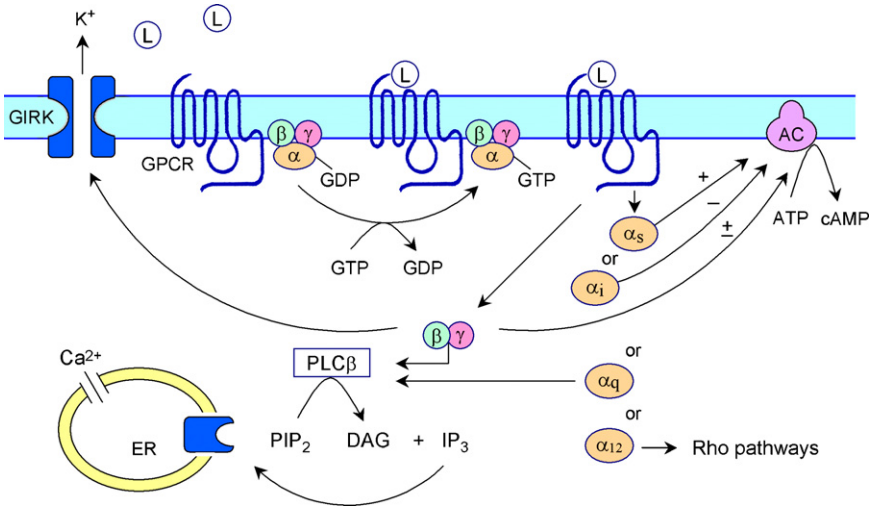

Fig. 5. Signal transduction pathways of G-protein coupled receptors. Upon activation of the G-protein coupled receptor (GPCR) by the binding of an appropriate ligand (L), the associated heterotrimeric G-protein, which contains an α, β and γ subunit, exchanges GDP for GTP. After dissociation of the G-protein into the signaling-competent Gα and Gβγ complexes, different signaling pathways can be triggered, depending on the nature of the Gα subunit. $G\alpha_i$ or $G\alpha_s$ result in the inhibition or activation of adenylyl cyclase (AC), respectively, subsequently leading to a decrease or increase of intracellular cAMP levels. $G\alpha_q$ can directly stimulate phospholipase Cβ (PLCβ), as is the case for Gβγ, leading to the release of intracellular Ca^{2+} from the endoplasmatic reticulum (ER). Gβγ has also been shown to regulate various AC isoforms in addition to G-protein-regulated inward-rectifier K^+ (GIRK) channels. DAG, Diacylglycerol; IP_3, inositol trisphosphate; PIP_2, phosphatidylinositol 4,5-bisphosphate.

case for $G\alpha_q$. PLCs hydrolyze the phosphoester bond of phosphatidylinositol 4,5-bisphosphate (PIP_2) in order to generate the second messengers diacylglycerol (DAG) and 1,4,5-trisphosphate (IP_3), which results in subsequent mobilization of Ca^{2+} from the endoplasmatic reticulum (ER). Monitoring the intracellular Ca^{2+} releases is a second strategy to screen for GPCR activation. Adenylyl cyclase (AC) generates cAMP from ATP and can be activated or inhibited by two other G proteins, $G\alpha_s$ or G α_i, respectively. $G\beta\gamma$ has also been shown to regulate various AC isoforms in addition to G-protein-regulated inward-rectifier K^+ (GIRK) channels.

To unravel possible functions of these seven-transmembrane receptors, whole-genome analysis by gene knockdown with RNA interference (RNAi) was performed (Keating et al., 2003). Doing so, 60 GPCRs that were predicted to bind either small-molecule neurotransmitters or neuropeptides were screened by using RNAi and quantitative behavioral assays. This showed that particular orphan receptors, C16D6.2, C25G6.5, F35G8.1 and F59C12.2, are likely to have a role in locomotion, whereas others, C15B12.5, C10C6.2, C24A8.4, F15A8.5, F59D12.1, T02E9.1 and T05A1.1, appear to be involved in reproduction (Keating et al., 2003). Many of these so-called orphan receptors, however, are still waiting to be assigned a natural ligand.

It is clear that agonist binding of GPCRs can result in a specific activation of several classical effectors and enzymes in order to exhibit complex behaviors or physiological responses. For example, natural variations in *C. elegans* having either Phe-215 or Val-215 in the neuropeptide Y-like GPCR NPR-1 result in either ''social'' or ''solitary'' feeding behavior (de Bono and Bargmann, 1998). Solitary worms feed alone and move slowly on the bacterial *Escherichia coli* lawn, while social animals move rapidly and aggregate together at the edges of the bacteria. These behavioral differences can be explained by a change of just one amino acid in the NPR-1 receptor. Social behavior is probably induced by neurons that detect aversive stimuli since ablation of the nociceptive neurons ASH and ADL transforms social animals into solitary feeders (de Bono et al., 2002). A natural variation in the acute response to ethanol was also shown to be mediated by the natural allelic variation in NPR-1 (Davies et al., 2004).

Obviously, nematodes also possess a wide variety of other downstream signaling molecules such as guanylate cyclases (GCs). These enzymes synthesize 3′,5′-cyclic guanosine monophosphate (cGMP), which is a common second messenger in sensory transduction. At least 26 receptor GC genes seem to be present in *C. elegans* (Yu et al., 1997; Bargmann, 1998), in addition to 7 soluble GCs (Morton et al., 1999). Soluble GCs are heterodimeric proteins having an α and a β subunit that bind a heme group and are activated by nitric oxide (NO). Surprisingly, however, the *C. elegans* genome does not appear to encode a gene for NO synthase.

Two soluble GCs, GCY-35 and CGY-36, have been shown to promote aggregation behavior in *C. elegans* (Cheung et al., 2004). These genes are expressed in neurons that also express the variable NPR-1 receptor and a heterodimeric cGMP-gated ion channel, TAX-2/TAX-4, which are also known to affect group foraging behavior of the social animals (Coates and de Bono, 2002). This bordering behavior, where the worms have the tendency to accumulate at the outer edges of the bacterial lawn, has also been correlated with an aversion to high O_2 concentrations. Interestingly, the same GCs were independently found to mediate an avoidance response to hyperoxia as analyzed by Cori Bargmann and colleagues (Gray et al., 2004). This example shows that neuropeptidergic signaling through GPCRs involves many downstream elements that can determine much of the phenotypic behavioral responses in *C. elegans*.

Table 4
Characterized *C. elegans* GPCRs and their activating ligands

C10C6.2 (FLP-15 receptor) [locomotion] (Kubiak et al., 2003b)	
flp-15	GGPQGPLRFamide
flp-15	RGPSGPLRFamide
flp-15	GPSGPLRFamide
C39E6.6 (NPR-1) 215-F [feeding behavior] (Rogers et al., 2003; Kubiak et al., 2003a)	
flp-21	GLGPRPLRFamide
C39E6.6 (NPR-1) 215-V [feeding behavior] (Rogers et al., 2003; Kubiak et al., 2003a)	
flp-18	EMPGVLRFamide
flp-18	DFDGAMPGVLRFamide
flp-18	SVPGVLRFamide
flp-18	EIPGVLRFamide
flp-18	SEVPGVLRFamide
flp-18	DVPGVLRFamide
flp-21	GLGPRPLRFamide
T19F4.1 a, b (FLP-2 receptor) [no phenotype] (Mertens et al., 2005)	
flp-2	SPREPIRFamide
	LRGEPIRFamide
C26F1.6 (VRFamide receptor 1) [egg-laying] (Mertens et al., 2004b)	
flp-7	TPMQRSSMVRFamide
flp-7	SPMQRSSMVRFamide
flp-7	SPMERSAMVRFamide
flp-11	AMRNALVRFamide
Y59H11AL.1 [function unknown] (Mertens et al., 2006)	
flp-7	TPMQRSSMVRFamide
flp-7	SPMQRSSMVRFamide
flp-7	SPMERSAMVRFamide
flp-7	SPMDRSKMVRFamide
flp-9	KPSFVRFamide
flp-11	AMRNALVRFamide
flp-11	NGAPQPFVRFamide
flp-11	ASGGMRNALVRFamide
flp-13	AADGAPLIRFamide
flp-13	ASPSAPLIRFamide
flp-13	SPSAVPLIRFamide
flp-13	ASSAPLIRFamide
flp-13	SAAAPLIRFamide
flp-22	SPSAKWMRFamide

So far, only FLP-activating G-protein coupled receptors (GPCRs) have been characterized. Gene names of the GPCRs with their descriptive names are indicated in bold, together with the observed phenotype between brackets, as mostly analyzed by RNAi studies (Simmer et al., 2002; Kamath et al., 2003; Keating et al., 2003). The solitary or social feeding behaviors of *C. elegans* strains having one of the two naturally occurring NPR-1 variants were originally described in (de Bono and Bargmann, 1998).

5.2. Deorphanizing GPCRs: reversed pharmacological approach

The ''orphan receptor strategy'' or ''reversed pharmacological approach'' tries to fish for appropriate synthetic ligands using a high throughput screen where the heterologuous expressed orphan receptors are used as baits (Civelli et al., 1999; Meeusen et al., 2003; Mertens et al., 2004a). For only five neuropeptide GPCRs in *C. elegans*, the activating ligand has been reported in papers (Table 4), while five additional receptors are described in a patent of the Upjohn/Pharmacia group (splice variants not taken into account) (Rogers et al., 2003; Lowery et al., 2003; Kubiak et al., 2003a,b; Mertens et al., 2004b, 2005, 2006). After ligand binding, GPCRs act on Gα-GDP/Gβγ heterotrimers to promote GDP release and GTP binding, which is thus the first reaction in the signal transduction pathway that can be monitored in order to characterize orphan receptors. Membrane fractions of CHO cells that express the receptor of interest are used to measure GTP-γS (a non-hydrolyzable GTP analog where an oxygen atom is substituted for a sulfur atom between the β and γ phosphates of GTP). These membrane fractions are incubated with the samples and GDP. Then, [^{35}S]GTP-γS is added and the samples are measured in a scintillation counter where non-specific binding is measured in the presence of cold GTP-γS. This way, C10C6.2 could be annotated as a FLP-15 peptide receptor (Kubiak et al., 2003b). They also showed that both Phe-215 and Val-215 forms of the NPR-1 receptor can be activated by the FLP-21 peptide, although FLP-21 is significantly more potent at the Val-215 NPR-1 receptor (Kubiak et al., 2003a). In addition, the group of Mario de Bono found that peptides encoded by *flp-18* solely activate NRP-1 Val-215 when using the *Xenopus* oocytes assay (Rogers et al., 2003). For the functional expression of GPCRs in *Xenopus* oocytes, cDNAs of the receptor and the co-expressed GIRK1 channel are mixed and injected into the oocytes. The gating of the GIRK channels is due to a direct interaction with free Gβγ and is measured by whole-cell voltage-clamp recordings.

In our lab, human embryonic kidney (HEK) cells which are co-transfected with the promiscuous G protein $G\alpha_{16}$ are used. After transfection, the cells are loaded with a fluorophore (Fluo-4-AM), which allows monitoring of a calcium response when an activating ligand binds to the transfected GPCR. Alternatively, Chinese hamster ovary (CHO) cells that are stably transfected with $G\alpha_{16}$ and the bioluminescent protein apoaequorin can also be used. After transient transfection of the orphan GPCR, these cells have to be incubated with coelenterazine to reconstitute the holoenzyme aequorin, prior to screening with a library of synthetic peptides or HPLC fractions of a *C. elegans* peptide extract. This way, three receptors (C26F1.6, T19F4.1 and Y59H11AL.1) were successfully characterized (Mertens et al., 2004b, 2005, 2006). In summary, two peptides ending in M/LVRFamide elicited a dose-dependent response of C26F1.6 or the VRFamide receptor 1, which appeared to be very specific since closely related FaRPs were less active (Mertens et al., 2004b). Pharmacological profiling of the most active peptide suggests that SMVRFamide is the most active minimal sequence core required to activate C26F1.6, because aminoterminal extensions decrease receptor activity. Disruption of the VRFamide receptor 1 gene resulted in a statistically significant increase in the number of progeny counted 48 h post-L4 as analyzed by RNAi (Keating et al., 2003). The receptor could therefore be involved in reproduction. The second receptor characterized by Mertens et al. is the so-called FLP-2 receptor of which there are two alternatively spliced GPCRs (T19F4.1a and T19F4.1b) (Mertens et al., 2005). Both receptors were characterized as being cognate receptors for the two FLP-2 peptides. Pharmacological profiling revealed that the active core of both peptides is EPIRFamide. A knock down of the FLP-2 receptor did not reveal any obvious visible phenotypes as observed in several RNAi studies (Simmer et al., 2002; Kamath et al., 2003; Keating et al., 2003). Finally, FaRPs mainly deduced from FLP-7, FLP-11 and FLP-13 were able to activate the tachykinin-related orphan receptor encoded by the Y59H11AL.1 gene (Mertens et al., 2006). The function of this GPCR, however, is unknown.

In *C. elegans*, only a single tyrosine kinase-like insulin receptor gene, *daf-2*, has been characterized so far (Kimura et al., 1997). In 2003, genetic evidence was provided that the novel insulin-like gene *daf-28* functions in the *daf-2* pathway of dauer arrest (Li et al., 2003). Nematodes enter this diapause life stage or dauer cycle when exposed to harsh environmental conditions like temperature changes or food depletion. Dauer larvae do not feed and their metabolism is dependent on internal food reserves. The decision to enter this alternative life stage is mediated by the DAF-2 pathway and a second pathway that uses Transforming Growth Factor-β (TGF-β) signaling (Kimura et al., 1997; Riddle and Albert, 1997). Adult worms, which express defects in the insulin/insulin-like growth factor receptor DAF-2 also display enhanced longevity. When conditions become favourable again, the worms leave the dauer cycle and resume their normal, reproductive life.

DAF-28 is not the only ligand for DAF-2. High levels of INS-4 and INS-6 can functionally substitute for DAF-28 since overexpression of *ins-4* and *ins-6* can suppress the *daf-28* mutation (Li et al., 2003). INS-6 has also been reported to activate the human insulin receptor (Hua et al., 2003). Since overexpression of *ins-1* causes arrest at the dauer stage and enhances dauer arrest in weak *daf-2* mutants, INS-1 is suggested to antagonize DAF-2 insulin-like signaling (Pierce et al., 2001). INS-1 and INS-18 are the only *C. elegans* insulin-like peptides containing a C-peptide that is predicted to be proteolytically removed. This is remarkable since only overexpression of *ins-1* and *ins-18* enhance dauer arrest in both wild type and *daf-2* mutant backgrounds, whereas *ins* genes without predicted C-peptide like, for example, *ins-9, ins-19, ins-22* and *ins-31* do not affect dauer arrest (Pierce et al., 2001). Despite the wide diversity among the 40 predicted INS peptides, only one receptor, *daf-2*, has been characterized until now. However, a series of insulin-like receptors from *C. elegans* have been identified by sequence profile searches using PSI-BLAST (Dlakic, 2002). Given the diversity of the *C. elegans ins* gene

family, it is very likely that some INS peptides act through these putative novel INS-receptors.

Although, there is a wide variety of *C. elegans* GPCRs and an overwhelming number of predicted NLPs, only a few FaRPs have been matched to their GPCRs in *C. elegans*. The first NLP-binding receptors, however, are currently being characterized (Janssen and Lindemans, personal communication). Some of them show remarkable sequence similarity with mammalian vasoactive intestinal peptide (VIP or VPAC) receptors, gonadotrophin-releasing hormone (GnRH) receptors and cholecystokinin receptors, suggesting that the key signaling pathways for reproduction, biological rhythms and digestion are evolutionarily conserved from nematodes through vertebrates.

5.3. Neuropeptide physiology

Despite the disadvantage of its size, electrophysiological studies to reveal the effects of FaRPs on pharyngeal tissues have been performed in *C. elegans*, demonstrating an impressive neurochemical complexity of the simple circuit that regulates feeding in the nematode (Rogers et al., 2001; Papaioannou et al., 2005). Applying synthetic peptides on muscle preparations of larger nematodes to investigate neuromuscular functions is another strategy to gather functional information. Numerous FaRPs originating from free living and parasitic nematodes have been shown to induce a diverse array of inhibitory and excitatory activities on dorsal or ventral muscle strip preparations of *A. suum* (Cowden and Stretton, 1995; Holden-Dye et al., 1995; Marks et al., 1997, 1999a, 2001). However, these data do not always facilitate a better understanding of the in vivo physiological functions. For example, the activity of KHEYLRFamide, which was isolated from *C. elegans* (Marks et al., 1995), *A. suum* (Cowden and Stretton, 1993) and *H. contortus* (Marks et al., 1999b), is very complex in that it induces a transient relaxation followed by an extended period of increased contractility in neuromuscular strips prepared from *A. suum* (Cowden and Stretton, 1993). Removal of the nerve cords abolishes the excitatory response even though cAMP levels remain elevated, as was the case for muscle strips with the nerve cords (Thompson et al., 2003), suggesting that increasing cAMP levels are associated with muscle relaxation, as earlier presupposed (Geary et al., 1999a). In the case of KSAYMRFamide, however, a unique differential activity comprising profound excitatory effects on ventral and slow relaxation effects on dorsal body wall muscle strip preparations were observed (Maule et al., 1995). This is the only known nematode peptide to show such differential neuromusculatory activity. Using the ovijector of *A. suum*, five categories of FaRP-induced responses could be identified, according to the pattern of changes in contractile behavior and baseline tension (Moffett et al., 2003). The influence of FLP peptides on pharyngeal pumping behavior has also been monitored in *A. suum* using a modified pressure transducer system that detects changes in intrapharyngeal pressure and therefore contraction of the radial muscle of the pharynx (Brownlee and Walker, 1999).

In order to understand in vivo roles of neuropeptides, comprehensive analysis on locomotory behaviors of intact adult *A. suum* have been carried out. Direct injections of FaRPs into the body cavity resulted in a wide variety of behavioral responses, including effects on body waveforms, body length, paralysis and altered cAMP concentrations (Reinitz et al., 2000).

5.4. Genetics approach

Knowledge about the neuropeptide expression profiles, efforts in gathering accurate peptide sequence information which can be used to characterize orphan GPCRs and the performance of (electro)physiological tests to unravel neuromuscular functions have already provided an overview of possible neuropeptide functions. Reverse genetics through overexpression or inactivation of (peptide) genes and of their receptor targets enables us to assign additional physiological functions to some of the neuropeptides. Inactivation of *flp-1*, for example, causes numerous behavioral defects, including uncoordination, hyperactivity, insensitivity to high osmolarity (Nelson et al., 1998) and egg-laying defects (Waggoner et al., 2000). Conversely, overexpression results in the reciprocal phenotypes. Deletions in ten other *flp* genes (*flp-3, 4, 6, 8, 9, 10, 12, 19, 20* and *21*) have been isolated until now, but most of them do not display an obvious behavioral phenotype (Li et al., 1999a; Li, 2005). One possible explanation could be functional redundancy, i.e. overlapping physiological effects and cellular expressions of the enormous quantity of different neuropeptides in *C. elegans*. Alternatively, it could be due to the scored phenotype that might not be observable in normal physiological conditions. No additional information regarding *flp* nor *nlp* mutants, with the exception of *flp-1*, is currently available.

The widespread neuronal expression and sequence similarities of NLPs with other invertebrate neuropeptides suggest an involvement in a wide variety of behaviors. Many genes are expressed in chemosensory neurons, suggesting a potential role in mediating responses to the changing environment, although no *nlp* mutant phenotypes have been described so far. *nlp-29, nlp-31* and *nlp-33*, all sharing YGGWG and YGGYG motifs, show unique profiles of induction after bacterial and fungal infections, as analyzed by microarrays (Couillault et al., 2004). These genes, presumably all encoding antimicrobial peptides, are mainly expressed in the hypodermis while no expression in neuronal tissues could be observed. A few recently discovered peptide-encoding genes having sequence similarity with *nlp* genes from the YGGWamide subfamily were also upregulated upon immunological challenge (Couillault et al., 2004). Searches of the *C. elegans* genome allowed the prediction of a family of closely related peptides, called caenacins (CNCs), which have, however, not (yet) been identified biochemically.

As the genetic work on the dauer pathway progressed, it became clear that this pathway was also involved with longevity. Kenyon et al. (1993) already showed that *daf-2* mutants live about twice as long as wild type. However, it took 4 years to establish the link between longevity and insulin

signaling by cloning and characterizing the insulin receptor *daf-2* (Kimura et al., 1997). Partial characterization of some of the ligands has been performed using deletion and overexpression studies (Pierce et al., 2001). Reduced signaling of insulin-like peptides increases the life-span of not only nematodes, but also of other organisms. A similar endocrine regulation of aging by insulin-like signals has been observed in flies (Tatar et al., 2001) and mice (Coschigano et al., 2000; Holzenberger et al., 2003). This implies that the insulin pathway's connection to lifespan is an ancient evolutionary invention that is conserved among many phyla. Despite the similarities among mammals and invertebrates in insulin-like peptides and their signaling pathways (Tatar et al., 2003), more research is needed to determine whether these signals control aging by similar mechanisms.

Phenotypical analysis of mutants lacking key enzymes needed for the processing of the peptide precursors, transport of the DCVs or exocytosis can yield valuable information about the importance of neuropeptide signaling. In *C. elegans*, CAPS/UNC-31 appears to be involved in the regulated secretion of a subset of neurotransmitters (Avery et al., 1993). This cytosolic protein is believed to be essential for the fusion of DCVs with the plasma membrane and the secretion of neuropeptides in *C. elegans* and *D. melanogaster* (Richmond and Broadie, 2002). *C. elegans* having mutations in *caps/unc-31* are lethargic, feed constitutively, are defective in egg-laying and produce dauer larvae that fail to recover (Avery et al., 1993). Mutations in the kinesin UNC-104/KIF1A, an essential motor for axonal transport of SVs and DCV to the synapse (Bloom, 2001; Zahn et al., 2004), result in a severely uncoordinated locomotion. Peptidomic analysis of different *unc-31* and *unc-104* mutant strains yielded a peptide map similar to that of the wild type *C. elegans* strain, indicating that the neuropeptides are still present in the DCV and that the observed phenotypes of these mutants are likely due to the impaired transport or secretion of neuropeptides from the DCVs (Husson et al., 2006).

Neuropeptides are also believed to modulate pre- and postsynaptic functions by modulating the release of neurotransmitters from presynaptic neurons or by downregulating their postsynaptic receptors. Therefore, neuropeptides, and thus the proprotein convertase gene *egl-3* may affect glutamatergic signaling by altering these pre-and postsynaptic functions. It has been shown that a mutation in *egl-3* not only suppresses the mechanosensory defects of animals having mutations in *glr-1*, a non-NMDA glutamate receptor subunit (Kass et al., 2001), but also restores osmotic avoidance responses (Mellem et al., 2002). The different signaling mechanisms of the polymodal ASH sensory neurons that release glutamate upon detection of mechanical, osmotic and chemical stimuli have elegantly been assessed by behavioral and electrophysiological studies of different mutants (Mellem et al., 2002). The polymodal signaling in *C. elegans* appears to occur via the differential activation of postsynaptic glutamate receptor subtypes. Since mutations in *egl-3* did not affect postsynaptic response to exogenously applied glutamate in either AVA or AVD interneurons, a presynaptic role for EGL-3 is suggested. Some critical neuropeptides might therefore function to modulate glutamatergic signaling from ASH.

6. Other neurotransmitters

Like in other animals, neuropeptides co-exist with a wide range of classic small molecule neurotransmitters (Brownlee and Fairweather, 1999). In addition to neurally directed roles, certain transmitters including adrenaline and noradrenaline might exhibit more hormone-like actions in nematodes. Progress in the biochemical identification, localization, synthesis and physiological actions of these transmitters has been reviewed elsewhere (Brownlee and Fairweather, 1999) and does not fall in the scope of this manuscript. Here, we briefly discuss signaling by monoamines in *C. elegans*.

Octopamine is a neurotransmitter found in many invertebrates and appears to play an important role in key processes of *C. elegans*. Exogenous octopamine inhibits egg-laying and pharyngeal pumping in contrast to serotonin, which stimulates egg-laying and pumping of the pharynx (Horvitz et al., 1982). The biosynthesis of octopamine requires the enzyme tyrosine decarboxylase (TDC) to convert tyrosine into tyramine, in addition to tyramine β-hydroxylase (TBH) to convert tyramine into octopamine. These two critical enzymes have recently been characterized in *C. elegans* by the Horvitz group (Alkema et al., 2005). Analysis of *C. elegans* strains having deletion alleles of the corresponding *tdc-1* and *tbh-1* genes shows that tyramine functions independently of octopamine and plays a specific role in the modulation of reversal behavior, the suppression of head oscillations in response to anterior touch and in the inhibition of egg-laying (Alkema et al., 2005). While putative octopamine receptors can only be tentatively identified in the *C. elegans* database, no octopamine receptors have been functionally characterized from any nematode. On the other hand, two tyramine receptors, SER-2 (Rex and Komuniecki, 2002; Rex et al., 2004) and TYRA-2 (Rex et al., 2005), have been functionally characterized from *C. elegans*, providing further evidence for tyramine-specific signaling in nematodes. The neurotransmitter serotonin or 5-hydroxytyramine (5-HT) is also widely used in the nervous system, mediating a wide variety of diverse responses by interacting with seven different 5-HT receptor subtypes in mammals. 5-HT mainly regulates locomotion, feeding behavior, pharyngeal pumping, mating behavior and egg-laying in *C. elegans*. Four nematode 5-HT receptors have been characterized: 5-HT_1- (SER-4 or formerly 5-HT-Ce) (Olde and McCombie, 1997), 5-HT_2- (SER-1) (Hamdan et al., 1999; Carnell et al., 2005) and 5-HT_7-like GPCRs (SER-7) (Hobson et al., 2003, 2006), as well as a 5-HT-gated Cl^- channel (MOD-1) (Ranganathan et al., 2000). In mammals, dopamine plays important roles in locomotion, reward, cognition and emotion. Dysfunctions of the dopaminergic system are linked to a variety of human disorders including Parkinson's disease, Tourette's syndrome, schizophrenia and drug addiction. In humans, five G-protein coupled dopamine receptors have been cloned (Missale et al., 1998). Based on their sequence similarities, biochemical properties and pharmacological profiles, they are divided into two

subfamilies: D1-like and D2-like receptors. In *C. elegans*, two D1-like receptors (DOP-1, Suo et al., 2002; Sanyal et al., 2004, and DOP-4, Sugiura et al., 2005) and two D2-like receptors (DOP-2, Suo et al., 2003, and DOP-3, Sugiura et al., 2005) have been cloned and characterized.

Although, each of these monoamines displays distinct and overlapping roles and functions in the worm's behavior, ACh is the only neurotransmitter that is essential for nematode viability. At the neuromuscular junction, excitatory neurons release ACh from bubble-like vesicles into the synaptic space, while inhibitory neurons release GABA. Two distinct types of cholinergic neurotransmission exist. Nicotinic receptors (nAChRs), which are pentameric cation ion channels, mediate a rapid transmission, whereas muscarinic ACh receptors (mAChRs) are typical GPCRs that enable a slow signal transduction pathway. Excess ACh is broken down by acetylcholinesterases, so that the synapse can be primed for the next signal. Analysis of the *C. elegans* genome revealed the largest known family of nAChR subunit genes in a single species, consisting of at least 27 members (Jones and Sattelle, 2004), while three genes (*gar-1*, *gar-2* and *gar-3*) have been shown to encode functional G-protein linked ACh receptors (Lee et al., 1999, 2000; Park et al., 2003). Steady-state ACh secretion in living nematodes can be assayed by measuring resistance to the acetylcholinesterase inhibitor aldicarb. Treatment of *C. elegans* with aldicarb causes accumulation of ACh at the neuromuscular junctions, leading to acute paralysis and subsequent death. To analyse synapse structure and function, approximately 10% of the *C. elegans* genome has been subjected to an impressive RNAi screen for aldicarb resistance, which yielded the identification of 185 genes (Sieburth et al., 2005). This collaborative effort was the first large-scale RNAi screen in neuroscience and forms an excellent example of innovative genetics that reaches the synapse (Bargmann, 2005).

7. Neuropeptidergic signaling in other organisms

In contrast to the wide variety of different FaRPs in invertebrates, only a few vertebrate -RFamide-like peptides are known while a tremendous variety of non-RFamide-like neuropeptides occur in mammals. At least five FaRP-encoding genes and at least five GPCRs on which they act have been found in mammals (Dockray, 2004). Immunoreactive material could be recorded by using antisera against the first discovered FMRFamide peptide from the clam *Macrocallista nimbosa* (Price and Greenberg, 1977a,b) and performing RIAs on tissue extracts of brain, gut and pancreas of various vertebrate species (chicken, frog, dog and rat) (Dockray et al., 1981). Immunocytochemical studies could demonstrate the peptide producing cells, but unraveling the primary sequence of the vertebrate -RFamide-like peptides was more challenging. The first vertebrate FaRP, LPLRFamide, was purified from brain tissues of chicken using antisera against the molluscan FMRFamide peptide (Dockray et al., 1983), followed by isolation of neuropeptide AF (NPAF) and neuropeptide FF (NPFF) from bovine brain (Yang et al., 1985). Later on, FaRP-encoding neuropeptide precursor genes could be identified in vertebrates. In order to simplify and avoid confusion about the different gene names, it is suggested that the annotations *farp-1* to *farp-5* are adopted to designate the genes in the order they were described (Dockray, 2004). The FaRP precursor gene, *farp-1*, encodes the two morphine modulating peptides NPFF and NPAF (Perry et al., 1997). Although the second gene, *farp-2*, encodes the so-called prolactin-releasing peptide (PrRP) (Hinuma et al., 1998), this peptide has also been proven to regulate feeding, energy homeostatis, stress and blood pressure (Fukusumi et al., 2006). The human RFamide-related peptide (hRFRP) precursor gene (Hinuma et al., 2000) or *farp-3* was predicted to encode at least three RFRPs. These peptides have been reported to stimulate prolactin secretion in rats and an involvement in pain modulation has been suggested (Fukusumi et al., 2006). The metastasis suppressor gene *kiss-1* (Lee et al., 1996) or *farp-4* was originally recognized as a tumor metastasis suppressor. Therefore, the derived product was initially referred to as metastin. The encoded RFamide peptide is now named kisspeptin and appears to be an important regulator of the reproductive axis (Kriegsfeld, 2006). The bioactive peptide from *farp-5* has differently been annotated as 26RFamide or P518 and QRFP (Jiang et al., 2003; Chartrel et al., 2003) and might be involved in food uptake (Chartrel et al., 2003, 2006; Dockray, 2004). All vertebrate RFamide-containing peptides that have been discovered could already be linked to at least one receptor.

Insulin is a peptide hormone with a remarkable historical record and has played a key role in medical and biological research. A decreased production or release of biologically active insulin results in an impaired carbohydrate metabolism, which can lead to diabetes. Although the structure and function of vertebrate insulins are strictly conserved, insulin-related peptides from invertebrates appear to be a structurally diverse group, encoded by large multi-gene families. While 40 *ins* genes (*ins-1* to *ins-39* and *daf-28*) have been predicted in *C. elegans*, seven molluscan insulin-related peptide (MIP) genes appear to be present in *Lymnea stagnalis* (Smit et al., 1998). Only one gene encoding the *Aplysia californica* insulin has been characterized (Floyd et al., 1999), in contrast to the surprising discovery of 38 different bombyxin or insulin-related peptide genes in *Bombyx mori* (Kondo et al., 1996). Seven *Drosophila* insulin-like peptide genes (*dilp1–7)* could be identified from the *D. melanogaser* genome database (Vanden Broeck, 2001; Brogiolo et al., 2001), while an equal number of insulin-like peptide genes (*AgamILP1–7*) have been predicted from the genome of the African malaria mosquito *Anopheles gambiae* (Riehle et al., 2002; Krieger et al., 2004). These invertebrate insulin-like peptides appear to be involved in growth, reproduction and aging, as recently reviewed (Wu and Brown, 2006).

Endogenous *C. elegans* peptides that do not contain the carboxyterminal RFamide moiety or lack sequence resemblance with insulin are contained within the NLP family of (neuro)peptides. Detailed comparative sequence analysis of some *C. elegans* NLP peptides with bioactive peptides from other species reveal profound similarities with insect neuropeptides

like periviscerokinins, leucokinins, tachykinins and allatostatins. For the latter group, no vertebrate orthologous peptides have been found (yet), although a huge variety of other non-RFamide-like neuropeptides appear to be present in mammals. This indicates that the concept of peptide-driven signal transduction and their players are evolutionary conserved, especially among the invertebrates. Not only neuropeptidergic signaling but also the mechanisms of proprotein peptide precursor processing, transport, exocytosis and classic small-molecule neurotransmitter functions are highly conserved in the nematode and vertebrate nervous system, although they diverged over 800 million years ago. This conservation indicates that the nervous system and its signaling components were not gradually perfected as more complex organisms evolved. Unraveling peptide sequences from nematodes and characterizing their corresponding GPCRs will obviously stimulate peptide research of other organisms.

8. Future prospects

Neuropeptides play many important roles in the regulation of physiological processes and behavior in both vertebrates and invertebrates. These regulatory peptides are processed, stored and released within the nervous system and can act through G-protein coupled receptors or peptide-gated ion channels to govern homeostatic processes in response to internal and external stimuli. This emphasizes the importance of detailed knowledge of the full complement of the wide diversity of actually expressed neuropeptides in relation to their roles in neuronal communication. In general, the analysis of biochemical entities like endogenous peptides, which are deduced from preproproteins or peptide precursors and thus from genes, is significantly more challenging than genome analysis. Peptide diversity can in principle be generated by cell-specific expression of distinct peptide precursor genes and their processing enzymes, alternative splicing of the resulting mRNAs, and alternative proteolytic processing of the precursor proteins, and, finally, cell-specific post-translational modifications can occur. Therefore, the neuropeptide complement of a cell, tissue or organism, displaying in vivo concentrations in the nano-molar range, is spatially and temporally dynamic. Moreover, lacking an amplification technique like the polymerase chain reaction that is widely used in genomic studies makes it extremely hard to identify and characterize individual peptides or to monitor whole peptide profiles from a tissue or organism.

Structural, physiological and expression data of bioactive peptides accumulate and will obviously stimulate the characterization of remaining orphan GPCRs in *C. elegans* since the success of a ''reversed pharmacology approach'' highly depends on the correctness of the sequences of the peptides used. These receptors are emerging as strong drug target candidates and can be handled in high throughput screening programs to identify useful peptide mimetics (Civelli et al., 1999; Geary and Kubiak, 2005). Since the market for anthelmintics is plagued by resistance, many efforts continuously have to be undertaken in order to discover novel key drug targets (Geary et al., 1999b). The ever-expanding family of FLP peptides is not limited to nematodes (McVeigh et al., 2005b), as platyhelminths or flatworms (Shaw et al., 1996; McVeigh et al., 2005a) also possess a wide variety of peptidergic signaling systems. Knowledge of these neuropeptidergic signaling pathways can contribute to the design of novel drug targets (Geary et al., 1995; Maule et al., 2002; Kimber and Fleming, 2005). These data suggest that there still is a great potential, which can be used to uncover a new generation of anthelmintics. The widespread occurrence of endogenous neuropeptides, which can act across the phylum barriers and their conserved sequence similarity, will also favor the development of broad-spectrum anthelmintics (Mousley et al., 2005). When screening for potential drug targets like endogenous peptides or their GPCRs, however, one has to consider potential cross-reactivities with mammalian receptors. Only nematode-specific systems can of course be attributed as potential targets for the development of nematode pharmacology.

The large number of similarities among the classic small molecule neurotransmission, biogenic amine receptor signaling pathways and neuropeptidergic signaling in *C. elegans* with mammalian counterparts suggests that the nematode can serve as an important model organism to study neurotransmission-related behaviors and disorders. Although recent behavioral studies indicate the worm is capable of quite sophisticated behaviors, the underlying signaling systems might be even more complex. The genetic, molecular and biochemical analysis of neuropeptidergic signaling strategies in *C. elegans* will obviously complement neuropeptide studies in other species and address the conservation of their signaling systems.

Acknowledgments

This project was sponsored by a grant of the Fund for Scientific Research-Flanders (FWO-Vlaanderen, G.0270.04). The authors wish to thank J. Puttemans for careful help with the artwork. S.J. Husson and M. Lindemans benefit from a Ph.D. grant of the Institute for Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). I. Mertens is a postdoctoral fellow of the Fund for Scientific Research-Flanders (FWO-Vlaanderen) and T. Janssen benefits from a Ph.D. scholarship of the FWO-Vlaanderen.

References

Alkema, M.J., Hunter-Ensor, M., Ringstad, N., Horvitz, H.R., 2005. Tyramine functions independently of octopamine in the *Caenorhabditis elegans* nervous system. Neuron 46, 247–260.

Atkinson, H.J., Isaac, R.E., Harris, P.D., Sharpe, C.M., 1988. FMRFamide-like immunoreactivity within the nervous system of the nematodes *Panagrellus redivivus*, *Caenorhabditis elegans* and *Heterodera glycines*. J. Zool. 216, 663–671.

Avery, L., Bargmann, C.I., Horvitz, H.R., 1993. The *Caenorhabditis elegans unc-31* gene affects multiple nervous system-controlled functions. Genetics 134, 455–464.

Baggerman, G., Boonen, K., Verleyen, P., De Loof, A., Schoofs, L., 2005. Peptidomic analysis of the larval *Drosophila melanogaster* central nervous

system by two-dimensional capillary liquid chromatography quadrupole time-of-flight mass spectrometry. J. Mass Spectrom. 40, 250–260.
Baggerman, G., Cerstiaens, A., De Loof, A., Schoofs, L., 2002. Peptidomics of the larval *Drosophila melanogaster* central nervous system. J. Biol. Chem. 277, 40368–40374.
Baggerman, G., Verleyen, P., Clynen, E., Huybrechts, J., De Loof, A., Schoofs, L., 2004. Peptidomics. J. Chromatogr. B 803, 3–16.
Bargmann, C., 2005. Neuroscience: genomics reaches the synapse. Nature 436, 473–474.
Bargmann, C.I., 1998. Neurobiology of the *Caenorhabditis elegans* genome. Science 282, 2028–2033.
Beinfeld, M.C., 1998. Prohormone and proneuropeptide processing. Recent progress and future challenges. Endocrine 8, 1–5.
Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340, 783–795.
Beyenbach, K.W., 2003. Transport mechanisms of diuresis in Malpighian tubules of insects. J. Exp. Biol. 206, 3845–3856.
Blackburn, M.B., Wagner, R.M., Kochansky, J.P., Harrison, D.J., Thomas-Laemont, P., Raina, A.K., 1995. The identification of two myoinhibitory peptides, with sequence similarities to the galanins, isolated from the ventral nerve cord of *Manduca sexta*. Regul. Pept. 57, 213–219.
Bloom, G.S., 2001. The UNC-104/KIF1 family of kinesins. Curr. Opin. Cell Biol. 13, 36–40.
Brenner, S., 1974. The genetics of *Caenorhabditis elegans*. Genetics 77, 71–94.
Brogiolo, W., Stocker, H., Ikeya, T., Rintelen, F., Fernandez, R., Hafen, E., 2001. An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth control. Curr. Biol. 11, 213–221.
Brooks, D.R., Appleford, P.J., Murray, L., Isaac, R.E., 2003. An essential role in molting and morphogenesis of *Caenorhabditis elegans* for ACN-1, a novel member of the angiotensin-converting enzyme family that lacks a metallopeptidase active site. J. Biol. Chem. 278, 52340–52346.
Brownlee, D.J., Fairweather, I., 1999. Exploring the neurotransmitter labyrinth in nematodes. Trends Neurosci. 22, 16–24.
Brownlee, D.J., Walker, R.J., 1999. Actions of nematode FMRFamide-related peptides on the pharyngeal muscle of the parasitic nematode, *Ascaris suum*. Ann. N.Y. Acad. Sci. 897, 228–238.
Canaff, L., Bennett, H.P., Hendy, G.N., 1999. Peptide hormone precursor processing: getting sorted? Mol. Cell Endocrinol. 156, 1–6.
Carnell, L., Illi, J., Hong, S.W., McIntire, S.L., 2005. The G-protein-coupled serotonin receptor SER-1 regulates egg laying and male mating behaviors in *Caenorhabditis elegans*. J. Neurosci. 25, 10671–10681.
Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., Prasher, D.C., 1994. Green fluorescent protein as a marker for gene expression. Science 263, 802–805.
Chartrel, N., Bruzzone, F., Leprince, J., Tollemer, H., Anouar, Y., Do-Rego, J.C., Segalas-Milazzo, I., Guilhaudis, L., Cosette, P., Jouenne, T., Simonnet, G., Vallarino, M., Beauvillain, J.C., Costentin, J., Vaudry, H., 2006. Structure and functions of the novel hypothalamic RFamide neuropeptides R-RFa and 26RFa in vertebrates. Peptides 27, 1110–1120.
Chartrel, N., Dujardin, C., Anouar, Y., Leprince, J., Decker, A., Clerens, S., Do-Rego, J.C., Vandesande, F., Llorens-Cortes, C., Costentin, J., Beauvillain, J.C., Vaudry, H., 2003. Identification of 26RFa, a hypothalamic neuropeptide of the RFamide peptide family with orexigenic activity. Proc. Natl. Acad. Sci. U.S.A. 100, 15247–15252.
Cheung, B.H., rellano-Carbajal, F., Rybicki, I., de, B.M., 2004. Soluble guanylate cyclases act in neurons exposed to the body fluid to promote *C. elegans* aggregation behavior. Curr. Biol. 14, 1105–1111.
Civelli, O., Reinscheid, R.K., Nothacker, H.P., 1999. Orphan receptors, novel neuropeptides and reverse pharmaceutical research. Brain Res. 848, 63–65.
Clynen, E., Baggerman, G., Veelaert, D., Cerstiaens, A., Van Der Horst, D., Harthoorn, L., Derua, R., Waelkens, E., De Loof, A., Schoofs, L., 2001. Peptidomics of the pars intercerebralis-corpus cardiacum complex of the migratory locust, *Locusta migratoria*. Eur. J. Biochem. 268, 1929–1939.
Clynen, E., De Loof, A., Schoofs, L., 2003. The use of peptidomics in endocrine research. Gen. Comp. Endocrinol. 132, 1–9.
Coates, D., Siviter, R., Isaac, R.E., 2000. Exploring the *Caenorhabditis elegans* and *Drosophila melanogaster* genomes to understand neuropeptide and peptidase function. Biochem. Soc. Trans. 28, 464–469.
Coates, J.C., de Bono, M., 2002. Antagonistic pathways in neurons exposed to body fluid regulate social feeding in *Caenorhabditis elegans*. Nature 419, 925–929.
Coschigano, K.T., Clemmons, D., Bellush, L.L., Kopchick, J.J., 2000. Assessment of growth parameters and life span of GHR/BP gene-disrupted mice. Endocrinology 141, 2608–2613.
Couillault, C., Pujol, N., Reboul, J., Sabatier, L., Guichou, J.F., Kohara, Y., Ewbank, J.J., 2004. 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, 488–494.
Cowden, C., Stretton, A.O., 1993. AF2, an *Ascaris* neuropeptide: isolation, sequence, and bioactivity. Peptides 14, 423–430.
Cowden, C., Stretton, A.O., 1995. Eight novel FMRFamide-like neuropeptides isolated from the nematode *Ascaris suum*. Peptides 16, 491–500.
Cuppen, E., van der Linden, A.M., Jansen, G., Plasterk, R.H.A., 2003. Proteins interacting with *Caenorhabditis elegans* Galpha subunits. Comp. Funct. Genom. 4, 479–491.
Davies, A.G., Bettinger, J.C., Thiele, T.R., Judy, M.E., McIntire, S.L., 2004. Natural variation in the *npr-1* gene modifies ethanol responses of wild strains of *C. elegans*. Neuron 42, 731–743.
de Bono, M., Bargmann, C.I., 1998. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. Cell 94, 679–689.
de Bono, M., Tobin, D.M., Davis, M.W., Avery, L., Bargmann, C.I., 2002. Social feeding in *Caenorhabditis elegans* is induced by neurons that detect aversive stimuli. Nature 419, 899–903.
Dlakic, M., 2002. A new family of putative insulin receptor-like proteins in *C. elegans*. Curr. Biol. 12, 155–157.
Dockray, G.J., 2004. The expanding family of -RFamide peptides and their effects on feeding behaviour. Exp. Physiol. 89, 229–235.
Dockray, G.J., Reeve Jr., J.R., Shively, J., Gayton, R.J., Barnard, C.S., 1983. A novel active pentapeptide from chicken brain identified by antibodies to FMRFamide. Nature 305, 328–330.
Dockray, G.J., Vaillant, C., Williams, R.G., Gayton, R.J., Osborne, N.N., 1981. Vertebrate brain-gut peptides related to FMRFamide and Met-enkephalin Arg6Phe7. Peptides 2 (Suppl. 2), 25–30.
Duckert, P., Brunak, S., Blom, N., 2004. Prediction of proprotein convertase cleavage sites. Protein Eng. Des. Sel. 17, 107–112.
Duret, L., Guex, N., Peitsch, M.C., Bairoch, A., 1998. New insulin-like proteins with atypical disulfide bond pattern characterized in *Caenorhabditis elegans* by comparative sequence analysis and homology modeling. Genome Res. 8, 348–353.
Eipper, B.A., Milgram, S.L., Husten, E.J., Yun, H.Y., Mains, R.E., 1993. Peptidylglycine alpha-amidating monooxygenase: a multifunctional protein with catalytic, processing, and routing domains. Protein Sci. 2, 489–497.
Floyd, P.D., Li, L., Rubakhin, S.S., Sweedler, J.V., Horn, C.C., Kupfermann, I., Alexeeva, V.Y., Ellis, T.A., Dembrow, N.C., Weiss, K.R., Vilim, F.S., 1999. Insulin prohormone processing, distribution, and relation to metabolism in *Aplysia californica*. J. Neurosci. 19, 7732–7741.
Fricker, L.D., 2005. Neuropeptide-processing enzymes: applications for drug discovery. AAPS J. 7, E449–E455.
Fricker, L.D., Lim, J., Pan, H., Che, F.Y., 2006. Peptidomics: identification and quantification of endogenous peptides in neuroendocrine tissues. Mass Spectrom. Rev. 25, 327–344.
Fukusumi, S., Fujii, R., Hinuma, S., 2006. Recent advances in mammalian RFamide peptides: the discovery and functional analyses of PrRP, RFRPs and QRFP. Peptides 27, 1073–1086.
Furuta, M., Yano, H., Zhou, A., Rouille, Y., Holst, J.J., Carroll, R., Ravazzola, M., Orci, L., Furuta, H., Steiner, D.F., 1997. Defective prohormone processing and altered pancreatic islet morphology in mice lacking active SPC2. Proc. Natl. Acad. Sci. U.S.A. 94, 6646–6651.
Gasnier, B., 2000. The loading of neurotransmitters into synaptic vesicles. Biochimie 82, 327–337.
Geary, T.G., Bowman, J.W., Friedman, A.R., Maule, A.G., Davis, J.P., Winterrowd, C.A., Klein, R.D., Thompson, D.P., 1995. The pharmacology of FMRFamide-related neuropeptides in nematodes: new opportunities for rational anthelmintic discovery? Int. J. Parasitol. 25, 1273–1280.

Geary, T.G., Kubiak, T.M., 2005. Neuropeptide G-protein-coupled receptors, their cognate ligands and behavior in *Caenorhabditis elegans*. Trends Pharmacol. Sci. 26, 56–58.

Geary, T.G., Marks, N.J., Maule, A.G., Bowman, J.W., Alexander-Bowman, S.J., Day, T.A., Larsen, M.J., Kubiak, T.M., Davis, J.P., Thompson, D.P., 1999a. Pharmacology of FMRFamide-related peptides in helminths. Ann. N.Y. Acad. Sci. 897, 212–227.

Geary, T.G., Sangster, N.C., Thompson, D.P., 1999b. Frontiers in anthelmintic pharmacology. Vet. Parasitol. 84, 275–295.

Goldstein, L.S., Yang, Z., 2000. Microtubule-based transport systems in neurons: the roles of kinesins and dyneins. Annu. Rev. Neurosci. 23, 39–71.

Gomez-Saladin, E., Luebke, A.E., Wilson, D.L., Dickerson, I.M., 1997. Isolation of a cDNA encoding a Kex2-like endoprotease with homology to furin from the nematode *Caenorhabditis elegans*. DNA Cell Biol. 16, 663–669.

Gomez-Saladin, E., Wilson, D.L., Dickerson, I.M., 1994. Isolation and in situ localization of a cDNA encoding a Kex2-like prohormone convertase in the nematode *Caenorhabditis elegans*. Cell Mol. Neurobiol. 14, 9–25.

Gray, J.M., Karow, D.S., Lu, H., Chang, A.J., Chang, J.S., Ellis, R.E., Marletta, M.A., Bargmann, C.I., 2004. Oxygen sensation and social feeding mediated by a *C. elegans* guanylate cyclase homologue. Nature 430, 317–322.

Gregoire, F.M., Chomiki, N., Kachinskas, D., Warden, C.H., 1998. Cloning and developmental regulation of a novel member of the insulin-like gene family in *Caenorhabditis elegans*. Biochem. Biophys. Res. Commun. 249, 385–390.

Hamdan, F.F., Ungrin, M.D., Abramovitz, M., Ribeiro, P., 1999. Characterization of a novel serotonin receptor from *Caenorhabditis elegans*: cloning and expression of two splice variants. J. Neurochem. 72, 1372–1383.

Hauser, F., Williamson, M., Grimmelikhuijzen, C.J., 1997. Molecular cloning of a peptidylglycine alpha-hydroxylating monooxygenase from sea anemones. Biochem. Biophys. Res. Commun. 241, 509–512.

Hayes, T.K., Pannabecker, T.L., Hinckley, D.J., Holman, G.M., Nachman, R.J., Petzel, D.H., Beyenbach, K.W., 1989. Leucokinins, a new family of ion transport stimulators and inhibitors in insect Malpighian tubules. Life Sci. 44, 1259–1266.

Hinuma, S., Habata, Y., Fujii, R., Kawamata, Y., Hosoya, M., Fukusumi, S., Kitada, C., Masuo, Y., Asano, T., Matsumoto, H., Sekiguchi, M., Kurokawa, T., Nishimura, O., Onda, H., Fujino, M., 1998. A prolactin-releasing peptide in the brain. Nature 393, 272–276.

Hinuma, S., Shintani, Y., Fukusumi, S., Iijima, N., Matsumoto, Y., Hosoya, M., Fujii, R., Watanabe, T., Kikuchi, K., Terao, Y., Yano, T., Yamamoto, T., Kawamata, Y., Habata, Y., Asada, M., Kitada, C., Kurokawa, T., Onda, H., Nishimura, O., Tanaka, M., Ibata, Y., Fujino, M., 2000. New neuropeptides containing carboxy-terminal RFamide and their receptor in mammals. Nat. Cell Biol. 2, 703–708.

Hobson, R.J., Geng, J., Gray, A.D., Komuniecki, R.W., 2003. SER-7b, a constitutively active Galphas coupled 5-HT7-like receptor expressed in the *Caenorhabditis elegans* M4 pharyngeal motorneuron. J. Neurochem. 87, 22–29.

Hobson, R.J., Hapiak, V.M., Xiao, H., Buehrer, K.L., Komuniecki, P.R., Komuniecki, R.W., 2006. SER-7, a *Caenorhabditis elegans* 5-HT7-like receptor, is essential for the 5-HT stimulation of pharyngeal pumping and egg laying. Genetics 172, 159–169.

Holden-Dye, L., Franks, C.J., Williams, R.G., Walker, R.J., 1995. The effect of the nematode peptides SDPNFLRFamide (PF1) and SADPNFLRFamide (PF2) on synaptic transmission in the parasitic nematode *Ascaris suum*. Parasitology 110, 449–455.

Holzenberger, M., Dupont, J., Ducos, B., Leneuve, P., Geloen, A., Even, P.C., Cervera, P., Le Bouc, Y., 2003. IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. Nature 421, 182–187.

Horvitz, H.R., Chalfie, M., Trent, C., Sulston, J.E., Evans, P.D., 1982. Serotonin and octopamine in the nematode *Caenorhabditis elegans*. Science 216, 1012–1014.

Hua, Q.X., Nakagawa, S.H., Wilken, J., Ramos, R.R., Jia, W., Bass, J., Weiss, M.A., 2003. A divergent INS protein in *Caenorhabditis elegans* structurally resembles human insulin and activates the human insulin receptor. Genes Dev. 17, 826–831.

Hua, Y.J., Tanaka, Y., Nakamura, K., Sakakibara, M., Nagata, S., Kataoka, H., 1999. Identification of a prothoracicostatic peptide in the larval brain of the silkworm, *Bombyx mori*. J. Biol. Chem. 274, 31169–31173.

Husson, S.J., Clynen, E., Baggerman, G., De Loof, A., Schoofs, L., 2005. Discovering neuropeptides in *Caenorhabditis elegans* by two dimensional liquid chromatography and mass spectrometry. Biochem. Biophys. Res. Commun. 335, 76–86.

Husson, S.J., Clynen, E., Baggerman, G., Janssen, T., Schoofs, L., 2006. Defective processing of neuropeptide precursors in *Caenorhabditis elegans* lacking proprotein convertase 2 (KPC-2/EGL-3): mutant analysis by mass spectrometry. J. Neurochem. 98, 1999–2012.

Isaac, R.E., Siviter, R.J., Stancombe, P., Coates, D., Shirras, A.D., 2000. Conserved roles for peptidases in the processing of invertebrate neuropeptides. Biochem. Soc. Trans. 28, 460–464.

Jacob, T.C., Kaplan, J.M., 2003. The EGL-21 carboxypeptidase E facilitates acetylcholine release at *Caenorhabditis elegans* neuromuscular junctions. J. Neurosci. 23, 2122–2130.

Jansen, G., Thijssen, K.L., Werner, P., van der, H.M., Hazendonk, E., Plasterk, R.H., 1999. The complete family of genes encoding G proteins of *Caenorhabditis elegans*. Nat. Genet. 21, 414–419.

Jiang, Y., Luo, L., Gustafson, E.L., Yadav, D., Laverty, M., Murgolo, N., Vassileva, G., Zeng, M., Laz, T.M., Behan, J., Qiu, P., Wang, L., Wang, S., Bayne, M., Greene, J., Monsma Jr., F., Zhang, F.L., 2003. Identification and characterization of a novel RF-amide peptide ligand for orphan G-protein-coupled receptor SP9155. J. Biol. Chem. 278, 27652–27657.

Jones, A.K., Sattelle, D.B., 2004. Functional genomics of the nicotinic acetylcholine receptor gene family of the nematode, *Caenorhabditis elegans*. Bioessays 26, 39–49.

Julius, D., Brake, A., Blair, L., Kunisawa, R., Thorner, J., 1984. Isolation of the putative structural gene for the lysine-arginine-cleaving endopeptidase required for processing of yeast prepro-alpha-factor. Cell 37, 1075–1089.

Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D.P., Zipperlen, P., Ahringer, J., 2003. Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. Nature 421, 231–237.

Kass, J., Jacob, T.C., Kim, P., Kaplan, J.M., 2001. The EGL-3 proprotein convertase regulates mechanosensory responses of *Caenorhabditis elegans*. J. Neurosci. 21, 9265–9272.

Keating, C.D., Kriek, N., Daniels, M., Ashcroft, N.R., Hopper, N.A., Siney, E.J., Holden-Dye, L., Burke, J.F., 2003. Whole-genome analysis of 60 G protein-coupled receptors in *Caenorhabditis elegans* by gene knockout with RNAi. Curr. Biol. 13, 1715–1720.

Kenyon, C., Chang, J., Gensch, E., Rudner, A., Tabtiang, R., 1993. A *C. elegans* mutant that lives twice as long as wild type. Nature 366, 461–464.

Kim, K., Li, C., 2004. Expression and regulation of an FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*. J. Comp. Biol. 475, 540–550.

Kimber, M.J., Fleming, C.C., 2005. Neuromuscular function in plant parasitic nematodes: a target for novel control strategies? Parasitology 131, S129–S142.

Kimura, K.D., Tissenbaum, H.A., Liu, Y., Ruvkun, G., 1997. *daf-2*, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. Science 277, 942–946.

Kolhekar, A.S., Roberts, M.S., Jiang, N., Johnson, R.C., Mains, R.E., Eipper, B.A., Taghert, P.H., 1997. Neuropeptide amidation in *Drosophila*: separate genes encode the two enzymes catalyzing amidation. J. Neurosci. 17, 1363–1376.

Kondo, H., Ino, M., Suzuki, A., Ishizaki, H., Iwami, M., 1996. Multiple gene copies for bombyxin, an insulin-related peptide of the silkmoth *Bombyx mori*: structural signs for gene rearrangement and duplication responsible for generation of multiple molecular forms of bombyxin. J. Mol. Biol. 259, 926–937.

Koushika, S.P., Nonet, M.L., 2000. Sorting and transport in *C. elegans*: a model system with a sequenced genome. Curr. Opin. Cell Biol. 12, 517–523.

Krieger, M.J., Jahan, N., Riehle, M.A., Cao, C., Brown, M.R., 2004. Molecular characterization of insulin-like peptide genes and their expression in the African malaria mosquito, *Anopheles gambiae*. Insect Mol. Biol. 13, 305–315.

Kriegsfeld, L.J., 2006. Driving reproduction: RFamide peptides behind the wheel. Horm. Behav. 50, 655–666.

Kubiak, T.M., Larsen, M.J., Nulf, S.C., Zantello, M.R., Burton, K.J., Bowman, J.W., Modric, T., Lowery, D.E., 2003a. 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, 33724–33729.

Kubiak, T.M., Larsen, M.J., Zantello, M.R., Bowman, J.W., Nulf, S.C., Lowery, D.E., 2003b. Functional annotation of the putative orphan *Caenorhabditis elegans* G-protein-coupled receptor C10C6.2 as a FLP15 peptide receptor. J. Biol. Chem. 278, 42115–42120.

Lamango, N.S., Nachman, R., Hayes, T.K., Strey, A., Isaac, R.E., 1997. Hydrolysis of insect neuropeptides by an angiotensin-converting enzyme from the housefly, *Musca domestica*. Peptides 18, 47–52.

Lee, J.H., Miele, M.E., Hicks, D.J., Phillips, K.K., Trent, J.M., Weissman, B.E., Welch, D.R., 1996. KiSS-1, a novel human malignant melanoma metastasis-suppressor gene. J. Natl. Cancer Inst. 88, 1731–1737.

Lee, Y.S., Park, Y.S., Chang, D.J., Hwang, J.M., Min, C.K., Kaang, B.K., Cho, N.J., 1999. Cloning and expression of a G protein-linked acetylcholine receptor from *Caenorhabditis elegans*. J. Neurochem. 72, 58–65.

Lee, Y.S., Park, Y.S., Nam, S., Suh, S.J., Lee, J., Kaang, B.K., Cho, N.J., 2000. Characterization of GAR-2, a novel G protein-linked acetylcholine receptor from *Caenorhabditis elegans*. J. Neurochem. 75, 1800–1809.

Li, C., 2005. The ever-expanding neuropeptide gene families in the nematode *Caenorhabditis elegans*. Parasitology 131, S109–S127.

Li, C., Chalfie, M., 1990. Organogenesis in *C. elegans*: positioning of neurons and muscles in the egg-laying system. Neuron 4, 681–695.

Li, C., Kim, K., Nelson, L.S., 1999a. FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*. Brain Res. 848, 26–34.

Li, C., Nelson, L.S., Kim, K., Nathoo, A., Hart, A.C., 1999b. Neuropeptide gene families in the nematode *Caenorhabditis elegans*. Ann. N.Y. Acad. Sci. 897, 239–252.

Li, W., Kennedy, S.G., Ruvkun, G., 2003. *daf-28* encodes a *C. elegans* insulin superfamily member that is regulated by environmental cues and acts in the DAF-2 signaling pathway. Genes Dev. 17, 844–858.

Lindberg, I., Tu, B., Muller, L., Dickerson, I.M., 1998. Cloning and functional analysis of *C. elegans* 7B2. DNA Cell Biol. 17, 727–734.

Liu, F., Baggerman, G., D'Hertog, W., Verleyen, P., Schoofs, L., Wets, G., 2006a. In silico identification of new secretory peptide genes in *Drosophila melanogaster*. Mol. Cell Proteomics 5, 510–522.

Liu, F., Baggerman, G., Schoofs, L., Wets, G., 2006b. Uncovering conserved patterns in bioactive peptides in Metazoa. Peptides 27, 3137–3153.

Lorenz, M.W., Kellner, R., Hoffmann, K.H., 1995. Identification of two allatostatins from the cricket, *Gryllus bimaculatus* de Geer (Ensifera, Gryllidae): additional members of a family of neuropeptides inhibiting juvenile hormone biosynthesis. Regul. Pept. 57, 227–236.

Lorenz, M.W., Kellner, R., Hoffmann, K.H., Gade, G., 2000. Identification of multiple peptides homologous to cockroach and cricket allatostatins in the stick insect *Carausius morosus*. Insect Biochem. Mol. Biol. 30, 711–718.

Lowery, D.E., Geary, T.G., Kubiak, T.M., Larsen, M.J., 2003. Pharmacia and Upjohn Company, G protein-coupled receptor-like receptors and modulators thereof. United States Patent 6632621.

Mair, G.R., Niciu, M.J., Stewart, M.T., Brennan, G., Omar, H., Halton, D.W., Mains, R., Eipper, B.A., Maule, A.G., Day, T.A., 2004. A functionally atypical amidating enzyme from the human parasite *Schistosoma mansoni*. FASEB J. 18, 114–121.

Marks, N.J., Maule, A.G., Geary, T.G., Thompson, D.P., Davis, J.P., Halton, D.W., Verhaert, P., Shaw, C., 1997. APEASPFIRFamide, a novel FMRFamide-related decapeptide from *Caenorhabditis elegans*: structure and myoactivity. Biochem. Biophys. Res. Commun. 231, 591–595.

Marks, N.J., Maule, A.G., Geary, T.G., Thompson, D.P., Li, C., Halton, D.W., Shaw, C., 1998. KSAYMRFamide (PF3/AF8) is present in the free-living nematode, *Caenorhabditis elegans*. Biochem. Biophys. Res. Commun. 248, 422–425.

Marks, N.J., Maule, A.G., Li, C., Nelson, L.S., Thompson, D.P., Alexander-Bowman, S., Geary, T.G., Halton, D.W., Verhaert, P., Shaw, C., 1999a. Isolation, pharmacology and gene organization of KPSFVRFamide: a neuropeptide from *Caenorhabditis elegans*. Biochem. Biophys. Res. Commun. 254, 222–230.

Marks, N.J., Sangster, N.C., Maule, A.G., Halton, D.W., Thompson, D.P., Geary, T.G., Shaw, C., 1999b. Structural characterisation and pharmacology of KHEYLRFamide (AF2) and KSAYMRFamide (PF3/AF8) from *Haemonchus contortus*. Mol. Biochem. Parasitol. 100, 185–194.

Marks, N.J., Shaw, C., Halton, D.W., Thompson, D.P., Geary, T.G., Li, C., Maule, A.G., 2001. Isolation and preliminary biological assessment of AADGAPLIRFamide and SVPGVLRFamide from *Caenorhabditis elegans*. Biochem. Biophys. Res. Commun. 286, 1170–1176.

Marks, N.J., Shaw, C., Maule, A.G., Davis, J.P., Halton, D.W., Verhaert, P., Geary, T.G., Thompson, D.P., 1995. Isolation of AF2 (KHEYLRFamide) from *Caenorhabditis elegans*: evidence for the presence of more than one FMRFamide-related peptide-encoding gene. Biochem. Biophys. Res. Commun. 217, 845–851.

Maule, A.G., Geary, T.G., Bowman, J.W., Marks, N.J., Blair, K.L., Halton, D.W., Shaw, C., Thompson, D.P., 1995. Inhibitory effects of nematode FMRFamide-related peptides (FaRPs) on muscle strips from *Ascaris suum*. Invert. Neurosci. 1, 255–265.

Maule, A.G., Mousley, A., Marks, N.J., Day, T.A., Thompson, D.P., Geary, T.G., Halton, D.W., 2002. Neuropeptide signaling systems—potential drug targets for parasite and pest control. Curr. Top. Med. Chem. 2, 733–758.

Mbikay, M., Seidah, N.G., Chretien, M., 2001. Neuroendocrine secretory protein 7B2: structure, expression and functions. Biochem. J. 357, 329–342.

McCudden, C.R., Hains, M.D., Kimple, R.J., Siderovski, D.P., Willard, F.S., 2005. G-protein signaling: back to the future. Cell Mol. Life Sci. 62, 551–577.

McKay, S.J., Johnsen, R., Khattra, J., Asano, J., Baillie, D.L., Chan, S., Dube, N., Fang, L., Goszczynski, B., Ha, E., Halfnight, E., Hollebakken, R., Huang, P., Hung, K., Jensen, V., Jones, S.J., Kai, H., Li, D., Mah, A., Marra, M., McGhee, J., Newbury, R., Pouzyrev, A., Riddle, D.L., Sonnhammer, E., Tian, H., Tu, D., Tyson, J.R., Vatcher, G., Warner, A., Wong, K., Zhao, Z., Moerman, D.G., 2004. Gene expression profiling of cells, tissues, and developmental stages of the nematode *C. elegans*. Cold Spring Harb. Symp. Quant. Biol. 68, 159–169.

McVeigh, P., Geary, T.G., Marks, N.J., Maule, A.G., 2006. The FLP-side of nematodes. Trends Parasitol. 22, 385–396.

McVeigh, P., Kimber, M.J., Novozhilova, E., Day, T.A., 2005a. Neuropeptide signalling systems in flatworms. Parasitology 131, S41–S55.

McVeigh, P., Leech, S., Mair, G.R., Marks, N.J., Geary, T.G., Maule, A.G., 2005b. Analysis of FMRFamide-like peptide (FLP) diversity in phylum Nematoda. Int. J. Parasitol. 35, 1043–1060.

Meeusen, T., Mertens, I., De Loof, A., Schoofs, L., 2003. G protein-coupled receptors in invertebrates: a state of the art. Int. Rev. Cytol. 230, 189–261.

Mellem, J.E., Brockie, P.J., Zheng, Y., Madsen, D.M., Maricq, A.V., 2002. Decoding of polymodal sensory stimuli by postsynaptic glutamate receptors in *C. elegans*. Neuron 36, 933–944.

Mertens, I., Clinckspoor, I., Janssen, T., Nachman, R., Schoofs, L., 2006. FMRFamide related peptide ligands activate the *Caenorhabditis elegans* orphan GPCR Y59H11AL.1. Peptides 27, 1291–1296.

Mertens, I., Meeusen, T., Janssen, T., Nachman, R., Schoofs, L., 2005. Molecular characterization of two G protein-coupled receptor splice variants as FLP2 receptors in *Caenorhabditis elegans*. Biochem. Biophys. Res. Commun. 330, 967–974.

Mertens, I., Vandingenen, A., Meeusen, T., De Loof, A., Schoofs, L., 2004a. Postgenomic characterization of G-protein-coupled receptors. Pharmacogenomics 5, 657–672.

Mertens, I., Vandingenen, A., Meeusen, T., Janssen, T., Luyten, W., Nachman, R.J., De Loof, A., Schoofs, L., 2004b. Functional characterization of the putative orphan neuropeptide G-protein coupled receptor C26F1.6 in *Caenorhabditis elegans*. FEBS Lett. 573, 55–60.

Missale, C., Nash, S.R., Robinson, S.W., Jaber, M., Caron, M.G., 1998. Dopamine receptors: from structure to function. Physiol. Rev. 78, 189–225.

Moffett, C.L., Beckett, A.M., Mousley, A., Geary, T.G., Marks, N.J., Halton, D.W., Thompson, D.P., Maule, A.G., 2003. The ovijector of *Ascaris suum*: multiple response types revealed by *Caenorhabditis elegans* FMRFamide-related peptides. Int. J. Parasitol. 33, 859–876.

Morton, D.B., Hudson, M.L., Waters, E., O'Shea, M., 1999. Soluble guanylyl cyclases in *Caenorhabditis elegans*: NO is not the answer. Curr. Biol. 9, R546–R547.

Mousley, A., Maule, A.G., Halton, D.W., Marks, N.J., 2005. Inter-phyla studies on neuropeptides: the potential for broad-spectrum anthelmintic and/or endectocide discovery. Parasitology 131, S143–S167.

Muller, L., Zhu, P., Juliano, M.A., Juliano, L., Lindberg, I., 1999. A 36-residue peptide contains all of the information required for 7B2-mediated activation of prohormone convertase 2. J. Biol. Chem. 274, 21471–21477.

Nathoo, A.N., Moeller, R.A., Westlund, B.A., Hart, A.C., 2001. Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species. Proc. Natl. Acad. Sci. U.S.A. 98, 14000–14005.

Nelson, L.S., Rosoff, M.L., Li, C., 1998. Disruption of a neuropeptide gene, *flp-1*, causes multiple behavioral defects in *Caenorhabditis elegans*. Science 281, 1686–1690.

O'Halloran, D.M., Fitzpatrick, D.A., McCormack, G.P., McInerney, J.O., Burnell, A.M., 2006. The molecular phylogeny of a nematode-specific clade of heterotrimeric G-protein alpha-subunit genes. J. Mol. Evol. 63, 87–94.

Olde, B., McCombie, W.R., 1997. Molecular cloning and functional expression of a serotonin receptor from *Caenorhabditis elegans*. J. Mol. Neurosci. 8, 53–62.

Orchard, I., Lange, A.B., Bendena, W.G., 2001. FMRFamide-related peptides: a multifunctional family of structurally related neuropeptides in insects. Adv. Insect Physiol. 28, 268–329.

Papaioannou, S., Marsden, D., Franks, C.J., Walker, R.J., Holden-Dye, L., 2005. Role of a FMRFamide-like family of neuropeptides in the pharyngeal nervous system of *Caenorhabditis elegans*. J. Neurobiol. 65, 304–319.

Park, Y.S., Kim, S., Shin, Y., Choi, B., Cho, N.J., 2003. Alternative splicing of the muscarinic acetylcholine receptor GAR-3 in *Caenorhabditis elegans*. Biochem. Biophys. Res. Commun. 308, 961–965.

Perry, S.J., Yi-Kung, H.E., Cronk, D., Bagust, J., Sharma, R., Walker, R.J., Wilson, S., Burke, J.F., 1997. A human gene encoding morphine modulating peptides related to NPFF and FMRFamide. FEBS Lett. 409, 426–430.

Pierce, S.B., Costa, M., Wisotzkey, R., Devadhar, S., Homburger, S.A., Buchman, A.R., Ferguson, K.C., Heller, J., Platt, D.M., Pasquinelli, A.A., Liu, L.X., Doberstein, S.K., Ruvkun, G., 2001. Regulation of DAF-2 receptor signaling by human insulin and *ins-1*, a member of the unusually large and diverse *C. elegans* insulin gene family. Genes Dev. 15, 672–686.

Price, D.A., Greenberg, M.J., 1977a. Purification and characterization of a cardioexcitatory neuropeptide from the central ganglia of a bivalve mollusc. Prep. Biochem. 7, 261–268.

Price, D.A., Greenberg, M.J., 1977b. Structure of a molluscan cardioexcitatory neuropeptide. Science 197, 670–671.

Prigge, S.T., Mains, R.E., Eipper, B.A., Amzel, L.M., 2000. New insights into copper monooxygenases and peptide amidation: structure, mechanism and function. Cell Mol. Life Sci. 57, 1236–1259.

Ranganathan, R., Cannon, S.C., Horvitz, H.R., 2000. MOD-1 is a serotonin-gated chloride channel that modulates locomotory behaviour in *C. elegans*. Nature 408, 470–475.

Reinitz, C.A., Herfel, H.G., Messinger, L.A., Stretton, A.O., 2000. Changes in locomotory behavior and cAMP produced in *Ascaris suum* by neuropeptides from *Ascaris suum* or *Caenorhabditis elegans*. Mol. Biochem. Parasitol. 111, 185–197.

Rex, E., Hapiak, V., Hobson, R., Smith, K., Xiao, H., Komuniecki, R., 2005. TYRA-2 (F01E11.5): a *Caenorhabditis elegans* tyramine receptor expressed in the MC and NSM pharyngeal neurons. J. Neurochem. 94, 181–191.

Rex, E., Komuniecki, R.W., 2002. Characterization of a tyramine receptor from *Caenorhabditis elegans*. J. Neurochem. 82, 1352–1359.

Rex, E., Molitor, S.C., Hapiak, V., Xiao, H., Henderson, M., Komuniecki, R., 2004. Tyramine receptor (SER-2) isoforms are involved in the regulation of pharyngeal pumping and foraging behavior in *Caenorhabditis elegans*. J. Neurochem. 91, 1104–1115.

Richmond, J.E., Broadie, K.S., 2002. The synaptic vesicle cycle: exocytosis and endocytosis in *Drosophila* and *C. elegans*. Curr. Opin. Neurobiol. 12, 499–507.

Riddle, D.L., Albert, P.S., 1997. Genetic and environmental regulation of dauer larva development. In: Riddle, D.L., Blumenthal, T., Meyer, B.J., Preiss, J.R. (Eds.), *C. elegans*, vol. II. Cold Spring Harbor Laboratory Press, New York, pp. 739–768.

Riehle, M.A., Garczynski, S.F., Crim, J.W., Hill, C.A., Brown, M.R., 2002. Neuropeptides and peptide hormones in *Anopheles gambiae*. Science 298, 172–175.

Rogers, C., Reale, V., Kim, K., Chatwin, H., Li, C., Evans, P., de Bono, M., 2003. Inhibition of *Caenorhabditis elegans* social feeding by FMRFamide-related peptide activation of NPR-1. Nat. Neurosci. 6, 1178–1185.

Rogers, C.M., Franks, C.J., Walker, R.J., Burke, J.F., Holden-Dye, L., 2001. Regulation of the pharynx of *Caenorhabditis elegans* by 5-HT, octopamine, and FMRFamide-like neuropeptides. J. Neurobiol. 49, 235–244.

Rosoff, M.L., Burglin, T.R., Li, C., 1992. Alternatively spliced transcripts of the *flp-1* gene encode distinct FMRFamide-like peptides in *Caenorhabditis elegans*. J. Neurosci. 12, 2356–2361.

Rosoff, M.L., Doble, K.E., Price, D.A., Li, C., 1993. The *flp-1* propeptide is processed into multiple, highly similar FMRFamide-like peptides in *Caenorhabditis elegans*. Peptides 14, 331–338.

Rouille, Y., Duguay, S.J., Lund, K., Furuta, M., Gong, Q., Lipkind, G., Oliva Jr., A.A., Chan, S.J., Steiner, D.F., 1995. Proteolytic processing mechanisms in the biosynthesis of neuroendocrine peptides: the subtilisin-like proprotein convertases. Front Neuroendocrinol. 16, 322–361.

Rutter, G.A., Tsuboi, T., 2004. Kiss and run exocytosis of dense core secretory vesicles. Neuroreport 15, 79–81.

Sajid, M., Keating, C., Holden-Dye, L., Harrow, I.D., Isaac, R.E., 1996. Metabolism of AF1 (KNEFIRFamide) in the nematode, *Ascaris suum*, by aminopeptidase, endopeptidase and deamidase enzymes. Mol. Biochem. Parasitol. 75, 159–168.

Sanyal, S., Wintle, R.F., Kindt, K.S., Nuttley, W.M., Arvan, R., Fitzmaurice, P., Bigras, E., Merz, D.C., Hebert, T.E., van der, K.D., Schafer, W.R., Culotti, J.G., Van Tol, H.H., 2004. Dopamine modulates the plasticity of mechanosensory responses in *Caenorhabditis elegans*. EMBO J. 23, 473–482.

Scalettar, B.A., 2006. How neurosecretory vesicles release their cargo. Neuroscientist 12, 164–176.

Schinkmann, K., Li, C., 1992. Localization of FMRFamide-like peptides in *Caenorhabditis elegans*. J. Comp. Neurol. 316, 251–260.

Schoofs, L., Holman, G.M., Hayes, T.K., Nachman, R.J., De Loof, A., 1991. Isolation, identification and synthesis of locustamyoinhibiting peptide (LOM-MIP), a novel biologically active neuropeptide from *Locusta migratoria*. Regul. Pept. 36, 111–119.

Schulz-Knappe, P., Zucht, H.D., Heine, G., Jurgens, M., Hess, R., Schrader, M., 2001. Peptidomics: the comprehensive analysis of peptides in complex biological mixtures. Comb. Chem. High Throughput. Screen. 4, 207–217.

Shaw, C., Maule, A.G., Halton, D.W., 1996. Platyhelminth FMRFamide-related peptides. Int. J. Parasitol. 26, 335–345.

Sieburth, D., Ch'ng, Q., Dybbs, M., Tavazoie, M., Kennedy, S., Wang, D., Dupuy, D., Rual, J.F., Hill, D.E., Vidal, M., Ruvkun, G., Kaplan, J.M., 2005. Systematic analysis of genes required for synapse structure and function. Nature 436, 510–517.

Simmer, F., Tijsterman, M., Parrish, S., Koushika, S.P., Nonet, M.L., Fire, A., Ahringer, J., Plasterk, R.H., 2002. Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi. Curr. Biol. 12, 1317–1319.

Smit, A.B., van Kesteren, R.E., Li, K.W., Van, M.J., Spijker, S., Van, H.H., Geraerts, W.P., 1998. Towards understanding the role of insulin in the brain: lessons from insulin-related signaling systems in the invertebrate brain. Prog. Neurobiol. 54, 35–54.

Southey, B.R., Amare, A., Zimmerman, T.A., Rodriguez-Zas, S.L., Sweedler, J.V., 2006. NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides. Nucleic Acids Res. 34, W267–W272.

Spanier, B., Sturzenbaum, S.R., Holden-Dye, L.M., Baumeister, R., 2005. *Caenorhabditis elegans* neprilysin NEP-1: an effector of locomotion and pharyngeal pumping. J. Mol. Biol. 352, 429–437.

Sudhof, T.C., 2005. CAPS in search of a lost function. Neuron 46, 2–4.

Sugiura, M., Fuke, S., Suo, S., Sasagawa, N., Van Tol, H.H., Ishiura, S., 2005. Characterization of a novel D2-like dopamine receptor with a truncated

splice variant and a D1-like dopamine receptor unique to invertebrates from *Caenorhabditis elegans*. J. Neurochem. 94, 1146–1157.

Suo, S., Sasagawa, N., Ishiura, S., 2002. Identification of a dopamine receptor from *Caenorhabditis elegans*. Neurosci. Lett. 319, 13–16.

Suo, S., Sasagawa, N., Ishiura, S., 2003. Cloning and characterization of a *Caenorhabditis elegans* D2-like dopamine receptor. J. Neurochem. 86, 869–878.

Tatar, M., Bartke, A., Antebi, A., 2003. The endocrine regulation of aging by insulin-like signals. Science 299, 1346–1351.

Tatar, M., Kopelman, A., Epstein, D., Tu, M.P., Yin, C.M., Garofalo, R.S., 2001. A mutant *Drosophila* insulin receptor homolog that extends life-span and impairs neuroendocrine function. Science 292, 107–110.

Thacker, C., Marra, M.A., Jones, A., Baillie, D.L., Rose, A.M., 1999. Functional genomics in *Caenorhabditis elegans*: an approach involving comparisons of sequences from related nematodes. Genome Res. 9, 348–359.

Thacker, C., Peters, K., Srayko, M., Rose, A.M., 1995. The *bli-4* locus of *Caenorhabditis elegans* encodes structurally distinct kex2/subtilisin-like endoproteases essential for early development and adult morphology. Genes Dev. 9, 956–971.

Thacker, C., Rose, A.M., 2000. A look at the *Caenorhabditis elegans* Kex2/Subtilisin-like proprotein convertase family. Bioessays 22, 545–553.

The *C. elegans* Sequencing Consortium, 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science 282, 2012–2018.

Thomas, J.H., 1990. Genetic analysis of defecation in *Caenorhabditis elegans*. Genetics 124, 855–872.

Thompson, D.P., Davis, J.P., Larsen, M.J., Coscarelli, E.M., Zinser, E.W., Bowman, J.W., Alexander-Bowman, S.J., Marks, N.J., Geary, T.G., 2003. Effects of KHEYLRFamide and KNEFIRFamide on cyclic adenosine monophosphate levels in *Ascaris suum* somatic muscle. Int. J. Parasitol. 33, 199–208.

Trent, C., Tsuing, N., Horvitz, H.R., 1983. Egg-laying defective mutants of the nematode *Caenorhabditis elegans*. Genetics 104, 619–647.

Turner, A.J., Isaac, R.E., Coates, D., 2001. The neprilysin (NEP) family of zinc metalloendopeptidases: genomics and function. Bioessays 23, 261–269.

Vanden Broeck, J., 2001. Neuropeptides and their precursors in the fruitfly, *Drosophila melanogaster.* Peptides 22, 241–254.

Veenstra, J.A., 2000. Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. Arch. Insect Biochem. Physiol. 43, 49–63.

Waggoner, L.E., Hardaker, L.A., Golik, S., Schafer, W.R., 2000. Effect of a neuropeptide gene on behavioral states in *Caenorhabditis elegans* egg-laying. Genetics 154, 1181–1192.

Wegener, C., Herbert, Z., Eckert, M., Predel, R., 2002. The periviscerokinin (PVK) peptide family in insects: evidence for the inclusion of CAP2B as a PVK family member. Peptides 23, 605–611.

Weimer, R.M., Jorgensen, E.M., 2003. Controversies in synaptic vesicle exocytosis. J. Cell Sci. 116, 3661–3666.

Westphal, C.H., Muller, L., Zhou, A., Zhu, X., Bonner-Weir, S., Schambelan, M., Steiner, D.F., Lindberg, I., Leder, P., 1999. The neuroendocrine protein 7B2 is required for peptide hormone processing in vivo and provides a novel mechanism for pituitary Cushing's disease. Cell 96, 689–700.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., 1994. 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. Nature 368, 32–38.

Wu, Q., Brown, M.R., 2006. Signaling and function of insulin-like peptides in insects. Annu. Rev. Entomol. 51, 1–24.

Yang, H.Y., Fratta, W., Majane, E.A., Costa, E., 1985. Isolation, sequencing, synthesis, and pharmacological characterization of two brain neuropeptides that modulate the action of morphine. Proc. Natl. Acad. Sci. U.S.A. 82, 7757–7761.

Yew, J.Y., Dikler, S., Stretton, A.O., 2003. De novo sequencing of novel neuropeptides directly from Ascaris suum tissue using matrix-assisted laser desorption/ionization time-of-flight/time-of-flight. Rapid Commun. Mass Spectrom. 17, 2693–2698.

Yew, J.Y., Kutz, K.K., Dikler, S., Messinger, L., Li, L., Stretton, A.O., 2005. Mass spectrometric map of neuropeptide expression in *Ascaris suum*. J. Comp. Neurol. 488, 396–413.

Yu, S., Avery, L., Baude, E., Garbers, D.L., 1997. Guanylyl cyclase expression in specific sensory neurons: a new family of chemosensory receptors. Proc. Natl. Acad. Sci. U.S.A. 94, 3384–3387.

Zahn, T.R., Angleson, J.K., MacMorris, M.A., Domke, E., Hutton, J.F., Schwartz, C., Hutton, J.C., 2004. Dense core vesicle dynamics in *Caenorhabditis elegans* neurons and the role of kinesin UNC-104. Traffic 5, 544–559.