Functional characterization of the putative orphan neuropeptide G-protein coupled receptor C26F1.6 in *Caenorhabditis elegans*☆

Inge Mertens[a], Anick Vandingenen[a], Tom Meeusen[a], Tom Janssen[a], Walter Luyten[b], Ronald J. Nachman[c], Arnold De Loof[a], Liliane Schoofs[a,*]

[a] *Laboratory of Developmental Physiology, Genomics and Proteomics, Katholieke Universiteit Leuven, Naamsestraat 59, B-3000 Leuven, Belgium*
[b] *IriDM, Naamsestraat 59, B-3000 Leuven, Belgium*
[c] *Southern Plains Agricultural Research Center, U.S. Department of Agriculture – Agricultural Research Service, 2881 F&B Road, College Station, TX 77845, USA*

Received 28 May 2004; revised 7 July 2004; accepted 14 July 2004

Available online 30 July 2004

Edited by Irmgard Sinning

Abstract In this study, we describe the cloning and the characterization of the third FMRFamide-related peptide (FaRP) receptor in *Caenorhabditis elegans*, the *VRFa receptor 1*. Numerous structurally different FaRPs were synthesized and used to screen the orphan C26F1.6 receptor for activation. Two peptides ending in M(orL)VRFamide elicited a calcium response in receptor expressing mammalian cells. The response is dose-dependent and appeared to be very specific, since very closely related FaRPs were less active, even the other peptides ending in M(orL)VRFamide. Pharmacological profiling of the most active peptide suggests that SMVRFa is the most active binding core. N-terminal extension decreases peptide activity.

Keywords: *Caenorhabditis elegans*; VRF-amide receptor 1; Reverse pharmacology; FMRFamide related peptide

1. Introduction

For a variety of reasons, the nematode worm *Caenorhabditis elegans* has been developed into a prime model system in animal biology. In 2002, the Nobel Prize for Medicine/Physiology was awarded to the pioneers in this domain, namely Brenner, Horvitz and Sulston.

The *C. elegans* genome (97 Mb, 19.099 genes in total) codes for about 1000 G-protein coupled receptors (GPCRs) [1]. About 100 of these encode neuropeptide GPCRs [2].

The best known neuropeptide family in invertebrates is the family of FMRFamide-related peptides (FaRPs). In *C. elegans*, a lot of FaRPs and their corresponding receptors remain uncharacterized. However, current problems of drug resistance in nematode parasites necessitate the identification of new targets and their exploitation through novel drug design.

In this study, we describe the characterization of the third FaRP receptor in *C. elegans*, receptor WP:CE06880 (gene C26F1.6). This receptor is structurally most closely related to other orphan *C. elegans* GPCRs from the rhodopsin family and the recently characterized FaRP receptors in *Drosophila melanogaster*, the FMRFamide receptor [3] and short NPF receptor [4].

We developed a cell line (HEK293 cells) expressing the C26F1.6 orphan receptor and screened several synthetic FaRPs, mined from the *C. elegans* genome, for a calcium response in the receptor expressing cells using a cellular fluorescent assay. We show that two peptides of *C. elegans* ending in M(L)VRFamide activate the cloned receptor in a specific and dose-dependent way.

2. Materials and methods

2.1. Cloning of the VRFamide receptor 1

The ORF of the C26F1.6 gene was amplified by PCR performed on the cDNA (SuperScript First-Strand Synthesis System for RT-PCR, Invitrogen, The Netherlands), synthetized from mRNA (QuickPrep *micro* mRNA Purification Kit, Roche, IN, USA) of whole nematodes. Specific oligonucleotide PCR primers (Eurogentec, Belgium) were used (forward primer 5′-CAG GAT CCG CCA CCA TGT TGC TAC TAC TCT CCC GGG-3′; reverse primer 5′-CAT CTA GAG ACA AAT TCC TAG ATA TCA CAT GGC-3′). The Advantage 2 PCR kit (Clontech, USA) was used under the following PCR conditions: 94 °C for 60 s and 68 °C for 60 s, 68 °C for 180 s (30 cycles). The obtained PCR product was first cloned in the pCRII-TOPO vector using the TOPO-TA Cloning kit (Invitrogen, The Netherlands). After selection and sequencing (310 Genetic Analyzer, Applied Biosystems, UK), the receptor was directionally cloned into the pcDNA3 mammalian expression vector (Invitrogen).

2.2. Creation of a cell line expressing C. elegans VRFa receptor 1

Human embryonic kidney (HEK293) cells were cultured at 37 °C in a humified atmosphere of 5% CO_2 in DMEM (BioWhittaker, Belgium), supplemented with 10% heat inactivated horse serum, non-essential amino acids, 100 U/ml penicillin and 100 μg/ml streptomycin. The C26F1.6/pcDNA3 construct was transfected into the HEK293 cells using FuGene 6 (Roche, IN, USA) according to the manufacturer's instructions. The promiscuous G-protein $G_{\alpha 16}$ was cotransfected.

☆ Nucleotide sequence data reported are available in the GenBank database under the Accession No. BK004131.

* Corresponding author. Fax: +32-16-32-39-02.
E-mail address: liliane.schoofs@bio.kuleuven.ac.be (L. Schoofs).

Abbreviations: FaRPs, FMRFamide-related peptides; *flp*, FMRFamide-like protein gene; GPCR, G-protein coupled receptor; CHO, Chinese hamster ovary; G16, G-protein 16; Drm-sNPF-R, *Drosophila melanogaster* short Neuropeptide F receptor

0014-5793/$22.00
doi:10.1016/j.febslet.2004.07.058

2.3. Peptide synthesis

All peptides were either custom-synthesized (Invitrogen) or made in house using conventional Fmoc chemistry.

2.4. Fluorescence assay

The HEK293 cells were detached 24 h after transfection. Cells were subsequently plated out in 96-well plates at approximately 70% confluence. Forty-eight hours after transfection, the cells were loaded with fluorophore, Fluo-4-AM (Molecular probes, The Netherlands) for 1 h, after which excess fluorophore was washed away with HBSS* buffer (HBSS, supplemented with 5 mM $CaCl_2$ and 10 mM HEPES). Excitation of the fluorophore was done at 488 nm. Fifty microliters of the different concentrations of the synthetic peptides, diluted in HBSS* buffer, was pipetted out of the compound plate and added to the 96-well plate containing the HEK-cells. The calcium response was measured for 2 min at 525 nm using the FLEXStation (Molecular Devices). Data were analyzed using Softmax Pro (Molecular Devices).

3. Results

3.1. Cloning of the full-length cDNA of C. elegans VRFa-R1

PCR amplification of the cDNA with oligonucleotide primers specific for the predicted ORF of C26F1.6 produced a single product of approximately 1100 bp (data not shown). Sequence determination of the TA-cloned PCR-product revealed a DNA insert of 1083 bp, corresponding to the sequence of the five predicted exons of the C26F1.6 gene (www.wormbase.org). The deduced protein encoded by the ORF of the C26F1.6 gene is 360 amino acids long. Analysis by the TMHMM program (www.cbs.dtu.dk/services/TMHMM-2.0/) revealed that this protein is predicted to have seven transmembrane domains along with the intracellular and extracellular loops, consistent with the known structure of GPCRs. The N-terminal extracellular region exhibits no O-glycosylation (www.cbs.dtu.dk/services/NetOGlyc/) and no N-glycosylation sites (www.cbs.dtu.dk/services/NetNGlyc/). The intracellular C-terminal region exhibits 10 possible phosphorylation sites (www.cbs.dtu.dk/services/NetPhos/).

A phylogenetic tree based on Clustal W alignment of the receptor sequence and various other related orphan *C. elegans* GPCRs as well as the FMRFamide receptor and the sNPF receptor of *D. melanogaster*, indicates that the receptor is most closely related to *C. elegans* orphan receptors CE29348 (T19F4.1) and CE06168 (K10C8.2) and to the Drm-FMRFa (CG2114) and Drm-sNPF receptors (CG7395) (Fig. 1). Sequence similarity among these receptors is depicted in their alignment by the AlignX program (Informax) (Fig. 2).

3.2. Fluorescence assay

Following successful cloning of the receptor and the construction of a cell line expressing the C26F1.6 gene, the cells were challenged with different synthetic FaRPs from *D. melanogaster* (sNPFs and FMRFamides), and from other arthropods as well as predicted peptides mined by BLAST analysis from the genome of *C. elegans* (Table 1). Only two peptides were able to clearly activate the receptor in a dose-dependent way: TPMQRSSMVRFamide encoded by the *flp 7* precursor gene and AMRNALVRFamide encoded by the *flp 11* precursor gene. Both peptides have the M(orL)VRFamide carboxyterminal sequence in common. Fig. 3 shows the fluorescence response of TPMQRSSMVRFamide, the most potent peptide ($EC_{50} = 1.02 \pm 0.24$ μM) and the response of AMRNALVRFamide, the second active peptide ($EC_{50} = 1.34 \pm 0.41$ μM).

The receptor response appears to be rather specific because the EC_{50} value of SPMQRSSMVRFamide, which differs only by a single N-terminal amino acid from the most potent peptide, is higher than 1.598 μM because at a concentration of 10 μM, a plateau is still not yet reached (Fig. 3). The EC_{50} value of SPMERSAMVRFamide, another closely related peptide, is higher than 1.732 μM (Fig. 3) and SPMDRKMVRFamide was no longer capable to activate the receptor at all (Fig. 3). ASGGMRNALVRFamide has the same C-terminus as the active AMRNALVRFamide peptide, but, in contrast, is not capable to elicit a response (Fig. 3). The receptor encoded by the C26F1.6 gene was named the first *C. elegans* VRFamide receptor or *C. elegans* VRFa-R1.

To search for the minimum core sequence for receptor activation, we tested an N-terminally truncated series of synthetic analogs of the most active peptide (TPMQRSSMVRFamide). Deletion of the N-terminal threonine yields a peptide with an EC_{50} value of 0.652 ± 0.083 μM, which is therefore more potent than the full-length peptide, but is still less potent than the carboxyterminal 5-mer. Further truncation of the proline yields an inactive peptide. The C-terminal SMVRFamide sequence seems to be the minimum core peptide sequence required for full activation of the receptor. This 5-mer is far more potent than the full-length peptide sequence (EC_{50} value $= 0.096 \pm 0.016$ μM). Without the interference of the N-terminal residues, the carboxyterminal SMVRFamide (and SSMVRFa) itself is apparently able to interact with the binding pocket region and to elicit a maximum receptor response. For activation of the receptor by the entire peptide sequence as predicted in the genome, the proline residue is needed to have the right conformational properties to bind the receptor.

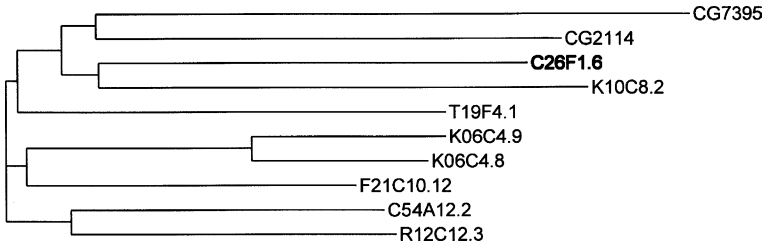

Fig. 1. Phylogenetic relationship of the *C. elegans* VRFamide receptor 1 (C26F1.6) and related peptide receptors (AlignX software-Informax). CG7395: Drm-sNPF-R or *Drosophila* short NPF receptor; CG2114: Drm-FMRFa-R or *Drosophila* FMRFamide receptor; K10C8.2, T19F4.1, K06C4.9, K06C4.8, F21C10.12, C54A12.2, R12C12.3: predicted orphan neuropeptide GPCRs in *C. elegans*.

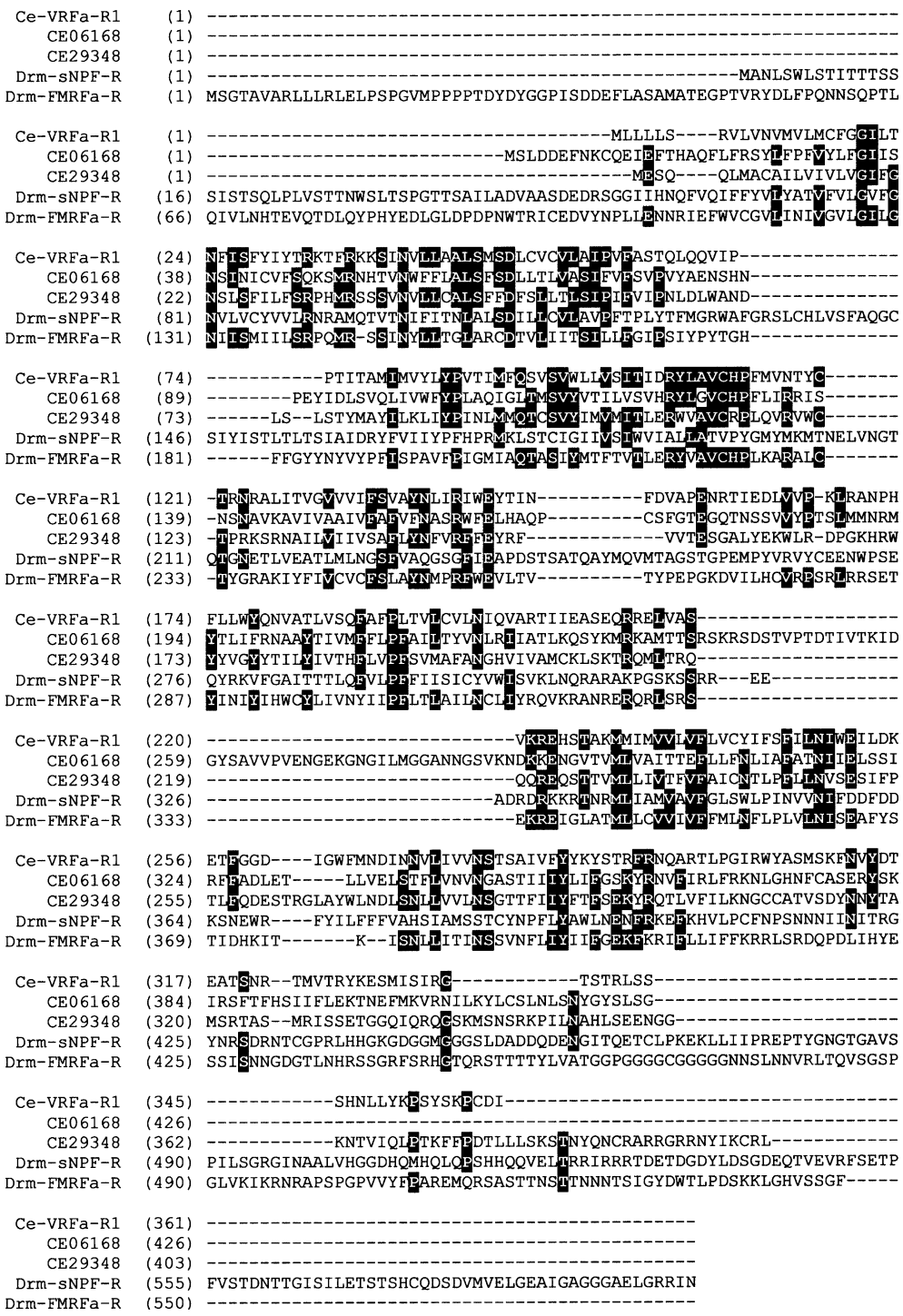

Fig. 2. Amino acid sequence alignments. Alignment of amino acid sequences of the *C. elegans* VRFamide receptor 1, *C. elegans* orphan GPCRs CE29348 and CE06168, the short Neuropeptide F receptor and the FMRFamide receptor of *Drosophila melanogaster* by AlignX software (Informax). Identical amino acids are indicated by the black background.

4. Discussion

BLAST searches of the *C. elegans* database reveal over 1000 GPCRs. Fifty-four of these receptors display similarities to vertebrate neuropeptide receptors and may include FaRP receptors [5]. To date, only two neuropeptide receptors have been identified in *C. elegans*, namely the FLP15 peptide receptor and the AF9 receptor 1 [6,7].

We decided to search for the ligand and the function of the C26F1.6 encoded receptor because this receptor shows very high similarity to the recently characterized *Drosophila* short neuropeptide F (sNPF) receptor [4] and the FMRFamide

Table 1
Peptide library of synthetic peptides, amino acid sequences and activity as tested in various concentrations (10^{-4}–10^{-11} M) for a calcium response using the fluorescence assay

Peptide	Sequence	EC_{50} (μM)
C. elegans peptides		
flp-1	KPNFMRYamide	n.a.
	PNFLRFamide	n.a.
flp-2	SPREPIRFamide	n.a.
flp-3	SPLGTMRFamide	n.a.
flp-4	PTFIRFamide	n.a.
flp-6	KSAYMRFamide	n.a.
flp-7	TPMQRSS**MVRFamide**	1.015 ± 0.240
	SPMQRSS**MVRFamide**	>1.598
	SPMERSA**MVRFamide**	>1.732
	SPMDRSK**MVRFamide**	n.a.
flp-10	QPKARSGYIRFamide	n.a.
flp-11	AMRNA**LVRFamide**	1.336 ± 0.414
	ASGGMRNA**LVRFamide**	n.a.
	NGAPQPFVRFamide	n.a.
flp-13	ASSAPLIRFamide	n.a.
flp-14	KHEYLRFamide	n.a.
flp-15	RGPSGPLRFamide	n.a.
flp-16	AQTFVRFamide	n.a.
flp-18	DVPGVLRFamide	n.a.
CE18432	EIVFHQISPIFFRFamide	n.a.
	SLLDYRFamide	n.a.
D. melanogaster peptides		
sNPF-1	AQRSPSLRLRFamide	n.a.
sNPF-2	SPSLRLRFamide	n.a.
sNPF-3	PQRLRWamide	n.a.
sNPF-4	PMRLRWamide	n.a.
FMRFa-1	DPKQDFMRFamide	n.a.
FMRFa-2	TPAEDTMRFamide	n.a.
FMRFa-3	SDNFMRFamide	n.a.
FMRFa-4	SPKQDFMRFamide	n.a.
FMRFa-5	PDNFMRFamide	n.a.
Other related peptides		
Pev SK	AGGSGGVGGE**Y**DD**Y**GHLRFamide	n.a.
Scg-FLRFa	PDVDHVFLRFamide	n.a.
Trunctated series of most active peptide		
PMQRSS**MVRFamide**		0.652 ± 0.083
MQRSS**MVRFamide**		n.a.
QRSS**MVRFamide**		n.a.
RSS**MVRFamide**		n.a.
SS**MVRFamide**		1.101 ± 0.212
S**MVRFamide**		0.096 ± 0.016
MVRFamide		n.a.

n.a., not active up to 10 μM. The Y residues of the sulfakinin indicated in bold are sulfated. sNPF, short Neuropeptide F. Pev SK, *Penaeus vennamei* sulfakinin. Scg-FLRFa, *Schistocerca gregaria* FLRFamide.
EC_{50} values were calculated using the results of three independent measurements. When the plateau was not reached at 10 μM, EC_{50} values are higher than the calculated value.
The receptor expressing cells were incubated at 37 °C post-transfection.

receptor [3] that are supposed to be involved in reproduction. sNPFs are members of the FaRP family. Since there are no peptides with the typical sNPF structure present in the *C. elegans* genome, we screened the receptor with all the different types of *C. elegans* FaRPs known to date.

The receptor specificity seems to be determined by the conformation of the entire peptide sequence, as indicated by the activity profile of closely related peptides. The presence of the N-terminal residues MQR of the truncated analogs appears to interfere with the ability of the C-terminal region to adapt to the right conformation to bind the receptor. So we can conclude that N-terminal extension of the C-terminal 5-mer (SMVRFamide) decreases receptor activation. It is hard to tell whether either (S)SMVRFamide or TPMQRSSMVRFamide is the endogenous ligand of the VRFamide receptor, as long as the endogenous peptide has not been purified. Bioinformatic analysis predicts the occurrence of TPMQRSSMVRFamide although it may be possible that this peptide is further processed to yield the shorter and fully potent SMVRFamide.

Flp-genes encode a family of peptides ending in RFamide, which belong to the FMRFamide-related peptide family (FaRPs). The first FaRP, the genuine FMRFamide, was isolated in 1977 as a cardioactive agent on the molluscan heart [8]. Since then, FARPs have been found in the nervous system of animals representing all major phyla [9–12]. These peptides have been shown to have diverse functions in invertebrates

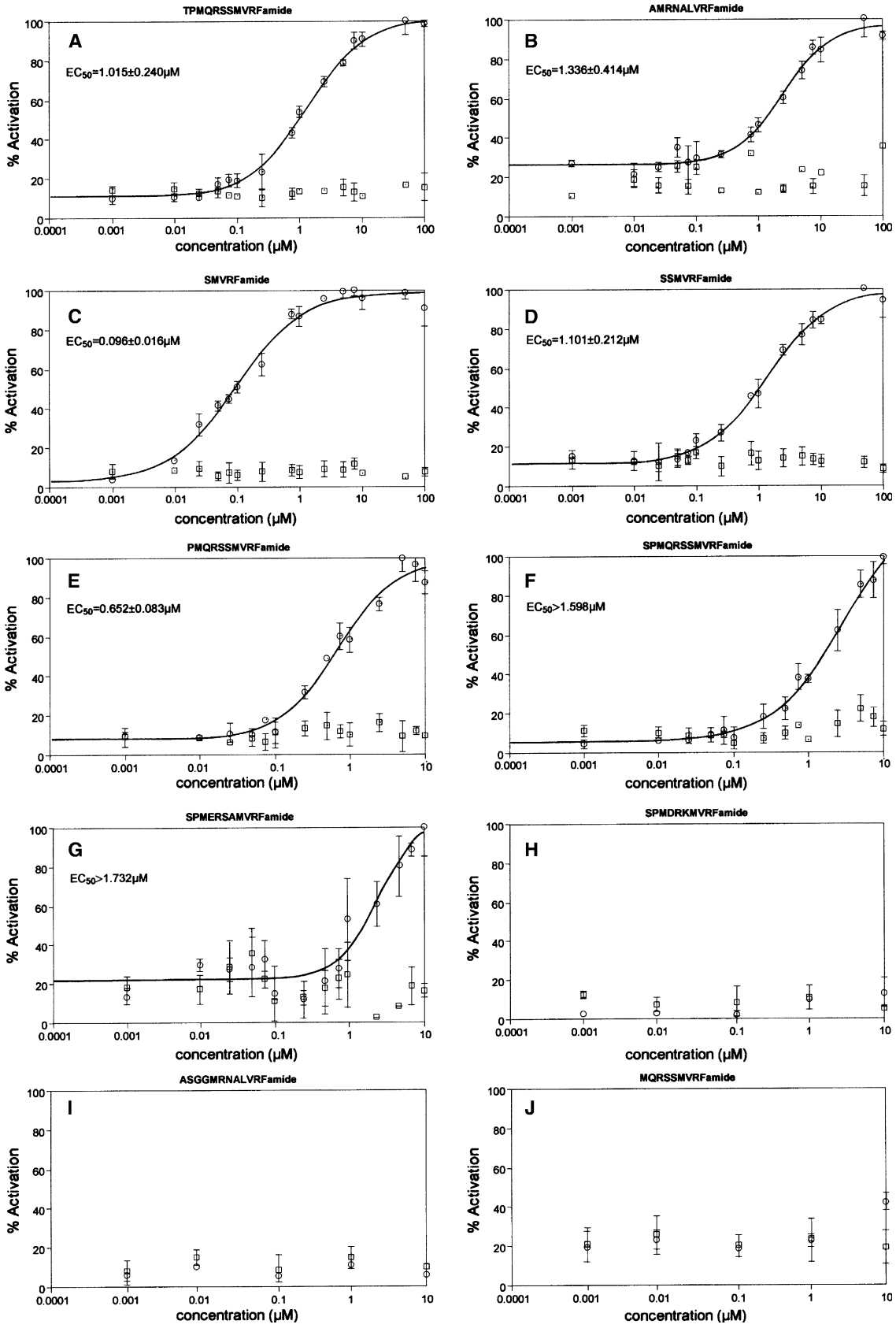

Fig. 3. Dose–response curves of the most important peptides tested on the VRFamide receptor 1 expressed in HEK293 cells. Fluorescent responses of the cloned cell line expressing the VRFa-R1 are expressed in % activation. Receptor responses (pcDNA/receptor) are represented by ○ and negative control responses (pcDNA3) are represented by □. These are the collected data of three independent measurements. The vertical bars represent standard deviations. Data were processed using Softmax Pro software (Molecular Devices).

such as cardioexcitation [13], control of the muscle contraction [14,15] and neuromodulation [16]. In vertebrates, they have anti-opioid effects [17].

The 23 *C. elegans flp*-genes, designated *flp-1* through *flp-23*, which encode 59 distinct FMRFamide-related neuropeptides, were identified by computional methods, i.e. through conventional screens of the cDNA libraries by GENEFINDER predictions from the *C. elegans* Genome Consortium and BLAST screens of the *C. elegans* genome [1]. *C. elegans* FaRPs are expressed in at least 10% of the neurons, including motor, sensory, and interneurons that are involved in movement, feeding, defecation and reproduction [6]. Hill et al. [13] made a transcriptional profiling of *C. elegans* in which the expression of all the genes is monitored by oligonucleotide arrays. They showed that there was no significant variation in the frequency in the transcript of genes F49E10.3 and K02G10.4 (*flp-7* and *flp-11* precursor) across eight developmental stages. The same results were found for the expression of the VRFamide receptor 1.

Out of all the *flp*-genes, only *flp-1* is extensively studied. The cognate ligands of the here described VRFamide receptor 1, TPMQRSSMVRFa and AMRNALVRFa, respectively, encoded by the *flp*-7 and *flp*-11 precursor, are expressed in *C. elegans*, since cDNAs were isolated for *flp-7* and EST sequences were found for *flp-11*. Also, alternative transcripts were isolated for *flp-11*. One *flp-11* transcript encodes three FaRPs, while the other transcript encodes only two of these FaRPs [5]. The functional consequences of these alternative transcripts are unknown. Animals in which the coding region of the *flp-7* and *flp-11* peptide precursor genes has been deleted, neither display an observable phenotype [18]. However, in nematodes in particular, it has been reported that FaRPergic neurons innervate essentially all muscular systems [19]. In *Ascaris suum*, AMRNALVRFamide induces a shortening of the ovijector, coupled with an increase in contraction frequency [17]. The neuropeptides encoded by *flp-7* produce paralysis and loss of waveforms, increased body length and decreased cAMP concentrations in *Ascaris* [20].

A genome-wide RNAi-study has recently been conducted in *C. elegans*. Disruption of *VRFamide receptor 1* gene (C26F1.6) revealed no visible phenotype considering viability, growth defects and defects in post-embryonic development [18].

However, when reproduction is considered, Keating and coworkers [21] very recently found that RNAi of the presently identified VRFamide receptor 1 on N2 animals resulted in a statistically significant increase in the number of progeny counted 48 h post-L4. The precise mechanism underlying the RNAi effects on egg laying and brood size were not determined in the latter study. One way of explaining the phenotype, considering the possible role of FaRPs, one of which is the ligand for the VRFa receptor 1, in muscle control, could be that the increase in progeny is due to physiological changes in the neurons or muscles that control egg laying. Future experiments will have to prove this assumption, since the observed phenotype could also be due to changes in the rate of ovulation, the number of sperm in the spermatheca or microchanges in the worm's anatomy.

It is worth noting that in this study, we described the characterization of the third *C. elegans* GPCR. A lot of orphan GPCRs await their characterization, which seems to be an arduous task, since the big publication boom is still to come.

Acknowledgements: This project was sponsored by the IWT (Vlaams instituut voor de bevordering van het wetenschappelijk-technologisch onderzoek in de industrie), the Flemish Science Foundation (FWO) (G.0356.98, G.0187.00 and G.0175.02) and the Belgian "Interuniversity Poles of Attraction Programme" (IUAP P5/30). We thank Prof. Dr. J. Vanfleteren for supplying the nematodes and finally L. Vanden Bosch and S. Van Soest for excellent technical assistance.

References

[1] Anon. Science 282 (1998) 2012–2018.
[2] Nathoo, A.N., Moeller, R.A., Westlund, B.A. and Hart, A.C. (2001) Proc. Natl. Acad. Sci. USA 98, 14000–14005.
[3] Meeusen, T., Mertens, I., Clynen, E., Baggerman, G., Nichols, R., Nachman, R.J., Huybrechts, R., De Loof, A. and Schoofs, L. (2002) Proc. Natl. Acad. Sci. USA 99, 15363–15368.
[4] Mertens, I., Meeusen, T., Huybrechts, R., De Loof, A. and Schoofs, L. (2002) Biochem. Biophys. Res. Commun. 297, 1140–1148.
[5] Li, C., Kim, C. and Nelson, L.S. (1999) Brain Res. 848, 26–34.
[6] Kubiak, T.M., Larsen, M.J., Nulf, S.C., Zantello, M.R., Burton, K.J., Bowman, J.W., Modric, T. and Lowery, D.E. (2003) J. Biol. Chem. 278, 33724–33729.
[7] Kubiak, T.M., Larsen, M.J., Zantello, M.R., Bowman, J.W., Nulf, S.C. and Lowery, D.E. (2003) J. Biol. Chem. 278, 42115–42120.
[8] Price, D.A. and Greenberg, M.J. (1977) Science 197, 670–671.
[9] Nichols, R. (1992) J. Mol. Neurosci. 3, 213–218.
[10] Rosoff, M.L., Doble, K.E., Price, D.A. and Li, C. (1993) Peptides 14, 331–338.
[11] Yang, H.Y.T., Fratta, W., Majane, E.A. and Costa, E. (1985) Proc. Natl. Acad. Sci. USA 82, 7757–7761.
[12] Orchard, I., Lange, A.B. and Bendena, W.G. (2001) Adv. Insect Physiol. 28, 267–329.
[13] Hill, A.A., Hunter, C.P., Tsung, B.T., Tucker-Kellogg, G. and Brown, E.L. (2000) Science 290, 809–812.
[14] Bowman, J.W., Friedman, A.R., Thompson, D.P., Ichhpurani, A.K., Kellman, M.F., Marks, N., Maule, A.G. and Geary, T.G. (1996) Peptides 17, 381–387.
[15] Marks, N.J., Maule, A.G., Halton, D.W., Geary, T.G., Shaw, C. and Thompson, D.P. (1997) Parasitology 114, 531–539.
[16] Brownlee, D.J.A., Holdendye, L., Fairweather, I. and Walker, R.J. (1995) Parasitology 111, 379–384.
[17] Moffett, C.L., Beckett, A.M., Mousley, A., Geary, T.G., Marks, N.J., Halton, D.W., Thompson, D.P. and Maule, A.G. (2003) Int. J. Parasitol. 33, 859–876.
[18] 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. and Ahringer, J. (2003) Nature 421, 231–237.
[19] 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. and Thompson, D.P. (1999) Ann. N. Y. Acad. Sci. 897, 212–217.
[20] Reinitz, C.A., Herfel, H.G., Messinger, L.A. and Stretton, A.O.W. (2000) Mol. Biochem. Parasitol. 111, 185–197.
[21] Keating, C.D., Kriek, N., Daniels, M., Ashcroft, N.R., Hopper, N.A., Siney, E.J., Holden-Dye, L. and Burke, J.F. (2003) Curr. Biol. 13, 1715–1720.