Available online at www.sciencedirect.com
SCIENCE DIRECT
Biochemical and Biophysical Research Communications 340 (2006) 696–701
BBRC
www.elsevier.com/locate/ybbrc

Molecular identification of the first SIFamide receptor ☆

Lars M. Jørgensen, Frank Hauser, Giuseppe Cazzamali, Michael Williamson, Cornelis J.P. Grimmelikhuijzen *

Center for Functional and Comparative Insect Genomics, Department of Cell Biology and Comparative Zoology, Institute of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark

Received 5 December 2005
Available online 20 December 2005

Abstract

SIFamide is the short name and also the C terminus of the *Drosophila* neuropeptide AYRKPPFNGSIFamide. SIFamide has been isolated or predicted from various insects and crustaceans, and appears to be extremely well conserved among these arthropods. However, the function of this neuropeptide is still enigmatic. Here, we have identified the *Drosophila* gene (CG10823) coding for the SIFamide receptor. When expressed in Chinese hamster ovary cells, the receptor is only activated by *Drosophila* SIFamide (EC_{50}, 2×10^{-8} M) and not by a library of 32 other insect neuropeptides and eight biogenic amines. Database searches revealed SIFamide receptor orthologues in the genomes from the malaria mosquito *Anopheles gambiae*, the silkworm *Bombyx mori*, the red flour beetle *Tribolium castaneum*, and the honey bee *Apis mellifera*. An alignment of the five insect SIFamide or SIFamide-like receptors showed, again, an impressive sequence conservation (67–77% amino acid sequence identities between the seven-transmembrane areas; 82–87% sequence similarities). The identification of well-conserved SIFamide receptor orthologues in all other insects with a sequenced genome, suggests that the SIFamide/receptor couple must have an essential function in arthropods. This paper is the first report on the identification of a SIFamide receptor.

Keywords: G protein-coupled receptor; GPCR; Neuropeptide; Neurohormone; Peptide hormone; Endocrinology; Insects; Crustaceans; *Drosophila*; Honey bee; Mosquito; Malaria

Neurohormones (neuropeptides, protein hormones, and biogenic amines) and their G protein-coupled receptors (GPCRs) occupy a high hierarchic position in the physiology of insects, because they steer important processes, such as reproduction, development, feeding, and behavior. The recently sequenced genome from *Drosophila melanogaster* contains 47 neuropeptide/protein hormone and 21 biogenic amine GPCRs [1]. About 60% of these GPCRs have been deorphanized, so far [1]. The identification of the remaining 40% of the GPCRs may be somewhat more difficult, but its results will be quite important, because it will give us a whole new view on *Drosophila* neuroendocrinology, for example by identifying all the target cells and organs of a certain neurohormone (using receptor localizations) or by getting additional insights into the neurohormone's actions (using receptor gene knock-outs).

Most of the neurohormone GPCR identifications in insects will be carried out in *Drosophila* and will be used to better understand *Drosophila*. These *Drosophila* GPCR data, however, are also essential to interpret (annotate) the neurohormone GPCR genes in the other insects with a sequenced genome, such as the malaria mosquito *Anopheles gambiae*, the silkworm *Bombyx mori*, the red flour beetle *Tribolium castaneum*, and the honey bee *Apis mellifera*. Using the information on deorphanized *Drosophila* neurohormone GPCRs, for example, we were able to identify the ligands or likely ligand for 63% of all honey bee neurohormone GPCRs (F. Hauser et al., unpublished).

☆ The sequence data from this study have been submitted to the GenBank database under Accession Nos. BK005264, BK005732–BK005734, and DQ285411.

* Corresponding author. Fax: +45 3532 1200.
E-mail address: cgrimmelikhuijzen@bi.ku.dk (C.J.P. Grimmelikhuijzen).
URL: http://www.bi.ku.dk/staff/staff-vip-details.asp?ID=90 (C.J.P. Grimmelikhuijzen).

0006-291X/$ - see front matter
doi:10.1016/j.bbrc.2005.12.062

Neuropeptides with the C-terminal sequence SIFamide have been isolated and identified from various crustaceans and insects [2,3]. The structure of *Drosophila* SIFamide is AYRKPPFNGSIFamide [3]. A comparison of the SIFamides from several insects and crustaceans showed that this peptide is evolutionarily extremely well conserved and that, in fact, only two isoforms, [G^1]-SIFamide and [A^1]-SIFamide, exist [3]. There has been one other variant published with the C-terminal sequence SLFamide [4], but this difference is probably due to a peptide sequencing error [3]. Also the SIFamide preprohormones have been determined both in crustaceans and in insects, as well as the neuronal localization of the peptide [2,3,5]. However, in spite of all these data, nothing is known about the physiological actions of the SIFamide neuropeptides.

In the present paper, we have cloned and characterized the SIFamide receptor from *Drosophila*. Our paper is the first report on the identification of a SIFamide receptor in invertebrates.

Materials and methods

Total RNA was isolated from all developmental stages of *Drosophila melanogaster*, using TRIzol Reagent (Invitrogen) and treated with DNase I (Ambion). cDNA was synthesized using the SMART RACE cDNA Amplification Kit (Clontech) and SuperScript III (Invitrogen).

The primers used for 3′RACE were: sense 5′-GTTTCCATGGA TGTGGGCCGGCGACAG-3′ and nested sense 5′-CCAGCGACCAGA GCAAAGGCGCCATC-3′, corresponding to positions 1957–1983 and 2058–2083 of Fig. 1. The primers used for 5′RACE were: antisense 5′-GCACTCGATGAGCCCAAGGAGCTG-3′ and nested antisense 5′-TCCCACCGCTACCACTAGTGCCATTC-3′, corresponding to positions 255–278 and 222–247 of Fig. 1. The primers used to amplify the coding region were: sense 5′-GAGATCTGCCACCATGATGGCAGCC AGCGGGCGGATTC-3′ and antisense 5′-GGAATTCCTATTCGTAG AGATCCGCATTCCTGCGAG-3′, where the underlined sequences correspond to positions 1–25 and 2249–2277 of Fig. 1.

The PCR programs used were: 3′RACE; primary PCR 94 °C for 3 min, then 5 cycles touchdown 94 °C for 30 s, 71 °C for 30 s decreasing 1 °C per cycle, and 72 °C for 3 min, followed by 15 cycles of 94 °C for 30 s, 66 °C for 30 s, and 72 °C for 3 min, and then a terminal elongation of 72 °C for 10 min. Nested PCR was performed as the primary PCR, except for using 30 cycles rather than 15. 5′RACE; primary PCR 94 °C for 3 min, then 5 cycles touchdown 94 °C for 30 s, 68 °C for 30 s decreasing 1 °C per cycle, 72 °C for 3 min, followed by 15 cycles of 94 °C for 30 s, 63 °C for 30 s, 72 °C for 3 min, and then a terminal elongation of 72 °C for 10 min. Nested PCR was performed as the primary PCR except for using 30 cycles rather than 15. Coding region; 94 °C for 3 min, then 40 cycles of 94 °C for 30 s, 65 °C for 45 s, and 72 °C for 4 min, followed by a terminal elongation of 72 °C for 10 min.

The PCR product of the coding region was cloned into pCR4-TOPO (Invitrogen), sequenced, and cloned into pIRES2-EGFP (Clontech) using *Bgl*II and *Eco*RI. We stably transfected Chinese hamster ovary (CHO) cells with this vector, which enabled us to clone cell lines (the greenest ones) that expressed the receptor most effectively. These cell lines were also stably transfected with DNA, coding for the promiscuous G protein, G-16, and transiently transfected with DNA, coding for apoaequorin [6,7]. Three hours before the assay we added the co-factor of apoaequorin, coelenterazine, to the cell culture medium. An activation of the expressed receptor in such pretreated cells would initiate an IP_3/Ca^{2+} cascade, leading to a strong bioluminescence response [6–10].

The correct sequence of the complete cDNA of Fig. 1, as well as that of the insertion in the expression construct, was verified by PCR and DNA sequencing.

DNA sequence comparisons were done using Vector NTI Advanced 9.0 (InforMax/Invitrogen). Protein sequence alignments were carried out, using ClustalW (http://npsa-pbil.ibcp.fr). BLAST searches were performed, using the sequences available at http://www.anobase.org/, http://www.bioinformatics.ksu.edu/BeetleBase/, http://www.flybase.org, http://www.hgsc.bcm.tmc.edu/projects/honeybee/, and (http://kaikoblast.dna.affrc.go.jp/cgi-bin/robo-blast/blast2.cgi?program=NA). Predictions of transmembrane helices were performed, using the TMHMM server (http://www.cbs.dtu.dk/services/TMHMM/).

Results

The *Drosophila* genome contains a gene, CG10823, annotated to code for an unidentified neuropeptide GPCR (http://www.flybase.org). We cloned the cDNA of this gene (Fig. 1), which contained the coding region, a 5′-noncoding region including several stop codons preceding the ATG start codon, and a very long 3′-noncoding region, which included the stop codon, but no polyadenylation signal. The coding region contained the information for a presumed GPCR, with seven-transmembrane α-helices, and five potential glycosylation sites, which were all located in the extracellular N terminus of the protein (Fig. 1).

We compared the cDNA of Fig. 1 with the genomic DNA sequences and found several nucleotide differences, which, however, did not result in amino acid differences of the receptor protein (Table 1). This comparison also showed that the gene had five introns (Table 2).

We stably transfected CHO cells with a vector containing the cDNA of the presumed receptor and isolated cloned cell lines that expressed the receptor most effectively. We further transformed and treated the cells in such a way that an activation of the expressed receptor would lead to a bioluminescence response (Materials and methods). We screened the transfected CHO cells with a library of 33 insect neuropeptides and eight biogenic amines. Only *Drosophila* SIFamide activated the receptor (Fig. 2), whereas the other compounds were without effect. A dose–response curve with *Drosophila* SIFamide showed an EC_{50} of 2×10^{-8} M (Fig. 2B).

We used the sequence of the *Drosophila* SIFamide receptor to carry out BLAST searches of the genomes from the other insect species with a sequenced genome, i.e., *A. gambiae*, *B. mori*, *T. castaneum*, and *A. mellifera*. These searches resulted in the identification of receptor orthologues in all four insects (Fig. 3). The receptor proteins from the five insects investigated (Fig. 3) are impressively similar and have high percentages of amino acid sequence identities. Compared to the seven-transmembrane area (TM I–VII) from *Drosophila*, *Anopheles* has 77%, *Tribolium* 73%, *Apis* 69%, and *Bombyx* 67% amino acid sequences in common. When similar (strongly conserved) residues were taken into account, these numbers were 87%, 84%, 82%, and 83%, respectively. Furthermore, the coding regions from all five receptor genes have two or three introns in common (at the same positions and with the same intron phasings) (Fig. 3). All these data strongly suggest that the GPCRs from Fig. 3 are SIFamide receptors.