The TRH-ortholog EFLamide in the migratory locust

Jan Veenstra, Ladislav Šimo

The TRH-ortholog EFLamide in the migratory locust

Jan A. Veenstra[1], Ladislav Šimo[2]

[1] INCIA UMR 5287 CNRS, University of Bordeaux, Pessac, France.

[2] UMR BIPAR INRA, Ecole Nationale Vétérinaire d'Alfort, ANSES, Université Paris-Est, Maisons-Alfort, France.

Corresponding Author:

Jan A. Veenstra[1]

INCIA UMR 5287 CNRS, Université de Bordeaux, allée Geoffroy St Hillaire, CS 50023, 33 615 Pessac Cedex, France

Email address: jan-adrianus.veenstra@u-bordeaux.fr

Abstract

Arthropod EFLamide genes in chelicerates, myriapods, decapods and non pterygote hexapods encode various EFLamide paracopies on a single precursor. However, in more advanced insect species such multiple EFLamide paracopies encoding genes are absent. In some Hemiptera putative exons of an EFLamide gene coding for a single EFLamide have been identified, while in the migratory locust a similar exon could potentially code for two EFLamide peptides. The recent identification of an EFLGamide from *Platynereis dumerilii* as the ligand for an ortholog of the TRH GPCR, suggested that the arthropod EFLamides might similarly activate TRH GPCR orthologs. We here identify the TRH GPCR ortholog from *Locusta migratoria* and show that it is activated in nanomolar concentrations by the two EFLamides previously predicted from this species. We also show that in the central nervous system there seems to be only a single bilateral neuron in the protocerebrum expressing this peptide. Given this very limited expression of EFLamide in locusts, it is perhaps not surprising that this gene and its receptor have been lost in many other insect species. This shows again that although neuropeptides and their receptors may persist in different evoltionary lineages, their functions can change dramatically.

Key words: *Locusta migratoria* – TRH – EFLamide – GPCR – evolution

1. Introduction

Many neuropeptides and their receptors originated very early during evolution and as a consequence their descendant vertebrate neuropeptides have orthologs not only in invertebrate deuterostomes, but also in protostomes (Mirabeau and Joly, 2013). In some cases both the structures of the ligand-receptor combinations and their functions show some similarity between the two major bilaterian lineages, but often the functions of what appear to be orthologs are distinctly different. For example, there is little evidence that the insect homologs of gonadotropin-releasing hormone are involved in the regulation of reproduction.

One such well known vertebrate neuropeptide is TRH (thyroid stimulating hormone (TSH)-releasing hormone), the peptide that forms part of the TRH-TSH-thyroid hormone feedback loop of the hypothalamus-pituitary-thyroid axis. The first indication for the existence of an insect TRH related peptide came from the identification of an ortholog of the TRH GPCR in the transcriptome from the brown planthopper *Nilaparvata lugens*, but there was no obvious ligand (Tanaka et al., 2014). Deorphanization of a TRH receptor ortholog from the annelid *Platynereis dumerlii* revealed its ligand to be FSEFLGamide (Bauknecht & Jékely, 2015). This peptide seems to be orthologous to the arthropod EFLamides, which were first described from chelicerates, where the genes have two different transcripts producing either EFLamides or EFLGGPamides (Veenstra, 2012), while some spiders have a second gene that produces exclusively a single EFLGGPamide (Veenstra, 2016a). The owl limpet *Lottia gigantea* and the annelids *Helobdella robusta*, *Capitella teleta* and *Platynereis dumerilii* have orthologous genes that similarly show alternative splicing (Veenstra 2010, 2011; Conzelmann et al., 2013). Although the alternative splicing of this gene thus likely predates the separation of the arthropods

and thelophotrochozoans, in pancrustaceans and the centipede *Strigamia maritima* EFLamide genes appear to have only a single transcript (Veenstra, 2016a,b).

EFLamide genes coding multiple EFLamide paracopies are present in apterygote hexapods (Derst et al., 2016) and in the genome of the mayfly *Danica ephemera*, but in more advanced insect species only partial putative cDNA or genomic sequences have been found. These include various hemiptera, including the bed bug *Cimex lectularius* (Predel et al., 2018), but also *Nilaparvata lugens*, the species with a GPCR similar to the TRH receptor, but also a dragonfly, a damselfly and the migratory locust (Veenstra, 2019). Some have questioned whether a small DNA sequence containing a sequence that could code for a single EFLamide is sufficient evidence for the existence of such a peptide. Indeed, some DNA sequences suggesting the existence of a neuropeptide gene can be misleading (Veenstra, 2017). So we felt it would be interesting to see whether we could get more convincing evidence for an EFLamide signaling system in insects. We choose to do this in the migratory locust *Locusta migratoria,* a species with a well studied neuropeptidome. Furthermore, in the locust the putative EFLamide precursor still has two EFLamide copies, which suggests that EFLamide may be physiologically more important in this species than in those in which the gene codes for only a single paracopy.

We prepared antisera to both EFLamide and its putative precursor and used these to provide evidence that in the locust this neuropeptide gene is only expressed in a single bilateral brain neuron. In order to show that EFLamide is a functional neuropeptide in this species we then identified the locust homolog of the TRH GPCR and showed that it is activated at nanomolar concentrations of the previously predicted *Locusta* EFLamides.

2. Materials & Methods

2.1. Insects

Locusta migratoria and the two-spotted cricket *Gryllus bimaculatus* were obtained from a local pet shop, the American cockroach *Periplaneta americana* was from a small laboratory culture started from animals gifted by Peter Kloppenburg (Köln, Germany). The common stick insect *Carausius morosus* was raised on blackberry leaves from eggs purchased on e-bay from Spain.

2.2. Peptides

The following peptides were custom synthetized by PhtdPeptides Co, Ltd (Zhengzhou City, China) : *Procambarus* AMGSEFLamide (95% purity), *Locusta* EFLamide-1 (NLGSEFL-amide, 95% purity), *Locusta* EFLamide-2 (MENLGSEFL-amide) 95% and *Locusta* EFLamide precursor peptide (CMQQSTTYNDLVQFRVV, 85% purity). The first of these corresponds to an EFLamide peptide predicted from the red swamp crayfish *Procambarus clarkii* (Veenstra, 2016) and the others are the predicted mature *Locusta* EFLamides as well as part of the predicted *Locusta* EFLamide precursor (Fig. 1). TRH (TSH releasing hormone) was from Genscript (Leiden, Netherlands).

2.3. EFLamide antisera

Three different polyclonal antisera were produced, a rabbit and a mouse antiserum to AMGSEFLamide conjugated to bovine serum albumen (BSA) through its N-terminal using Difluorodinitrobenzene as described (Tager, 1976) and a rabbit antiserum to part the putative

Locusta EFLamide precursor (Veenstra, 2019). For the latter CMQQSTTYNDLVQFRVV was conjugated through its N-terminal cysteine to BSA using sulfo-SMCC (Sulfosuccinimidyl-4-[N-maleimidomethyl]cyclohexane-1-carboxylate). Antisera were produced by Pineda Antikörper-Service (Berlin, Germany).

2.4. Immunohistology

Insects were dissected and fixed in 4% paraformaldehyde in PBS for 1 to 4 hrs at room temperature. After eight 30 minute washes in PBSAT (PBS containing 0.1% sodium azide and 1% Triton X-100) tissues were incubated in 10% normal goat serum for one hour, followed by incubation in primary antiserum for three to six days. Tissues were washed eight times for 30 minutes in PBSAT and incubated for one hour in 10% normal goat serum. This was followed by incubation in secondary antiserum for one night to three days. Tissues were again washed eight times for 30 minutes and then transferred to 20% glycerol for 15 minutes. Incubation for 15 minutes in 40%, 60% and 80% glycerol followed and tissues were then mounted n 80% glycerol. All washes and antibody incubations were done at room temperature under gentle shaking. Primary antiserum dilutions were 1:2000 for the mouse anti-EFLamide, 1: 8,000 for the rabbit EFLamide and 1: 8,000 for the EFLamide precursor antiserum. Other primary antisers used were anti-corazonin (Veenstra, 1991), anti-leucokinin IV antiserum (Chen et al., 1994) and an anti-Crustacean Hyperglycemic Hormone (CHH) from the crayfish *Orconectes limosus* (Keller, 1988); all these antisera were raised in rabbits. DyLight-488-conjugated goat anti-mouse IgG and DyLight-549-conjugated goat anti-rabbit IgG (Jackson ImmunoResearch Europe, UK) were used as secondary antibodies at a 1:1,000 dilution. Incubations in primary antisera were done for three

to seven days and secondary antisera for two to three days. Preabsorption experiments were performed using 10 nmol of peptide per ml of diluted antiserum.

2.5. Bioinformatics

Transcripts for neuropeptides and GPCRs were produced as described in detail elsewhere (*e.g.* Veenstra and Khammassi, 2017) essentially using the SRA Toolkit (https://www.ncbi.nlm.nih.gov/sra/docs/toolkitsoft/) and Trinity (Grabherr et al., 2011). A listing of the various SRAs used is provided as supplementary data (page 2). A phylogenetic tree was made for the EFLamide and TRH GPCRs using only the conserved parts of the receptors, *i.e.* from the start of transmembrane region (TM) 1 through the end of TM 5 and then from the beginning of TM6 through the end of TM7. Sequence alignment was done using Clustal Omega (Sievers et al., 2011) on the desktop and was manually inspected using Seaview (Gouy et al., 2010). Seaview was also used for selecting the conserved protein regions used for making the trees using PhyML (Guindon and Gascuel, 2003).

2.6. Putative EFLamide precursors and receptors

Initial attempts to find the *Locusta* ortholog of *Nilaparvata* GPCR A45 (Tanaka et al., 2014) from the public *Locusta* SRAs were unsuccessful. We therefore tried to find it using the public SRAs from three *Gryllus* species – *G. texensis, G. bimaculatus and G. pennsylvanicus.* Incomplete and very similar sequences were obtained from all three species (Supp. Data, page 4). These sequences were then used in an attempt to find the EFLamide receptor (EFLaR) exons in the genome of *Locusta migratoria.* However, neither the published genome (Wang et al.,

2014), nor the one available at the i5k Workspace (https://i5k.nal.usda.gov/locusta-migratoria) permitted reconstruction of the *Locusta* receptor as only some exons could be identified and some were present in several complete and incomplete copies. Using the obtained partial genomic sequences as bait to search for RNAseq reads in the available *Locusta* transcriptome SRAs two large parts of the *Locusta* receptor were reconstructed. Those two pieces were joined using the genome sequences from the i5k Workspace. This putative cDNA sequence of an EFLamide receptor still lacked the 3'end and this was obtained using a 3' Race protocol on *Locusta* brain cDNA. For this brains were dissected from adult *Locusta* and total RNA extracted using a kit from Macherey-Nagel. cDNA was produced with reverse trancription of 1 µg or total RNA using 5'-CGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGTTTTTTTTTTTTTTTTTT T-3' as primer and M-MuLV reverse transcriptase (New England Biolabs). The resulting cDNA was then amplified with two rounds of nested PCR using Q5® High-Fidelity DNA polymerase (New England Biolabs) first with primers 5'-CAGTGCAGGGTCCGAGGTAT-3' and 5'-GACCTGCGTCTTCGTCAACA-3' and then in the second with step 5'-CCGAGGTATTCGCACTGGATACGTTT-3' and 5'-ACAGCGCCATCAACCCCATC-3'. After the second PCR a single band was gel purified and sequenced. It yielded the missing 3'-sequence of the deduced putative *Locusta* EFLamide receptor. The complete sequence was submitted to Genbank and received accession number MN398190. Attempts to amplify the complete EFLamide receptor cDNA from brain cDNA were unsuccessful, possibly because the sequence is very GC-rich and/or because its expression is very low. The deduced amino acid sequence of this GPCR was codon optimized for expression in CHO cells by Biomatik Corporation

(Cambridge, ON Canada) and after addition of a Kozak site synthesized by the same company and cloned in pcDNA3.1(+).

Partial putative EFLamide precursors have been described from a number of insect species, including *L. migratoria* (Veenstra, 2019). Attempts were made to obtain the complete coding sequences using the same methodology as that described above for the putative EFLamide receptors.

2.7. Receptor assays

Transient coexpression of the *Locusta* EFLamide receptor with the human cytoplasmic aequorin reporter (Vernon and Printen, 2002) in Chinese hamster ovary (CHO-K1, Sigma) cells was used for the monitoring of intracellular calcium mobilization-triggered bioluminescence upon activation of the receptor (Simo et al 2011, 2013). The assays were performed in opaque 96-well microplates (Nunc) using the Fluostar Omega microplate reader (BMG Labtech). The cells were simultaneously co-transfected by the pcDNA3.1(+)/EFLaR, pcDNA3/Zeo(+)/human cytoplasmic aequorin and pcDNA3.1(+)/wild type human G protein alpha 15 subunit Gα15(16) (cDNA Resource Centre, Bloomsburg, University of Pennsylvania) constructs. The use of chimeric Gα15(16) subunit is advocated by its high effectivity to link the calcium mobilization signaling pathway to the transfected Gi/o coupled receptors (Offermans and Simon, 1995; Park et al., 2002, 2003). The cells were preequilibrated with coelenterazine-h (Promega) for the 3h in room temperature. Various doses of agonist ligands in 50 μL were plated in each well followed by the injection of the 50 μl cell suspension (~15,000 cells). Immediately after the injections, the changes in luminescence were monitored every 0.5 second during 25 seconds interval and their

integrated values over time were normalized to the largest positive control response in each plate after background subtractions. The data obtained were analyzed with Excel 16.16.13 (Microsoft Office) and the dose response curves including the half maximum response values (EC_{50}) were calculated using Prism 5 software package (GraphPad Software, San Diego California USA).

The negative control consisted of cells transfected with the reporter construct aequorin and Gα15(16) only, which did not show any response to the ligands tested. In addition the 10μM concentration of *Ixodes scapularis* myoinhibitory peptide 2 and *Drosophila melanogaster* SIFamide were used as another negative control for transfected cells.

3. Results

3.1. Immunohistology

Antisera against EFLamide recognizes two different cell types in the brain of the migratory locust (Fig. 2A,B). Most prominently there is a bilateral interneuron located laterally below the calyx that innervates a large number of different areas in the brain. Of these the lower divisions of the anterior optic tubercles and the ventral layer of the anterior lobe of the lobula are very prominently innervated. Other easily recognizable areas are the central body and lower units of the anterior optic tubercles, while a very fine axon is often found innervating the olfactory glomeruli in the deutocerebrum (Fig. 2A).

At two- to four-fold higher concentrations there are lateral neuroendocrine cells that project to the corpora cardiaca. The latter cells are located close to but different from the corazonin neuroendocrine cells (Fig. 2B) and similarly project through the nervus corporis

cardiaci II to the corpus cardiacum. The EFLamide antisera also recognize neuroendocrine
216 lateral cells in *Carausius morosus* and in that species those cells are strongly immunoreactive
(Fig. 2C). In the American cockroach, *Periplaneta americana*, the bilateral pair is missing and
218 the lateral neuroendocrine cells are also strongly immunoreactive (Fig. 2D). Like the various
RFamide peptides, EFLamide has both an aromatic and a charged amino acid in its C-terminal.
220 Antisera to RFamides are known to easily cross-react with other neuropeptides sharing just a C-
terminal RFamide or even RYamide (*e.g.* Veenstra and Schooneveld, 1984; Veenstra and
222 Khammassi, 2017). Given the absence of any trace of an EFLamide receptor in the *Periplaneta*
genome (Li et al., 2018) or the various brain transcriptome SRAs, it seemed likely that the
224 immunoreactivity in the lateral neuroendocrine cells was due to cross-reactivity with a different
neuropeptide. Some of these lateral cells were also immunoreactive with a leucokinin antiserum
226 and all of them were recognized by an antiserum to CHH from *Orconectes limosus*. As this
antiserum is known to recognize its insect ortholog ion transport peptide (ITP; Dircksen, 2009),
228 it seemed plausible that both the EFLamide antisera also recognize at least one of the ITP
isoforms. Analysis of brain transcriptome SRAs from *P. americana* and *Carausius morosus*
230 revealed the short form of ITP from this species to have a C-terminal EFLamide, while in
Locusta the C-terminal has an EILamide sequence (Table 1). This convincingly shows that the
232 presence of EFLamide immunoreactivity in neurons does not prove that such neurons express
EFLamide. We therefore made another antiserum to part of the previously identified putative
234 *Locusta* EFLamide precursor (Veenstra, 2019). This antiserum only labels the the bilateral
neuron in the brain (Fig. 2E), thus suggesting that this is the only neuron that expresses
236 EFLamide in *Locusta*. No other immunoreactive EFLamide or EFLamide precursor neurons

were found in the nervous system. Immunoreactive enteroendocrine cells were neither found and although we attempted to find EFLamide immunoreactivity at the periphery, such efforts were without success.

Immunoreactivity was reduced but not completely abolished after preincubation of the EFLamide antisera with each of the three EFLamide peptides. The immunoreactivity of the antiserum to EFLamide precursor was similary diminished after preincubation with the peptide used to make the antiserum.

3.2. EFLamide precursors

Various attempts were made to use the transcriptome SRAs from three *Gryllus* species and *L. migratoria* to construct complete cDNAs for the putative EFLamide precursors. The results of these attempts resemble those obtained after similar efforts for RYamide precursor transcript from *Drosophila melanogaster* (Veenstra and Khammassi, 2017). Thus, one can find the exon coding the EFLamide peptide(s), but exons upstream are not identified. Those RNAseq reads that contain sequence 5' to the coding exons either contain in frame stop codons (*Gryllus*) or coding sequences from a different protein (*Locusta*). In the case of the locust some Trinity produced transcripts were obtained that at first sight looked credible as they have a signal peptide. However, the RNAseq reads that had been added to the putative EFLamide precursor reads belong to the *Locusta* ortholog of the protein first identified from honeybees as IDLSRFYGHFNT-containing or prohormone 4 (Hummon et al., 2006). The latter protein is very well conserved (Suppl. Data, page 3), and analysis of RNAseq data show that the very large

majority of RNAseq reads support a typical IDLSRFYGHFNT-containing ortholog and hence this transcript has to be an artifact.

3.3. EFLamide receptor

It should probably not come as a surprise that when it is difficult or impossible to construct the cDNA sequence for the neuropeptide that the one for its receptor is also complicated. In the end a complete transcript could be constructed by combining RNAseq derived partial transcripts with genomic sequences and a PCR product. The translated sequence of the *Locusta* EFLamide receptor cDNA is predicted to have a signal peptide in addition to seven transmembrane regions typical of a GPCR.

Together with public and ad hoc constructed orthologs and a few vertebrate TRH GPCRs (for complete protein sequences and alignment see Suppl. Data, pages 5 and 10 respectively) a phylogenetic tree was constructed that was rooted using the A isoform of the *Drosophila* ETH receptor (the *Drosophila* GPCR that is most closely related to the EFLamide receptor). The results show the arthropod EFLamide GPCRs as a single cluster in which the branch length is considerably smaller in the chelicerates than in insects and crustaceans, suggesting that in the latter species natural selection to maintain the sequence of this receptor is less intense than in chelicerates.

3.4. EFLamide receptor activation

The two predicted EFLamide peptides from *L. migratoria* were highly active on the EFLamide receptor, with EC_{50} values of 130.2 nM and 138.3nM for *Locusta* EFLamide-1 and

EFLamide-2 respectively (Fig. 4A). Slightly less robust luminescent responses were also generated by *Procambarus* EFLamide with an EC_{50} value of 532.3 nM (Fig. 4A). The EFLamide receptor was much less sensitive to different doses of TRH with EC_{50} value of 1.499 μM, when compared to those of tested EFLamide peptides. For all of the active doses of the EFLamide, the increase in luminescence was observed after ~3 second after the ligand was applied, and this luminescence peaked within the 5-12s depending of ligand concentration (Fig. 4B). The ligands tested in this study showed no activity on cells that were not transfected with EFLamide receptor indicating that the luminescent responses were specifically mediated by the transfected receptor. The cells that expressed EFLamide receptor were activated only by EFLamides and TRH, and did not respond to the other tested ligands.

4. Discussion

Antisera raised to EFLamide recognized two different cell types in the migratory locust, one of which was only weakly labeled in this species, but much more intensely in a stick insect and the American cockroach. The latter cell type produces ITP. One of the ITP isoforms has an EFLamide C-terminal in *Periplaneta* and *Carausius,* two species in which the ITP neuroendocrine cells react strongly with the EFLamide antisera. In *Locusta* on the other hand the corresponding ITP isoform has a C-terminal EILamide and the ITP cells are only weakly EFLamide immunoreactive. This suggested that the EFLamide antisera cross-reacts with these ITP isoforms. We thus produced another antiserum to a different part of the *Locusta* EFLamide precursor. This second antiserum only recognizes the bilateral neuron in the brain and we thus conclude that these are the only neurons that express the EFLamide gene in *Locusta.*

302 Antisera to RFamide neuropeptides similarly cross-react with different neuropeptides like e.g. FMRFamide and myosuppressin. Although both of FMRFamide and myosuppressin have
304 their own typical receptor, it has been shown that in the prothoracic gland of *Bombyx mori* the myosuppressin receptor in this tissue can be activated by both myosuppressin, presumably as a
306 hormone, and FMRFamide as a neuromodulator directly released in the tissue by efferent neurons (Yamanaka et al., 2005, 2006). This raises the question whether there might be similar
308 interactions between EFLamide and the short isoforms of ITP. Although this may seem speculative, it has already been demonstrated that in *Bombyx* the tachykinin receptor can be
310 activated by both the long form of ITP and tachykinins (Nagai et al., 2014; Nagai-Okatani et al., 2016).

312 To support the argument that despite its very limited expression the EFLamide neuropeptide is functional in the locust we identified its putative receptor based on previous
314 work in the annelid *Platynereis dumerilii* that showed that in this species a TRH GPCR ortholog is activated by EFLGamide (Bauknecht and Jékely, 2015). The scarcity of RNAseq reads in the
316 various *Locusta* transcriptome SRAs and the difficulty of amplifying this GPCR by RT-PCR suggest that the expression of the receptor like that of the EFLamide precursor is very low.
318 Nevertheless we were able to show that it is activated at nanomolar concentrations of the previously predicted EFLamides from this species.

320 Neuropeptides that are encoded by genes coding multiple structurally similar paracopies are typically released into the hemolymph. The potent physiological effects of EFLamide on the
322 motor patterns produced by the cardiac and stomatogastric ganglia in the lobster *Homarus americanus* (Dickinson et al., 2019) suggests that in Decapods EFLamide is likely released into

the hemolymph as a hormone. Such multiple paracopies coding EFLamide genes are also present in non-pterygote hexapods and mayflies, but in those advanced insects that still have this gene the number of paracopies is reduced to only one or two (Derst et al., 2016; Veenstra, 2019). This in combination with a relatively high EC_{50} value for the interaction of the peptides with their receptor in *Locusta* suggests that in the latter species, the peptide may have lost its hormonal function. The presence of only a single bilateral interneuron in the central nervous system that expresses the EFLamide gene reinforces this hypothesis.

The bilateral pair of ELFamide-cells in the locust brain have unusually wide ramifications in superior protocerebral brain areas. Interestingly, the areas innervated most densely, the lower layer of the anterior lobe of the lobula, the lower units of the anterior optic tubercles, and the central body and lateral complex are all involved in the processing of skylight polarization signals used for spatial orientation of the locust (Homberg et al., 2003, 2011). It is therefore, possible that the ELFamide neurons exert a modulatory input onto the navigation system of the locust by simultaneaous targeting several stages of the sky compass system in the locust brain.

Concentrations of neuropeptides are much lower in the hemolymph than locally inside the nervous system. It is therefore not surprising that the EC_{50} values of around 130 nM for the two *Locusta* EFLamides are higher than those typically found for neuropeptides released into the hemolymph and one would expect these values to be significantly lower in species where the peptides still function as hormones. This suggests that the evolutionary constraints on maintaining peptide and receptor structures will be relaxed once a neuropeptide is released exclusively within the central nervous system and no longer used as a neurohormone. The long

346 branch length of the *Locusta* EFLamide GPCR in the phylogenetic tree (Fig. 3) indeed suggests that this might be the case here. It is interesting to note that the branch length of the *Nilaparvata* 348 GPCR on this tree is even longer and that its apparent place on the phylogenetic tree does not correspond to its true phylogenetic position, thus suggesting that here the evolutionary constraint 350 on maintaining the receptor is still much weaker. Indirectly, it suggests that EFLamide in *Nilaparvata* is not released in large quantities.

352 Our results convincingly show that the EFLamide neuropeptide is expressed in a very small number of neurons. This no doubt explains that it has never been physically identified as 354 such in insects. Nevertheless, the identification of a *Locusta* EFLamide receptor that is activated by nanomolar concentrations of two predicted EFLamides must lead to the conclusion that this 356 neuropeptide signaling system is still functional in this species. However, this very limited expression of this neuropeptide also suggests that whatever its original function as an arthropod 358 neurohormone may have been, that function must have become superfluous as insects evolved.

360 *Acknowledgements.* We thank Uwe Homberg (Philipps-Universität, Marburg, Germany) for for helpful suggestions to improve the manuscript. Institutional funding from the CNRS is also 362 acknowledged.

364 **References**

366 Bauknecht P, Jékely G. 2015. Large-scale combinatorial deorphanization of *Platynereis* neuropeptide GPCRs. Cell Reports 12:684-693. doi: 10.1016/j.celrep.2015.06.052.

368

Chen, Y., Veenstra, J.A., Davis, N.T., Hagedorn, H.H. 1994. A comparative study of leucokinin-
370 immunoreactive neurons in insects. Cell Tissue Res. 276, 69-83. doi: 10.1007/bf00354786.

372 Derst, C., Dircksen, H., Meusemann, K., Zhou, X., Liu, S., Predel, R. 2016. Evolution of neuropeptides in non-pterygote hexapods. BMC Evol. Biol. 16:51. doi: 10.1186/s12862-016-
374 0621-4.

376 Dickinson, P.S., Dickinson, E.S., Oleisky, E.R., Rivera, C.D., Stanhope, M.E., Stemmler, E.A., Hull, J.J., Christie, A.E. 2019. AMGSEFLamide, a member of a broadly conserved peptide
378 family, modulates multiple neural networks in *Homarus americanus*. J. Exp. Biol. 222(Pt 2). pii: jeb194092. doi: 10.1242/jeb.194092.

380

Dircksen, H. 2009. Insect ion transport peptides are derived from alternatively spliced genes and
382 differentially expressed in the central and peripheral nervous system. J. Exp. Biol. 212:401-12. doi: 10.1242/jeb.026112.

384

Gouy, M., Guindon, S., Gascuel, O. 2010. SeaView version 4: a multi platform graphical user
386 interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol.27, 221–224. doi: 10.1093/molbev/msp259.

388

Grabherr, M.G., Haas, B.J,, Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X.,
390 Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N.,
di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A. 2011. Full-
392 length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol.
29:644-52. doi: 10.1038/nbt.1883.
394
Guindon, S., Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large
396 phylogenies by maximum likelihood. Syst. Biol.52, 696–670. doi:10.1080/10635150390235520.

398 Homberg, U., Hofer, S., Pfeiffer, K., Gebhardt, S. 2003. Organization and neural connections of
the anterior optic tubercle in the brain of the locust, *Schistocerca gregaria*. J. Comp. Neurol.
400 462, 415-430. doi: 10.1002/cne.10771.

402 Homberg, U., Heinze, S., Pfeiffer, K., Kinoshita, M., el Jundi, B. 2011. Central neural coding of
sky polarization in insects. Philos. Trans. R. Soc. B 366, 680-687. doi: 10.1098/rstb.2010.0199.
404
Hummon, A.B., Richmond, T.A., Verleyen, P., Baggerman, G., Huybrechts, J.,Ewing, M.A.,
406 Vierstraete, E., Rodriguez-Zas, S.L., Schoofs, L., Robinson, G.E.,Sweedler, J.V. 2006. From the
genome to the proteome: uncovering peptides in the *Apis* brain. Science 314, 647–649. doi:
408 10.1126/science.1124128.

410 Keller, R., 1988. Radioimmunoassays and ELISAs: peptides, in: Gilbert, L.I., Miller, T.A. (Eds.), Immunological Techniques in Insect Biology. Springer, Heidelberg, pp. 253 -272.

412

Li, S., Zhu, S., Jia, Q., Yuan, D., Ren, C., Li, K., Liu, S., Cui, Y., Zhao, H., Cao, Y., Fang, G.,

414 Li, D., Zhao, X., Zhang, J., Yue, Q., Fan, Y., Yu, X., Feng, Q., Zhan, S. 2018. The genomic and functional landscapes of developmental plasticity in the American cockroach. Nat. Commun.

416 9:1008. doi: 10.1038/s41467-018-03281-1.

418 Mirabeau, O., Joly, J.S. 2013. Molecular evolution of peptidergic signaling systems in bilaterians. Proc. Natl. Acad. Sci. USA 110:E2028-E2037. doi: 10.1073/pnas.1219956110.

420

Nagai, C., Mabashi-Asazuma, H., Nagasawa, H., Nagata, S. 2014. Identification and

422 characterization of receptors for ion transport peptide (ITP) and ITP-like (ITPL) in the silkworm *Bombyx mori*. J. Biol. Chem. 289:32166-32177. doi: 10.1074/jbc.M114.590646.

424

Nagai-Okatani, C., Nagasawa, H., Nagata, S. 2016. Tachykinin-related peptides share a G

426 pProtein-coupled receptor with ion transport peptide-like in the silkworm *Bombyx mori.* PLoS One. 11:e0156501. doi: 10.1371/journal.pone.0156501.

428

Offermanns, S., Simon, M.I., 1995. $G\alpha_{15}$ and $G\alpha_{16}$ couple a wide variety of receptors to

430 phospholipase C. J. Biol. Chem. 270:15175-15180. doi: 10.1074/jbc.270.25.15175.

Park, Y., Kim, Y.J., Adams, M.E. 2002. Identification of G protein-coupled receptors for *Drosophila* PRXamide peptides, CCAP, corazonin, and AKH supports a theory of ligand-receptor coevolution. Proc. Natl. Acad. Sci. USA 99, 11423-11428. doi: 10.1073/pnas.162276199.

Park, Y., Kim, Y.J., Dupriez, V., Adams, M.E. 2003. Two subtypes of ecdysis-triggering hormone receptor in *Drosophila melanogaster.* J. Biol. Chem. 278, 17710-17715. doi: 10.1074/jbc.M301119200.

Predel, R., Neupert, S., Derst, C., Reinhardt, K., Wegener, C. 2018. Neuropeptidomics of the bed bug *Cimex lectularius*. J. Proteome Res. 17, 440–454. doi: 10.1021/acs.jproteome.7b00630.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol.7, 539. doi: 10.1038/msb.2011.75.

Šimo, L., Koči, J., Žitňan, D., Park, Y. 2011. Evidence for D1 dopamine receptor activation by a paracrine signal of dopamine in tick salivary glands. PLoS One 6, e16158. doi: 10.1371/journal.pone.0016158.

Šimo, L., Koči, J., Park, Y. 2013 Receptors for the neuropeptides, myoinhibitory peptide and
454 SIFamide, in control of the salivary glands of the blacklegged tick *Ixodes scapularis*. Insect
Biochem. Mol. Biol. 43:376–387. doi: 10.1016/j.ibmb.2013.01.002.
456

Tager, H.S. 1976. Coupling of peptides to albumin with difluorodinitrobenzene. Anal. Biochem.
458 71, 367-375.

460 Tanaka, Y., Suetsugu, Y., Yamamoto, K., Noda, H., Shinoda, T. 2014. Transcriptome analysis of
neuropeptides and G-protein coupled receptors (GPCRs) for neuropeptides in the brown
462 planthopper *Nilaparvata lugens.* Peptides 53:125-133. doi: 10.1016/j.peptides.2013.07.027.

464 Veenstra, J.A. 1991. Presence of corazonin in three insect species, and isolation and
identification of [His^7]corazonin from *Schistocerca americana*. Peptides 12:1285-1289. doi:
466 10.1016/0196-9781(91)90208-7.

468 Veenstra, J.A. 2000. Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine
peptide precursors. Arch. Insect Biochem. Physiol. 43: 49-63. doi: 10.1002/(SICI)1520-
470 6327(200002)43:2<49::AID-ARCH1>3.0.CO;2-M

472 Veenstra, J.A. 2016a. Neuropeptide evolution: Chelicerate neurohormone and neuropeptide
genes may reflect one or more whole genome duplications. Gen. Comp. Endocrinol. 229:41-55.
474 doi: 10.1016/j.ygcen.2015.11.019.

476 Veenstra, J.A. 2016b. Similarities between decapod and insect neuropeptidomes. PeerJ 4:e2043. doi: 10.7717/peerj.2043.

478

Veenstra, J.A. 2017. The salivary gland salivation stimulating peptide from *Locusta migratoria* 480 (Lom-SG-SASP) is not a typical neuropeptide. PeerJ 5:e3619. doi: 10.7717/peerj.3619.

482 Veenstra, J.A. 2019. Two Lys-vasopressin-like peptides, EFLamide, and other phasmid neuropeptides. Gen. Comp. Endocrinol. 278:3-11. doi: 10.1016/j.ygcen.2018.04.027.

484

Veenstra, J.A., Khammassi, H. 2017. Rudimentary expression of RYamide in *Drosophila* 486 *melanogaster* relative to other *Drosophila* species points to a functional decline of this neuropeptide gene. Insect Biochem. Mol. Biol. 83:68-79. doi: 10.1016/j.ibmb.2017.03.001.

488

Veenstra, J.A., Rombauts, S., Grbić, M. 2012. In silico cloning of genes encoding neuropeptides, 490 neurohormones and their putative G-protein coupled receptors in a spider mite. Insect Biochem. Mol. Biol. 42:277-295. doi: 10.1016/j.ibmb.2011.12.009.

492

Veenstra, J.A., Schooneveld, H. 1984. Immunocytochemical localization of neurons in the 494 nervous system of the Colorado potato beetle with antisera against FMRFamide and bovine pancreatic polypeptide. Cell Tissue Res. 235, 303-308. doi: 10.1007/bf00217854.

496

Vernon, W.I., Printen, J.A. 2002. Assay for intracellular calcium using a codon optimized
498 aequorin. Biotechniques 33, 730-734. doi: 10.2144/02334bm02.

500 Wang, X., Fang, X., Yang, P., Jiang, X., Jiang, F., Zhao, D., Li, B., Cui, F., Wei, J., Ma, C.,
Wang, Y., He, J., Luo, Y., Wang, Z., Guo, X., Guo, W., Wang, X., Zhang, Y., Yang, M., Hao,
502 S., Chen, B., Ma, Z., Yu, D., Xiong, Z., Zhu, Y., Fan, D., Han, L., Wang, B., Chen, Y., Wang, J.,
Yang, L., Zhao, W., Feng, Y., Chen, G., Lian, J., Li, Q., Huang, Z., Yao, X., Lv, N., Zhang, G.,
504 Li, Y., Wang, J., Wang, J., Zhu, B., Kang, L. 2014. The locust genome provides insight into
swarm formation and long-distance flight. Nat. Commun. 5:2957. doi: 10.1038/ncomms3957.
506

Yamanaka, N., Hua, Y.J., Mizoguchi, A., Watanabe, K., Niwa, R., Tanaka, Y., Kataoka, H.
508 2005. Identification of a novel prothoracicostatic hormone and its receptor in the silkworm
Bombyx mori. J. Biol. Chem. 280(15):14684-14690. doi: 10.1074/jbc.M500308200.
510

Yamanaka, N., Zitnan, D., Kim, Y.J., Adams, M.E., Hua, Y.J., Suzuki, Y., Suzuki, M., Suzuki,
512 A., Satake, H., Mizoguchi, A., Asaoka, K., Tanaka, Y., Kataoka, H. 2006. Regulation of insect
steroid hormone biosynthesis by innervating peptidergic neurons. Proc. Natl. Acad. Sci. USA
514 103:8622-8627. doi: 10.1073/pnas.0511196103.

Table 1

Peptide sequences. Sequences of the two predicted EFLamides from *Locusta migratoria* (Veenstra, 2019). For comparison the *Procambarus* EFLamide used as antigen for the EFLamide antiserum as well as the C-termini of the short isoforms of ion transport peptide from *L. migratoria, P. americana* and *C. morosus* are also provided. Note that these peptides show sufficient structural similarity to EFLamide to explain the cross-reactivity of the EFLamide antiserum with these ion transport peptides. The peptide used to make a *Locusta* EFLamide precursor antiserum is also shown to demonstrate that it has no sequence similarity to EFLamide.

Locusta EFLamide-1	NLGSEFLamide
Locusta EFLamide-2	MENLGSEFLamide
Procambarus EFLamide	AMGSEFLamide
C-terminal of *Locusta* ITP	---KFNQMVEILamide
C-terminal of *Periplaneta* ITP	---KFNQMAEFLamide
C-terminal of *Carausius*	---NFSQMVEFLamide
Locusta EFLamide precursor peptide	CQQSTTYNDLVQFRVV

Legends to the figures.

Fig. 1. Partial genomic sequence of *Locusta migratoria* and its plausible translation into the last coding exon of the putative EFLamide gene of this species. Small lettering corresponds to the DNA sequence, with the untranslated intron and 3'-sequence in italic lower case and the translated in bold capitals while the large lettering corresponds to the predicted protein sequence. The two predicted *Locusta* EFLamides are indicated in light blue, with the Gly residues predicted to be transformed in C-terminal amides in dark blue. Predicted convertase cleavage sites are highlighted in red [processing at the single Arg residue is supported by the presence of an Arg residue in the P4 pocket (Veenstra, 2000)]. Yellow highlighting indicates the peptide used for making *Locusta* EFLamide precursor specific antiserum.

Fig. 2. Immunohistological localization of EFLamide immunoreactivity. **A.** Adult female *Locusta migratoria* brain showing the EFLamide immunoreactive bilateral neuron (1). Note the strong immunoreactivity in the lower units of the anterior optic tubercles (2) and the ventral layer of the anterior lobe of the lobula (3); immunostaining is also prominent in the central body (4) and lateral complex (5). **B.** EFLamide- (Bi and Bii) and corazonin-immunoreactive cells (Bii and Biii) in green and magenta respectively in the pars lateralis of *Locusta*. Note that the immunoreactivity in the bilateral neuron (in the lower part of the pictures) is much stronger than that in the lateral neuroendocrine cells. **C.** EFLamide- (Ci and Cii) and corazonin-immunoreactive cells (Cii and Ciii) in green and magenta respectively in the pars lateralis of *Carausisu morosus*. Note that the immunoreactivity in the EFLamide immunoreactive

554 neuroendocrine cells in this species is much stronger than in *Locusta* (Bi and Bii). **D.** EFLamide-(Di and Dii) and CHH-immunoreactive cells (Dii and Diii) in green and magenta respectively in
556 the pars lateralis of *Periplaneta americana.* Note that the immunoreactivity in the EFLamide immunoreactive neuroendocrine cells in this species is much stronger than in *Locusta* and also
558 note the two antisera recognize exactly the same cells. **E.** EFLamide precursor (Ei and Eii) and EFLamide immunoreactivity (Eii and Eiii) in green and magenta respectively in a last instar
560 *Locusta* larva. Note that the two antisera label exactly the same neurons and axons.

562 **Fig. 3.** Phylogenetic analysis of EFLamide and TRH GPCRs. The Arthropod EFLamide receptor clade is highlighted in blue, the vertebrate TRH GPCR clade in yellow. Note that the *Platynereis*
564 EFLGamide receptor seems somewhat more similar to the vertebrate TRH than to the arthropod EFLamide receptors. Whereas the branch lengths of the chelicerate EFLamide receptors are
566 relatively short, those of the other arthropod GPCRs are longer and that the branch lengths for the *Locusta* and *Nilaparvata* receptors are the longest. Also note that the location of the
568 *Nilaparvata* GPCR on the tree does not correspond to its true phylogenetic position, suggesting significant relaxation of natural selection to maintain its primary sequence. For individual
570 sequences and other details see Supplementary Data. PhyML ln(L)=−7073.0 281 sites LG 100 replic. 4 rate classes.

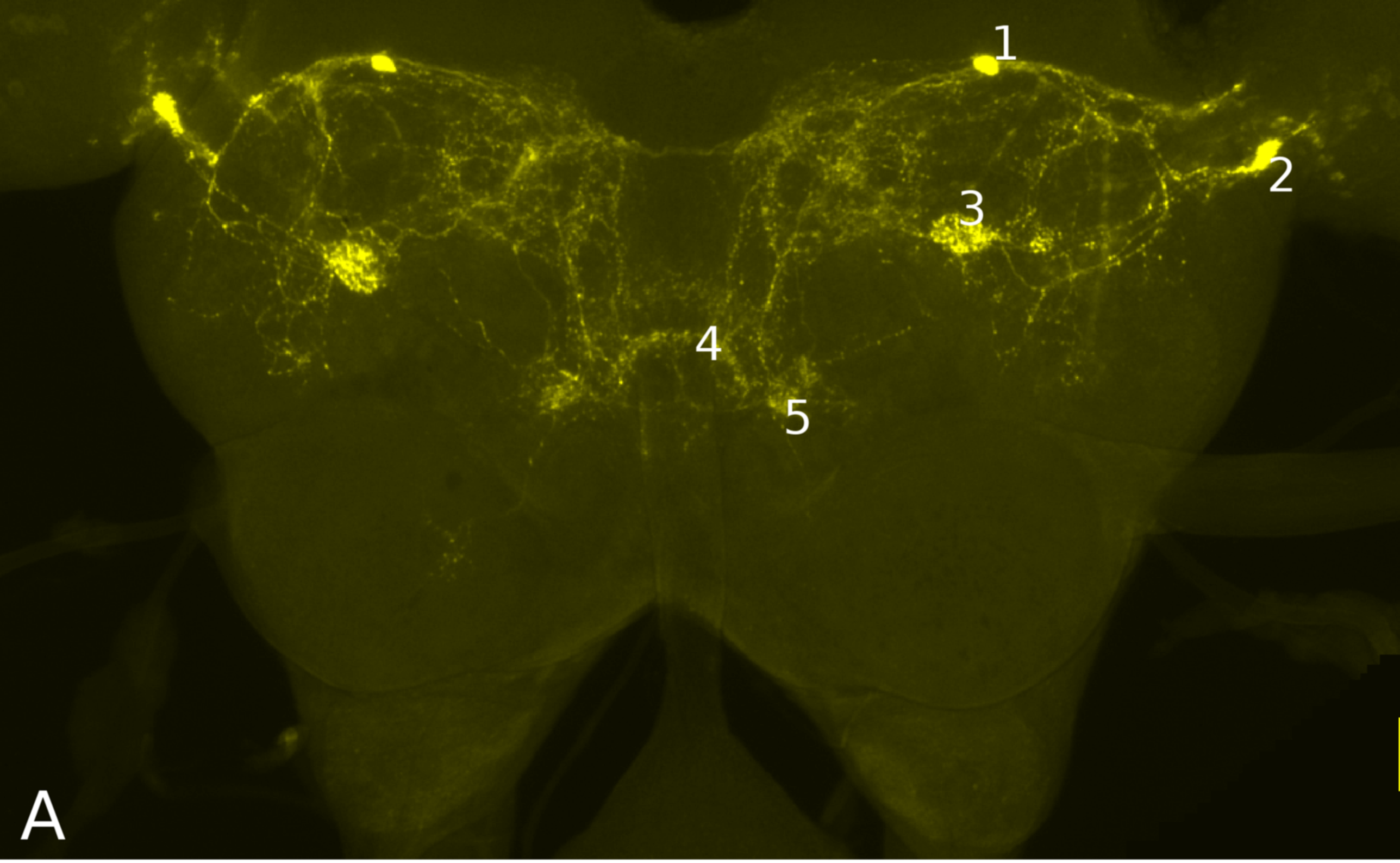
572

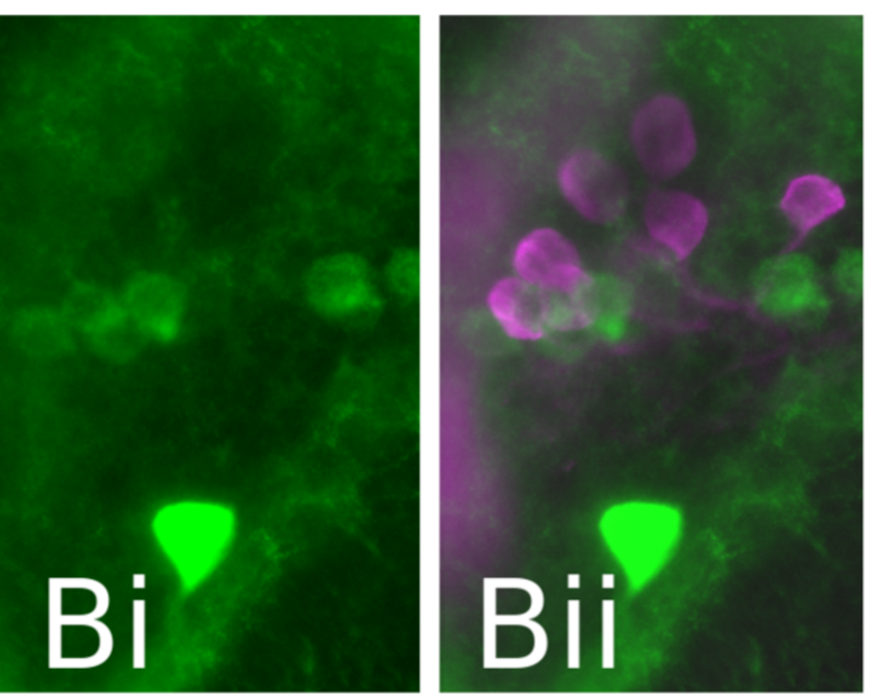
Fig. 4. Cellular responses observed after ligand-mediated calcium mobilization in aequorin
574 reporter assays in CHO-K1 cells that were transfected with EFLamide receptor. (A) The dose response curves for *Locusta* EFLamide-1, EFLamide-2, *Procambarus* EFLamide and TRH. (B)

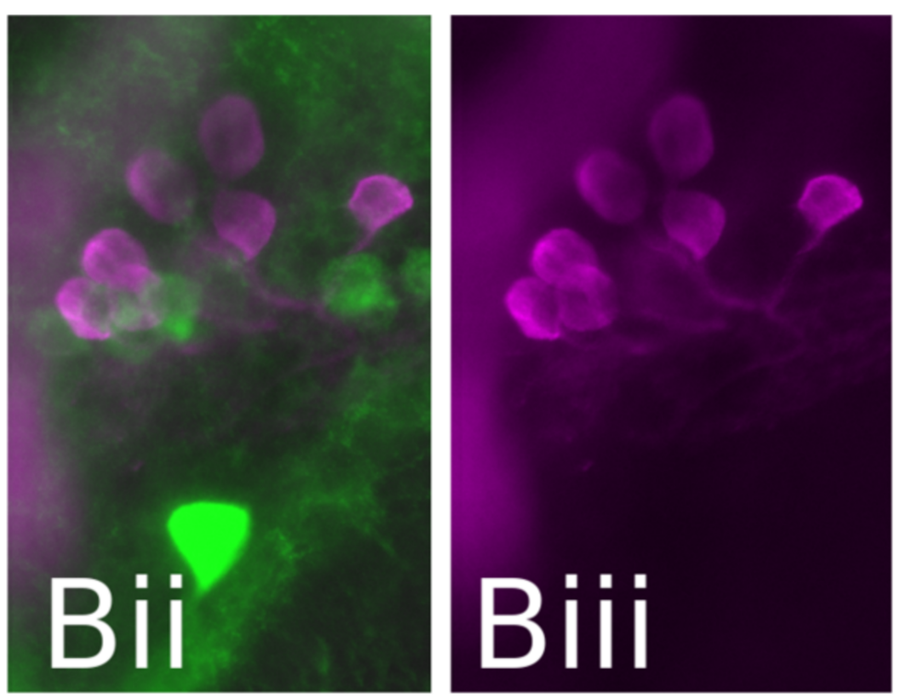
576 Typical representative responses of the cells when treated with different doses of EFLamides (*Locusta* EFLamide-2 in this case). Inset in (B) shows the integrated relative luminescent values
578 (RLU) calculated from 50 intervals within the 25 second responses. The bars in A indicate the standard error for a minimum of three replicated plates.

580

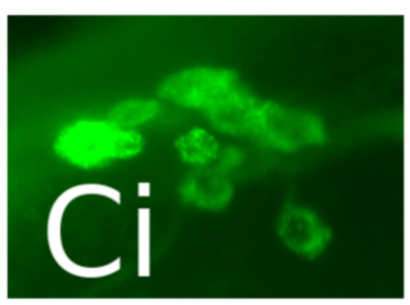

tgaacttttctgttccagAATTTCTGAAGACTCTCTCTCTTTTGAACGAGCGCCTCTCAGAGATGCAACAATCCACTACCTACAACGACTTAGTTCAATTTCGTGTTGTCAGAAACCTG
XFLKTLSLLNERLSEMQQSTTYNDLVQFRVVRNL
GGATCCGAATTCCTAGGAAAACGAATGGAAAACTTGGGATCAGAATTTCTAGGAAAACGAATGCAAAACTTGAAAAATATACTAGGAGTCAAAtaatgtaatatgtctcacctgctaca
GSEFLGKRMENLGSEFLGKRMQNLKNILGVK*

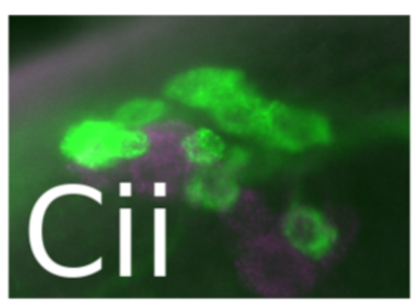

1
2
3
4
5
A


Bi

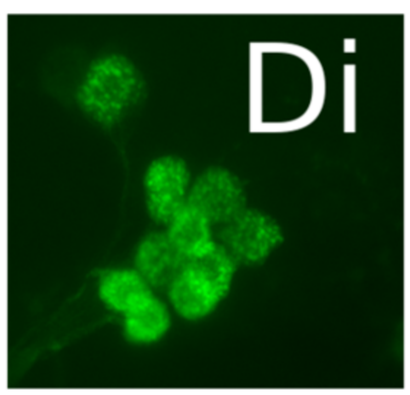

Bii

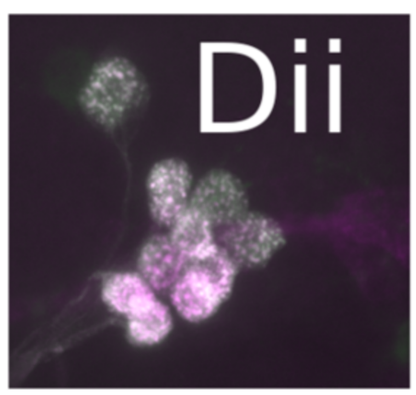
Biii

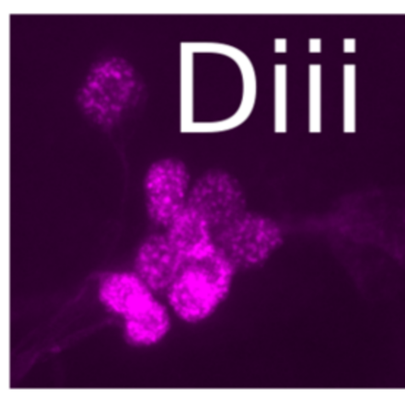

Ci

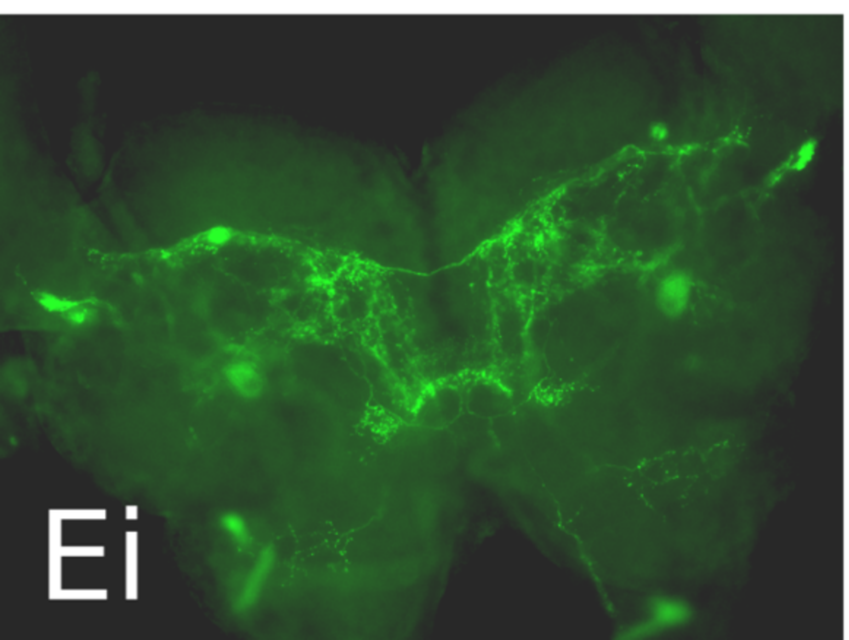

Cii

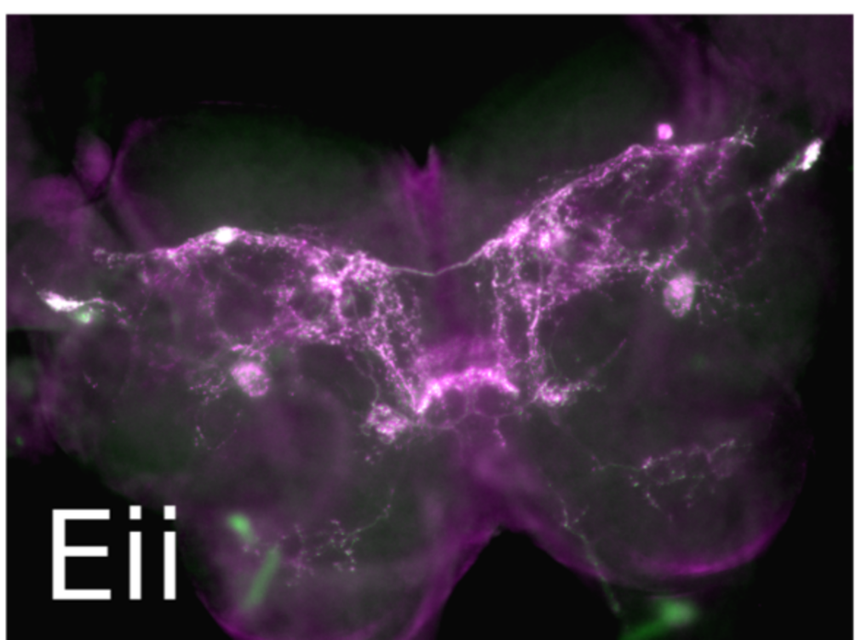

Cii

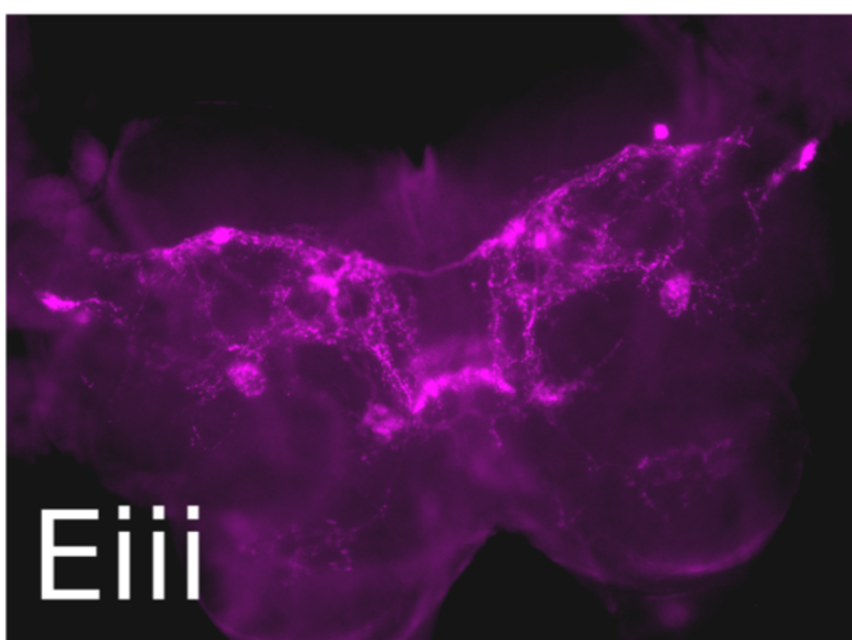

Di

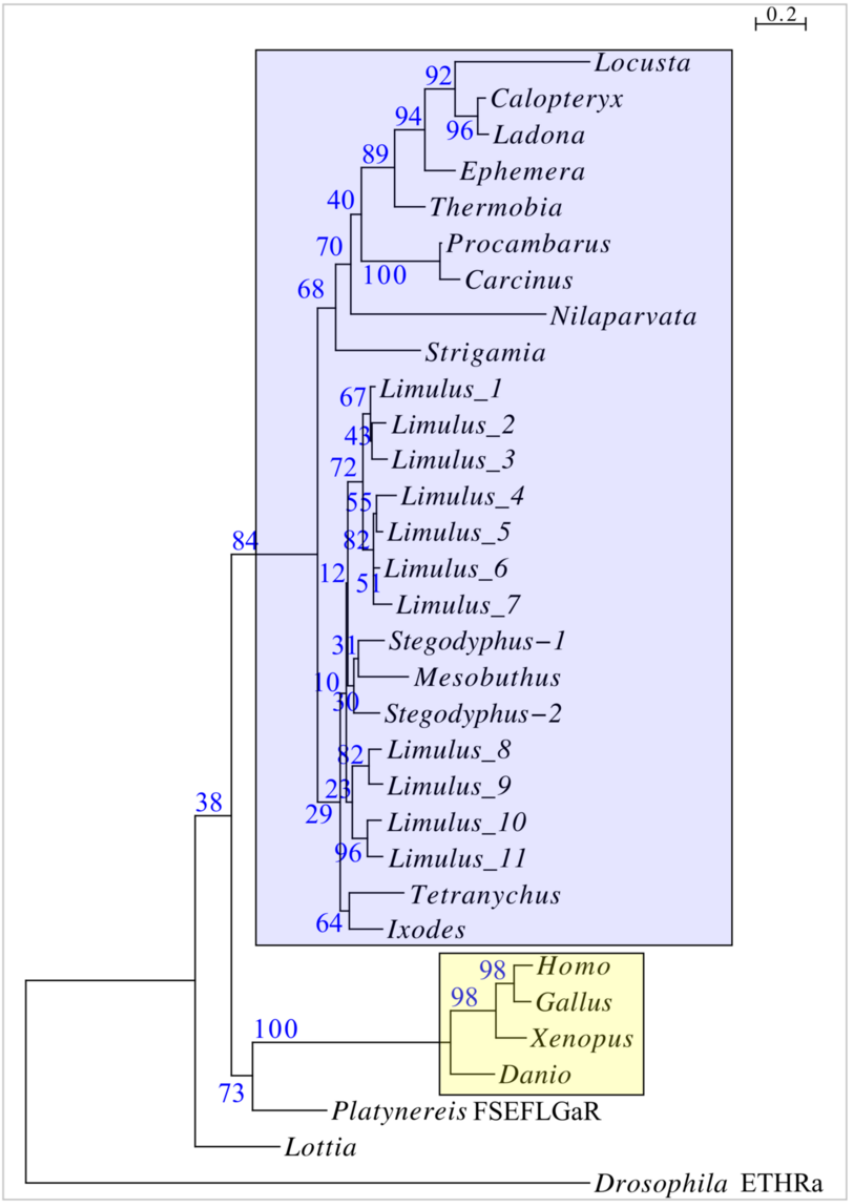

Dii

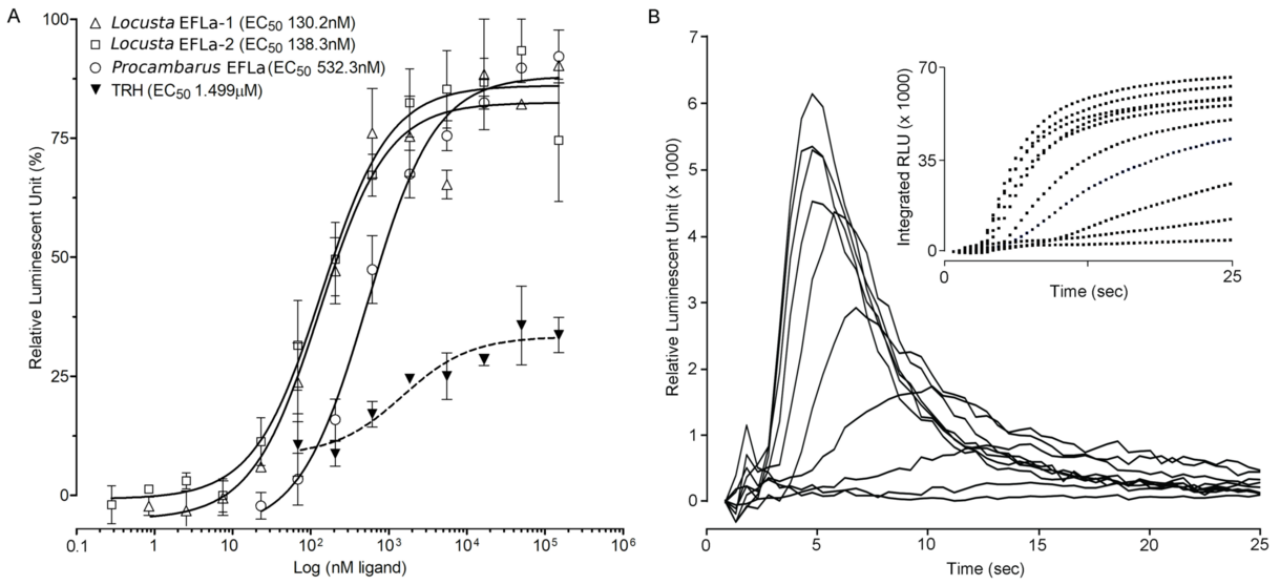

Diii

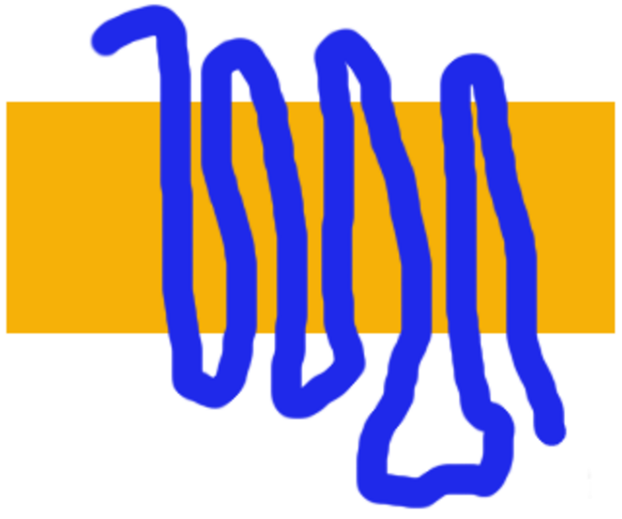

Ei

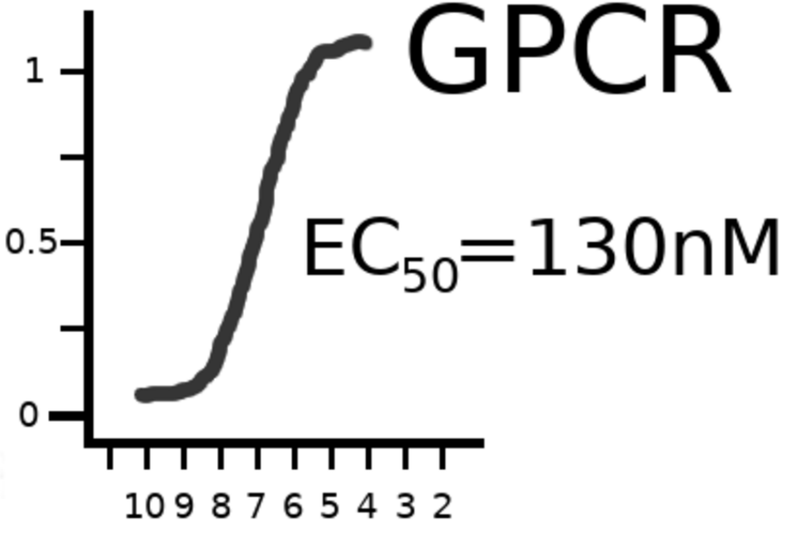

Eii


Eiii


0.2
Locusta
Calopteryx
Ladona
Ephemera
Thermobia
Procambarus
Carcinus
Nilaparvata
Strigamia
Limulus_1
Limulus_2
Limulus_3
Limulus_4
Limulus_5
Limulus_6
Limulus_7
Stegodyphus−1
Mesobuthus
Stegodyphus−2
Limulus_8
Limulus_9
Limulus_10
Limulus_11
Tetranychus
Ixodes
Homo
Gallus
Xenopus
Danio
Platynereis FSEFLGaR
Lottia
Drosophila ETHRa
92
94
96
89
40
70
100
68
67
43
72
55
82
84
12
51
31
10
30
82
23
38
29
96
64
98
98
100
73


A
Locusta EFLa-1 (EC50 130.2nM)
Locusta EFLa-2 (EC50 138.3nM)
Procambarus EFLa (EC50 532.3nM)
TRH (EC50 1.499μM)
Relative Luminescent Unit (%)
Log (nM ligand)
B
Relative Luminescent Unit (x 1000)
Time (sec)
Integrated RLU (x 1000)
Time (sec)





EFLamide