Biochemical and Biophysical Research Communications 404 (2011) 184–189

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

The *Drosophila* genes CG14593 and CG30106 code for G-protein-coupled receptors specifically activated by the neuropeptides CCHamide-1 and CCHamide-2 ☆

Karina K. Hansen, Frank Hauser, Michael Williamson, Stine B. Weber [1], Cornelis J.P. Grimmelikhuijzen *

Center for Functional and Comparative Insect Genomics, Department of Biology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark

ARTICLE INFO

Article history:
Received 18 November 2010
Available online 24 November 2010

Keywords:
GPCR
Neuropeptide
Genomics
Drosophila
Insects
Arthropods

ABSTRACT

Recently, a novel neuropeptide, CCHamide, was discovered in the silkworm *Bombyx mori* (L. Roller et al., Insect Biochem. Mol. Biol. 38 (2008) 1147–1157). We have now found that all insects with a sequenced genome have two genes, each coding for a different CCHamide, CCHamide-1 and -2. We have also cloned and deorphanized two *Drosophila* G-protein-coupled receptors (GPCRs) coded for by genes CG14593 and CG30106 that are selectively activated by *Drosophila* CCH-amide-1 (EC_{50}, 2×10^{-9} M) and CCH-amide-2 (EC_{50}, 5×10^{-9} M), respectively. Gene CG30106 (symbol synonym CG14484) has in a previous publication (E.C. Johnson et al., J. Biol. Chem. 278 (2003) 52172–52178) been wrongly assigned to code for an allatostatin-B receptor. This conclusion is based on our findings that the allatostatins-B do not activate the CG30106 receptor and on the recent findings from other research groups that the allatostatins-B activate an unrelated GPCR coded for by gene CG16752. Comparative genomics suggests that a duplication of the CCHamide neuropeptide signalling system occurred after the split of crustaceans and insects, about 410 million years ago, because only one CCHamide neuropeptide gene is found in the water flea *Daphnia pulex* (Crustacea) and the tick *Ixodes scapularis* (Chelicerata).

1. Introduction

Insects comprise more than 80% of all animal species. They play an extremely important role in ecology and agriculture, because most flowering plants depend on insects for their pollination. But insects can also be agricultural pests, destroying about 30% of our potential annual harvest and can be vectors (intermediate pathogen carriers), transmitting serious diseases, such as malaria, sleeping disease, Dengue fever, yellow fever, West Nile encephalitis, Chagas disease, elephantiasis, and many others. Because insects are so important, about 30 insect genomes have been sequenced during the last 10 years [1–5]. This large number of newly sequenced insect genomes represents an invaluable resource for insect biologists, as it enables them to perform detailed comparative molecular studies on all aspects of insect physiology.

Recently, when analyzing the sequenced genome from the silkworm *Bombyx mori*, Roller and coworkers [6] discovered the gene for a novel insect neuropeptide, CCHamide. This peptide is 13 amino acid residues long, containing two cysteines (forming a cystine bond) and a C-terminal histidine-amide group. In *B. mori* the CCHamide gene is expressed in several dispersed neurons of the central nervous system and in endocrine cells of the gut [6], but otherwise nothing is known about the biological activities of this novel brain-gut neuropeptide. Roller and coworkers [6] found that the CCHamide gene is also present in mosquitoes (one copy), the honey bee *Apis mellifera* (2 copies), the red flour beetle *Tribolium castaneum* (one copy), and the fruitfly *Drosophila melanogaster* (2 copies).

In our present paper, we find that all insects with a sequenced genome contain two copies of the CCHamide gene, each coding for a specific CCHamide (CCHamide-1 or -2). Furthermore, we identify for the first time the receptors for the CCHamides. In *Drosophila* these receptors are specifically activated by either CCHamide-1 or CCHamide-2, suggesting different functions for the two novel neuropeptide signalling systems.

2. Materials and methods

The cDNAs from the *Drosophila* genes CG30106 (symbol synonym CG14484) and CG14593 were recloned based on our GenBank database submissions with accession numbers AY282787, and AY842252, using standard procedures and subsequently recloned into the vector pIRES2-EGFP (Clontech, Mountain View, USA).

☆ The sequence data of this paper have been submitted to the GenBank database under Accession Numbers AY282787 and AY842252.

* Corresponding author. Fax: +45 3532 1200.
E-mail address: cgrimmelikhuijzen@bio.ku.dk (C.J.P. Grimmelikhuijzen).
URL: http://www2.bio.ku.dk/insect_genomics/ (C.J.P. Grimmelikhuijzen).

[1] Present address: Department of Clinical Biochemistry, Hillerød Hospital, Dyrehavevej 29, DK-3400 Hillerød, Denmark (until 2009 her name was Stine B. Frederiksen).

0006-291X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2010.11.089

Chinese hamster ovary (CHO) cells stably expressing the human G-protein G16 (CHO/G16) were grown as described previously [7] and transfected with the vector using JetPEI™ (Polyplus, Illkirch, France). The bioluminescence assay was performed as described earlier [8–10] with the only difference that JetPEI™ was used as transfection medium. We tested a library of eight biogenic amines and 25 *Drosophila* neuropeptides (Supporting Information, Table S1 and the novel *Drosophila* neuropeptides CCHamide-1 and -2 (synthesized by Genemed Synthesis, San Antonio, USA).

TBLASTN homology searches were performed in the nucleotide collection, nonhuman nonmouse expressed sequence tags (EST), or whole-genome shotgun reads (WGS) databases at NCBI (http://blast.bcbi.nlm.nih.gov/Blast.cgi), VectorBase (http://www.vectorbase.org/Tools/BLAST/) and wfleabase (http://wfleabase.org/blast/) using known CCHamide sequences. Signal peptides were predicted using the SignalP server (http://www.cbs.dtu.dk/services/SignalP/). Multiple sequence alignments were performed in ClustalW http://www.ebi.ac.uk/Tools/clustalw/), phylogenetic analysis was done

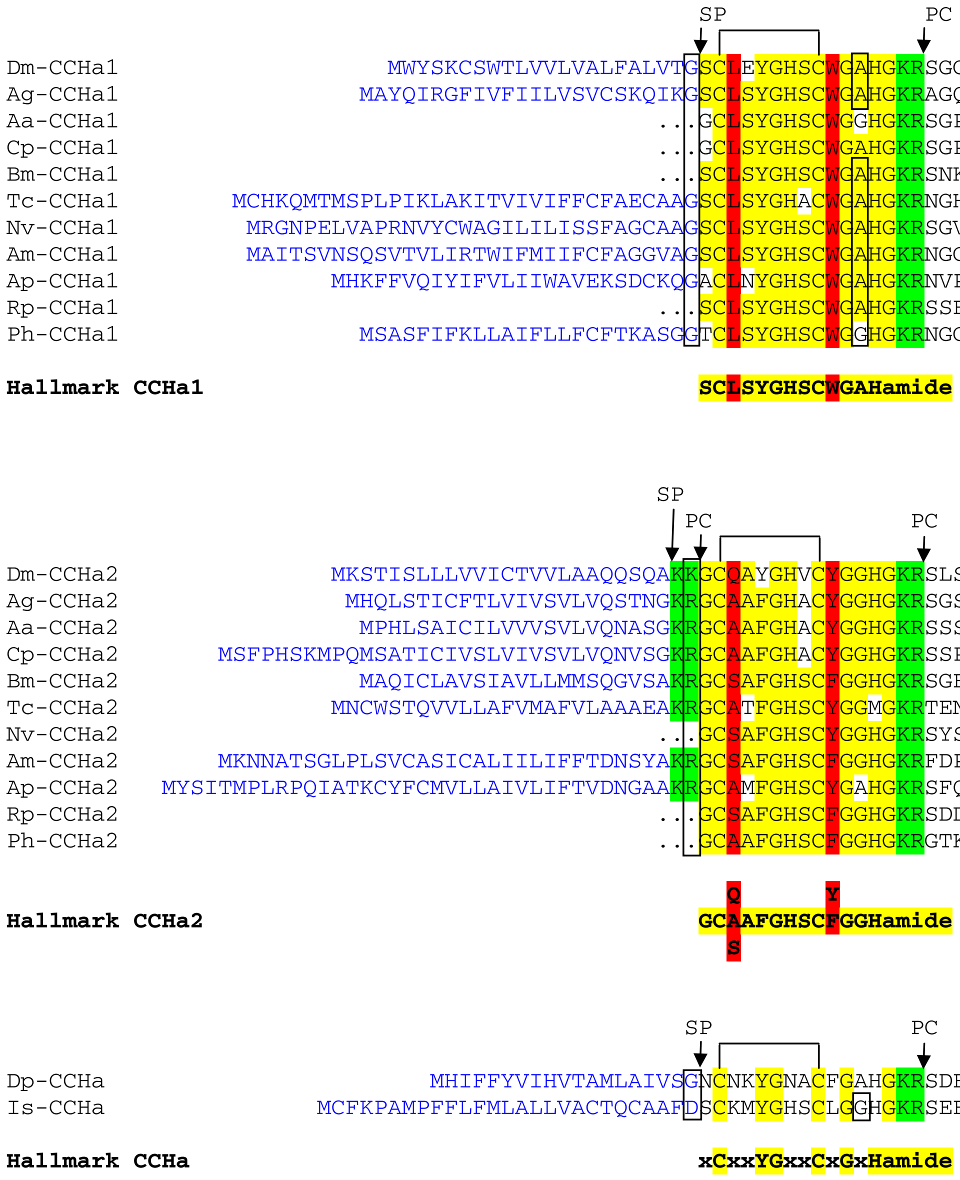

Fig. 1. Alignment of CCHamide preprohormone sequences from *Drosophila melanogaster* (Dm-CCHa1, NP_001097784; Dm-CCHa2, AAL90314), *Anopheles gambiae* (Ag-CCHa1, from EST BM585352 and WGS ABKP02020978; Ag-CCHa2, XP_001237550), *Aedes aegypti* (Aa-CCHa-1, from WGS AAGE02019353; Aa-CCHa2, XP_001649947), *Culex pipiens* (Cp-CCHa1, from AAWU01008744; Cp-CCHa2, from AAWU01038417), *Bombyx mori* (Bm-CCHa1, from BABH01008847; Bm-CCHa2, NP_001123587), *Nasonia vitripennis* (Nv-CCHa1, from WGS AAZX01006578; Nv-CCHa2, from WGS AAZX01016958), *Apis mellifera* (Am-CCHa1, XP_625263; Am-CCHa2, XP_001120020), *Tribolium castaneum* (Tc-CCHa1, from AAJJ01001703.1; Tc-CCHa2, from AAJJ01000164), *Acyrthosiphon pisum* (Ap-CCHa1, from WGS ABLF02035218; Ap-CCHa2, from WGS ABLF02028552), *Rhodnius proxilus* (Rp-CCHa1, from WGS ACPB01040406; Rp-CCHa2, from WGS ACPB01045070), *Pedunculus humanus* (Ph-CCHa1, from WGS DS235222; Ph-CCHa2, from WGS DS235745), *Daphnia pulex* (Dp-CCHa, from EST FE354142 and WGS scaffold_25) and *Ixodes scapularis* (Is-CCHa, from EST EL516199 and WGSs ABJB011135768, ABJB010914573 and ABJB011109090). The complete annotated preprohormone sequences could not be given in this figure, due to space constraints, but are listed in Supporting Information, Fig. S2. Furthermore, to keep this figure at a reasonable size, we have omitted the CCHamide preprohormone sequences from eleven other *Drosophila* species and two other *Nasonia* species with a sequenced genome [3]. Vertical boxes represent introns. The predicted signal peptides are highlighted in blue. The amino acid residues present in more than 60% of CCHamide peptides in each block are marked with yellow (or red) background. The residues highlighted with red background are the residues that can be used to identify whether a processed peptide is CCHa1 or CCHa2. Green background indicates putative cleavage sites for prohormone convertase (PC). The arrows indicate the cleavage positions of signal peptidase (SP) or PC. The cystine bridge formed in each CCHa is indicated above the aligned sequences.

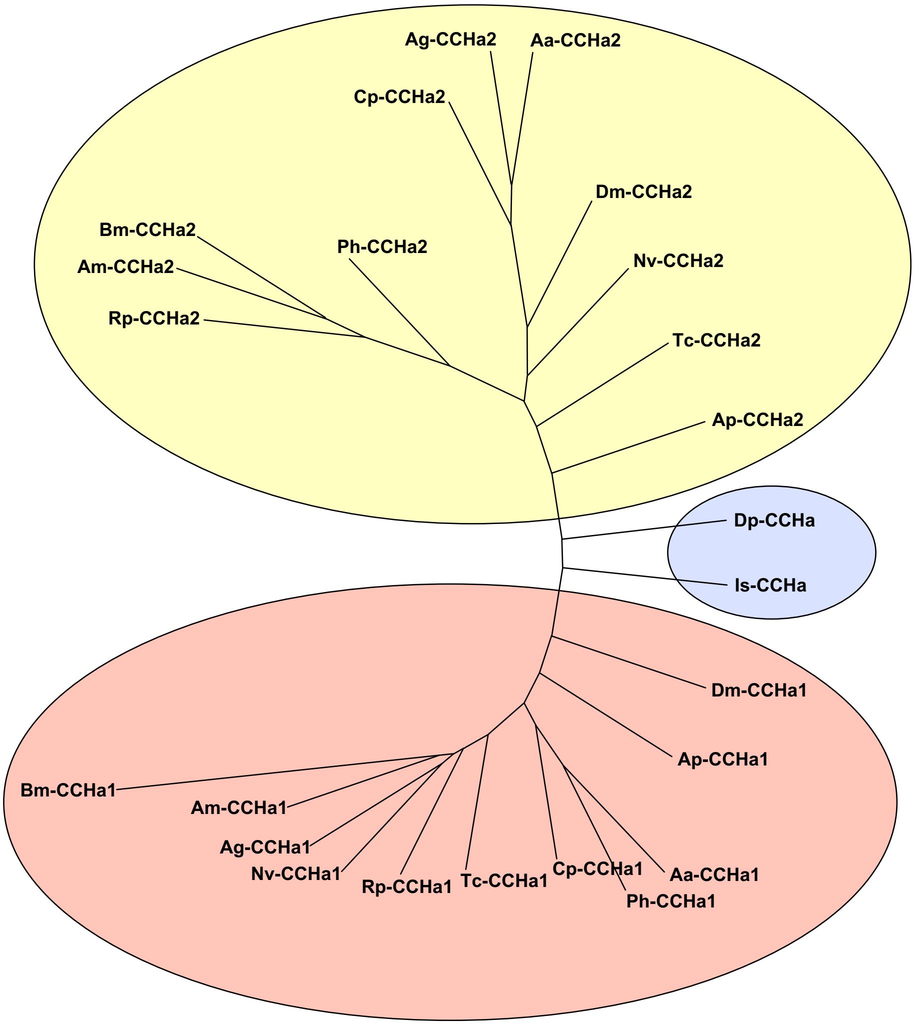

Fig. 2. Phylogenetic tree analysis of various CCHamide peptide sequences. The sequences used are the processed peptide sequences taken from Fig. 1. It can clearly be seen that CCHamides-1 and -2 form distinct clusters and that the CCHamides from *D. pulex* and *I. scapularis* are at a position intermediate to these clusters.

using the programs Phylip protdist (http://mobyle.pasteur.fr/cgi-bin/portal.py?form=protdist) and Phylodendron (http://iubio.bio.indiana.edu/treeapp/treeprint-form.html).

3. Results and discussion

3.1. Identification of arthropod CCHamide genes

We investigated most insects with a sequenced genome and found that all of them have two CCHamide genes, coding for two different CCHamides (Fig. 1). The genomic organizations of the two genes differ, because the CCHamide-1 genes have generally two introns in their coding regions, whereas the CCHamide-2 genes have only one (Fig. 1). The two CCHamide genes code for two CCHamide preprohormones, which, again, are organized in a different way. In the prepro-CCHamides-1 the signal peptides are immediately preceding the active CCHamide-1 sequences, indicating that the signal peptidase liberates the N-termini of the CCHamides-1, whereas in the prepro-CCHamides-2 there are dibasic cleavage sites between the signal peptides and the active CCHamide-2 sequences, suggesting that a prohormone convertase liberates the N-termini of the CCHamides-2 (Fig. 1). In both the CCHamide-1 and -2 preprohormones a dibasic cleavage site liberates the carboxytermini of the CCHamides, a process catalyzed again by prohormone convertase, after which the residual carboxyterminal basic amino acid residues are removed by a specific carboxypeptidase. Finally, the resulting C-terminal glycine residues of the immature CCHamides are converted into amide groups by a peptidylglycine alpha-amidating monooxygenase (Fig. 1) [11].

Also at the peptide level, the insect CCHamides can clearly be assigned to either the CCHamide-1 or -2 groups. CCHamides-1 have always a leucine residue in position 3 and a tryptophan in position 10; CCHamides-2 always have glutamine/alanine/serine residues in position 3 and tyrosine/phenylalanine residues in position 10 (marked red in Fig. 1).

In addition to insects, we also found CCHamides in other arthropods. In the water flea *Daphnia pulex* (Crustacea) and the tick *Ixodes scapularis* (Chelicerata) we found that only one CCHamide gene is present in each species. Both the genomic organization and the structure of its preprohormone would classify the CCHamide gene from *I. scapularis* as a CCHamide-1 gene (Fig. 1). For the *D. pulex* gene, the situation is less clear. The absence of an N-terminal dibasic cleavage site in its preprohormone would classify it as a CCHamide-1 gene; the absence of a second intron, however, would assign it as a CCHamide-2 gene (Fig. 1). At the peptide level the *I. scapularis* and *D. pulex* CCHamides have mixed properties, resembling both insect CCHamides-1 and -2 (Fig. 1).

A phylogenetic tree analysis of the amino acid sequences of processed arthropod CCHamides supports our classification of these peptides in CCHamides-1 and -2 (Fig. 2). Again, the *I. scapularis* and *D. pulex* CCHamides lie at an intermediate position between both groups (Fig. 2). An additional phylogenetic tree analysis of

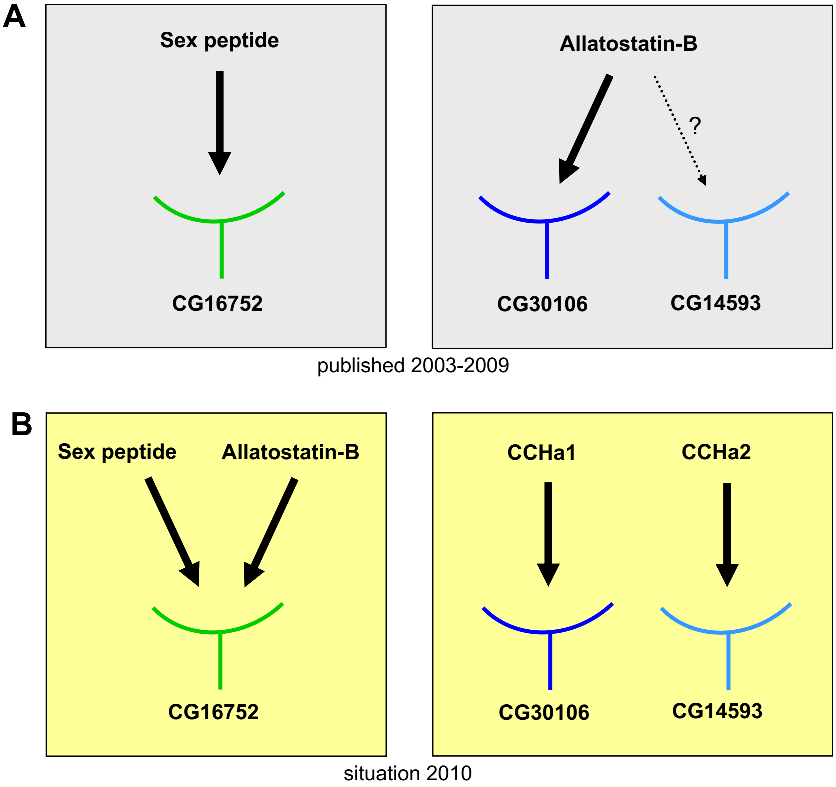

Fig. 3. Schematic drawing of the published ligand specificities of the *Drosophila* receptors CG16752, CG30106 (symbol synonym CG14484), and CG14593. (A) As published 2003–2009. (B) The current situation. Note that CG30106 and CG14593 are closely related genes (paralogues), whereas CG16752 is quite distant. For further information please see the text.

the arthropod CCHamide preprohormones (instead of processed CCHamides) gives a similar picture (Supporting Information, Fig. S1). These phylogenetic data, together with our findings that only one chelicerate and one crustacean CCHamide gene could be identified, suggest that, during evolution, a CCHamide gene duplication event has occurred after the split of Crustacea and Insecta about 410 million years (MYR) ago [12].

3.2. Deorphanizing the CCHamide receptors in Drosophila

Already in 2002 we cloned two orphan G protein-coupled receptors (GPRCs) from *Drosophila*, one coded for by gene CG30106 (GenBank accession no. AY842252) and the other coded for by gene CG14593 (GenBank accession no. AY282787). In 2002, we expressed these GPCRs in Chinese hamster ovary (CHO) cells in cell culture, but were unable to find the ligands for these receptors, using the insect neuropeptide library that was available during that time [13]. In the following year Johnson and coworkers published that the GPCR coded for by CG30106 (symbol synonym CG14484) was an allatostatin-B receptor [14]. We could not confirm these results, but because our bioassay system was different from the one used by Johnson and coworkers [14], we imagined that these differences must be due to the different cell types and assay conditions used by the two groups. Because CG30106 and CG14593 are two paralogue genes, we assumed that also CG14593 could code for an allatostatin-B receptor (stipled line and question mark drawn in Fig. 3A) [2,15].

Seven years later, in 2010, several research groups [16–18] reported that a completely unrelated *Drosophila* gene, CG16752 (Fig. 3), which was previously identified as the sex peptide receptor [19], codes for an allatostatin-B receptor. These results suggested to us that the CG30106 receptor was perhaps identified wrongly as an allatostatin-B receptor by Johnson et al. [14] and that the CG30106 and CG14593 receptors might indeed still be orphans. This idea prompted us to reinvestigate the two receptors and test them with recently discovered *Drosophila* neuropeptides not known in 2002.

Fig. 4B shows that the CG30106 receptor is activated by nanomolar concentrations of CCHamide-1 (EC_{50}, 2×10^{-9} M). Also CCHamide-2 activated this receptor (now called CCHamide-1 receptor), but about 5000 times higher concentrations than CCHamide-1 were needed for this effect (Fig. 4A and B). As we found earlier, the allatostatins-B did not activate the CCHamide-1 receptor. Also a library of more than twenty other *Drosophila* neuropeptides (Supporting Information, Table S1) did not activate the receptor with the exception of sex peptide, which, however, only had a very minor effect and only at concentrations 100,000 times higher than CCHamide-1 (Fig. 4B). We, thus, regard the effect of sex peptide on this receptor as physiologically irrelevant.

In contrast, Fig. 4D shows that the CG14593 receptor is activated by nanomolar concentrations of CCHamide-2 (EC_{50}, 5×10^{-9} M). CCHamide-1 is also activating this receptor (now called CCHamide-2 receptor), but about 2000 times higher concentrations were needed for this effect (Fig. 4C and D). No other known *Drosophila* neuropeptide activated the receptor, including allatostatins-B and sex peptide (Fig. 4D).

4. Conclusions

The recently discovered insect brain-gut neuropeptide CCHamide [6] is present in all insects with a sequenced genome, where it occurs in two discrete forms, CCHamide-1 and -2, each coded for by a separate gene (Figs. 1 and 2). In other arthropods with a sequenced genome, such as the water flea *D. pulex* (Crustacea), and

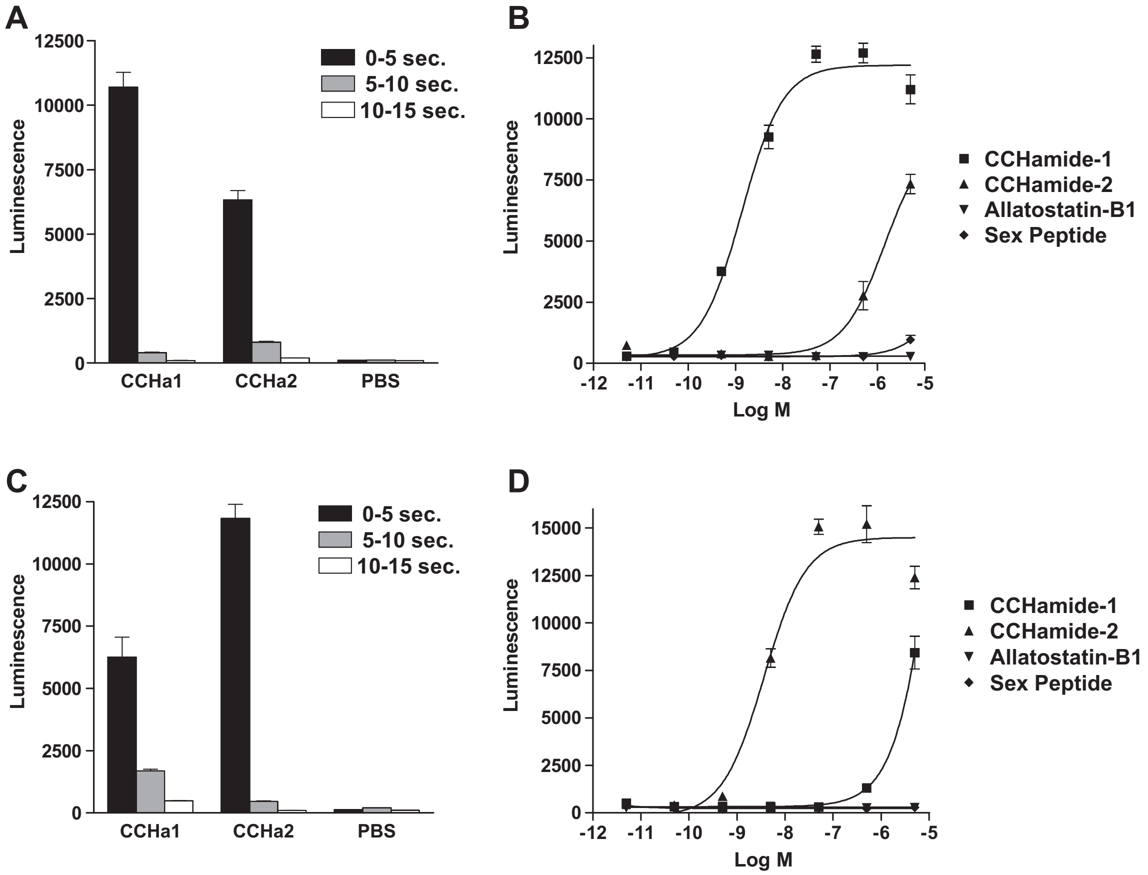

Fig. 4. Bioluminescence responses of cloned CHO/G-16 cell lines transfected with DNA coding for either the *Drosophila melanogastor* CCHamide-1 receptor (CG30106) or the CCHamide-2 receptor (CG14593). The vertical bars represent SEM ($n = 3$ in (A); $n = 5$ in (B); $n = 3$ in (C); $n = 5$ in (D)), which sometimes are smaller than the symbols (or lines) used. In these cases, only the symbols (or lines) are given. (A) Bioluminescence response of CHO/G-16/CG30106 cells after addition of Drm-CCHamide-1 (final concentration, 5×10^{-6} M), Drm-CCHamide-2 (final concentration, 5×10^{-6} M) or phosphate-buffered saline (PBS). (B) Dose–response curves of the effects of CCHamide-1, CCHamide-2, allatostatin-B1 and sex peptide on these cells. The transfected cells were only activated by CCHamide-1 (and to a much lower degree also by CCHamide-2). (C) Bioluminescence response of CHO/G-16/CG14593 cells after addition of Drm-CCHamide-1 (final concentration, 5×10^{-6} M), Drm-CCHamide-2 (final concentration, 5×10^{-6} M) or phosphate-buffered saline (PBS). (D) Dose–response curves of the effects of CCHamide-1, CCHamide-2, allatostatin-B1 and sex peptide on these cells. The transfected cells were only activated by CCHamide-2 (and to a much lower degree also by CCHamide-1). The CHO/G-16/CG30106 cells and CHO/G-16/CG14593 cells did not react with more than twenty other *Drosophila* neuropeptides and eight biogenic amines mentioned in Supporting Information, Table S1.

the tick *I. scapularis* (Chelicerata), only one CCHamide gene was identified (Figs. 1 and 2), suggesting that a CCHamide gene duplication occurred after the split of crustaceans and insects, about 410 MYR ago [12]. In accordance with the presence of two different CCHamide peptides in *Drosophila*, we could identify two different CCHamide receptors in this fruitfly coded for by the two paralogue genes, CG30106 and CG14593. The two receptors are specifically activated by either CCHamide-1 (CG30106), or CCHamide-2 (CG14593) (Figs. 3 and 4). We have, therefore, identified two novel, separate neuropeptide signalling systems in *Drosophila* and their presence in *Drosophila* and other insects with a sequenced genome, will enable us to elucidate their so far unknown functions, using knock-out mutants or RNAi.

Acknowledgments

We thank Anders Bo Rønnegaard Hansen for typing the manuscript and the Danish Research Agency, and Novo Nordisk Foundation for financial support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2010.11.089.

References

[1] C.J.P. Grimmelikhuijzen, G. Cazzamali, M. Williamson, et al., Perspective. The promise of insect genomics, Pest. Manag. Sci. 63 (2007) 413–416.
[2] F. Hauser, G. Cazzamali, M. Williamson, et al., A genome-wide inventory of neurohormone GPCRs in the red flour beetle *Tribolium castaneum*, Front. Neuroendocrinol. 29 (2008) 142–165.
[3] J.H. Werren, S. Richards, C.A. Desjardins, et al., Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species, Science 327 (2010) 343–348.
[4] S. Richards, R.A. Gibbs, N.M. Gerardo, et al., Genome sequence of the pea aphid *Acyrthosiphon pisum*, PLoS Biol. 8 (2010) e1000313.
[5] P. Arensburger, K. Megy, R.M. Waterhouse, et al., Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics, Science 330 (2010) 86–88.
[6] L. Roller, N. Yamanaka, K. Watanabe, et al., The unique evolution of neuropeptide genes in the silkworm *Bombyx mori*, Insect Biochem. Mol. Biol. 38 (2008) 1147–1157.
[7] J. Stables, A. Green, F. Marshall, et al., A bioluminescent assay for agonist activity at potentially any G-protein-coupled receptor, Anal. Biochem. 252 (1997) 115–126.
[8] T. Secher, C. Lenz, G. Cazzamali, et al., Molecular cloning of a functional allatostatin gut/brain receptor and an allatostatin preprohormone from the silkworm *Bombyx mori*, J. Biol. Chem. 276 (2001) 47052–47060.
[9] F. Staubli, T.J.D. Jørgensen, G. Cazzamali, et al., Molecular identification of the insect adipokinetic hormone receptors, Proc. Natl. Acad. Sci. USA 99 (2002) 3446–3451.
[10] K.K. Hansen, E. Stafflinger, M. Schneider, et al., Discovery of a novel insect neuropeptide signaling system closely related to the insect adipokinetic hormone and corazonin hormonal systems, J. Biol. Chem. 285 (2010) 10736–10747.

[11] B.A. Eipper, D.A. Stoffers, R.E. Mains, The biosynthesis of neuropeptides: peptide α-amidation, Ann. Rev. Neurosci. 15 (1992) 57–85.

[12] H. Glenner, P.F. Thomsen, M.B. Hebsgaard, et al., Evolution. The origin of insects, Science 314 (2006) 1883–1884.

[13] S.B. Frederiksen, Molecular cloning of three G protein-coupled receptors in *Drosophila melanogaster*, Master Degree Thesis in Biochemistry, University of Copenhagen, 2004.

[14] E.C. Johnson, L.M. Bohn, L.S. Barak, et al., Identification of *Drosophila* neuropeptide receptors by G protein-coupled receptors-beta-arrestin2 interactions, J. Biol. Chem. 278 (2003) 52172–52178.

[15] F. Hauser, G. Cazzamali, M. Williamson, et al., A review of neurohormone GPCRs present in the fruitfly *Drosophila melanogaster* and the honey bee *Apis mellifera*, Prog. Neurobiol. 80 (2006) 1–19.

[16] Y.J. Kim, K. Bartalska, N. Audsley, et al., MIPs are ancestral ligands for the sex peptide receptor, Proc. Natl. Acad. Sci. USA 107 (2010) 6520–6525.

[17] N. Yamanaka, Y.J. Hua, L. Roller, et al., Bombyx prothoracicostatic peptides activate the sex peptide receptor to regulate ecdysteroid biosynthesis, Proc. Natl. Acad. Sci. USA 107 (2010) 2060–2065.

[18] J. Poels, T. Van Loy, H.P. Vandersmissen, et al., Myoinhibiting peptides are the ancestral ligands of the promiscuous *Drosophila* sex peptide receptor, Cell Mol. Life Sci. 67 (2010) 3511–3522.

[19] N. Yapici, Y.J. Kim, C. Ribeiro, et al., A receptor that mediates the post-mating switch in *Drosophila* reproductive behaviour, Nature 451 (2008) 33–37.