



ACADEMIC
PRESS

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Biochemical and Biophysical Research Communications 303 (2003) 146–152

BBRC

www.elsevier.com/locate/ybbrc

Molecular identification of a *Drosophila* G protein-coupled receptor specific for crustacean cardioactive peptide[☆]

Giuseppe Cazzamali, Frank Hauser, Sune Kobberup, Michael Williamson,
and Cornelis J.P. Grimmelikhuijzen*

Department of Cell Biology, Zoological Institute, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark

Received 12 February 2003

Abstract

The *Drosophila* Genome Project website (www.flybase.org) contains the sequence of an annotated gene (CG6111) expected to code for a G protein-coupled receptor. We have cloned this receptor and found that its gene was not correctly predicted, because an annotated neighbouring gene (CG14547) was also part of the receptor gene. DNA corresponding to the corrected gene CG6111 was expressed in Chinese hamster ovary cells, where it was found to code for a receptor that could be activated by low concentrations of crustacean cardioactive peptide, which is a neuropeptide also known to occur in *Drosophila* and other insects (EC₅₀, 5.4×10^{-10} M). Other known *Drosophila* neuropeptides, such as adipokinetic hormone, did not activate the receptor. The receptor is expressed in all developmental stages from *Drosophila*, but only very weakly in larvae. In adult flies, the receptor is mainly expressed in the head. Furthermore, we identified a gene sequence in the genomic database from the malaria mosquito *Anopheles gambiae* that very likely codes for a crustacean cardioactive peptide receptor.

© 2003 Elsevier Science (USA). All rights reserved.

The recent sequencing of the *Drosophila* genome by the *Drosophila* Genome Project consortium [1] represents a milestone in insect research, because it enables researchers to identify all proteins in *Drosophila* and, thereby, to understand an insect at all possible levels (molecular, cellular, organismic, behavioural, etc.). Furthermore, the new knowledge that we will gain from *Drosophila* will certainly also have a strong impact on our understanding of other insects, other arthropods, or even other invertebrates. Our research group is particularly interested in neuropeptide receptors and their ligands (the neuropeptides), because these proteins/peptides occupy a high hierarchic position in the physiology of an insect and steer important processes such as reproduction, development, feeding, and all types of behaviours.

The website of the *Drosophila* Genome Project consortium (www.flybase.org) contains a list of 40–45 annotated (i.e., hypothetical) neuropeptide G protein-coupled receptors. We have earlier used this list as a starting point to clone various *Drosophila* neuropeptide receptors [2–5], but we have often found that the annotations were not correct. Furthermore, the ligands for the annotated receptors are unknown, i.e., they are orphan receptors. Therefore, proper cloning of the annotated receptors, expression in cells, and identification of their cognate ligands are still necessary processes.

In the present paper we have cloned the receptor encoded by the annotated gene CG6111, subsequently corrected its annotation (gene structure), and identified its cognate ligand. We have chosen gene CG6111, because its putative gene product is structurally somewhat related to the *Drosophila* adipokinetic hormone (Drm-AKH) and *Drosophila* corazonin (Drm-corazonin) receptors [3,6]. This would open the possibility for the existence of a second Drm-AKH or -corazonin receptor. We found, however, that the receptor was highly specific

[☆] The nucleotide sequence reported in this paper has been submitted to the GenBank Data Bank with Accession No. [AY219842](http://www.ncbi.nlm.nih.gov/nuclot/AY219842).

* Corresponding author. Fax: +45-35321200.

E-mail address: cgrimmelikhuijzen@zi.ku.dk (C.J.P. Grimmelikhuijzen).

URL: www.zi.ku.dk/cellbiology/.

TGCCAGACTGAGGCGCAATC																								-1		
ATG	CTG	CAC	TTA	AGG	CTA	TTC	GAT	AGC	TCA	CTC	TAC	TAC	ACC	CTG	GCC	TCC	GCC	TCC	GAA	TCC	TCG	GGG	CTG	GCG	TCC	78
Met	Leu	His	Leu	Arg	Leu	Phe	Asp	Ser	Ser	Leu	Tyr	Tyr	Thr	Leu	Ala	Ser	Ala	Ser	Glu	Ser	Ser	Gly	Leu	Ala	Ser	26
TCC	ACC	TCC	ACG	GAA	AGG	TCG	TTC	AAT	GGC	ACC	CAG	GGG	GCG	GGG	GGC	GTG	GCC	GCT	GGC	GGC	GAG	TCC	CTA	ACG	CCA	156
Ser	Thr	Ser	Thr	Glu	Arg	Ser	Phe	Asn	Gly	Thr	Gln	Gly	Ala	Gly	Gly	Val	Ala	Ala	Gly	Gly	Glu	Ser	Leu	Thr	Pro	52
ACC	GAC	GTG	GCT	GCA	GTT	AAT	CTG	ACA	TAT	TTC	ACA	CCT	GCG	ATA	TCA	CAC	GTG	ATG	CTG	GCG	CCA	ACA	ACG	ATA	GCA	234
Thr	Asp	Val	Ala	Ala	Val	Asn	Leu	Thr	Tyr	Phe	Thr	Pro	Ala	Ile	Ser	His	Val	Met	Leu	Ala	Pro	Thr	Thr	Ile	Ala	78
ACC	ACG	ACA	GCC	TCG	GCC	ACA	ATG	GTC	CAG	ATC	CAG	ACG	ACA	GCG	GCG	CCA	AGT	CAC	GAC	TTG	GAG	ACG	GGC	GGC	AAC	312
Thr	Thr	Thr	Ala	Ser	Ala	Thr	Met	Val	Gln	Ile	Gln	Thr	Thr	Ala	Ala	Pro	Ser	His	Asp	Leu	Glu	Thr	Gly	Gly	Asn	104
TCC	ACA	TCC	AGC	GAC	CCC	GGC	GAA	TTT	GAC	AAT	TTA	AAT	TCG	TTC	TAC	TTT	TAT	GAG	ACC	GAA	CAG	TTT	GCT	GTG	CTC	390
Ser	Thr	Ser	Ser	Asp	Pro	Gly	Glu	Phe	Asp	Asn	Leu	Asn	Ser	Phe	Tyr	Phe	Tyr	Glu	Thr	Glu	Gln	Phe	Ala	Val	Leu	130
TMI																										
TGG	ATC	CTG	TTC	ACC	GTC	ATC	GTT	CTG	GGC	AAT	TCA	GCT	GTT	CTG	TTC	GTG	ATG	TTC	ATC	AAC	AAG	AAT	CGC	AAG	TCG	468
Trp	Ile	Leu	Phe	Thr	Val	Ile	Val	Leu	Gly	Asn	Ser	Ala	Val	Leu	Phe	Val	Met	Phe	Ile	Asn	Lys	Asn	Arg	Lys	Ser	156
CGG	ATG	AAC	TAC	TTC	ATT	AAA	CAG	CTG	GCA	TTG	GCA	GAT	CTG	TGC	GTG	GGA	CTG	CTC	AAC	GTC	CTC	ACC	GAC	ATC	ATA	546
Arg	Met	Asn	Tyr	Phe	Ile	Lys	Gln	Leu	Ala	Leu	Ala	Asp	Leu	Cys	Val	Gly	Leu	Leu	Asn	Val	Leu	Thr	Asp	Ile	Ile	182
TMII																										
TGG	CGC	ATC	ACG	ATT	TCG	TGG	CGG	GCA	GGC	AAC	CTG	GCC	TGC	AAG	GCC	ATC	CGC	TTC	TCG	CAG	GTC	TGC	GTC	ACA	TAC	624
Trp	Arg	Ile	Thr	Ile	Ser	Trp	Arg	Ala	Gly	Asn	Leu	Ala	Cys	Lys	Ala	Ile	Arg	Phe	Ser	Gln	Val	Cys	Val	Thr	Tyr	208
TCG	TCC	ACC	TAC	GTG	CTG	GTG	GCC	ATG	AGC	ATC	GAC	AGA	TAC	GAT	GCC	ATC	ACA	CAC	CCC	ATG	AAC	TTC	TCA	AAG	TCG	702
Ser	Ser	Thr	Tyr	Val	Leu	Val	Ala	Met	Ser	Ile	Asp	Arg	Tyr	Asp	Ala	Ile	Thr	His	Pro	Met	Asn	Phe	Ser	Lys	Ser	234
TMIV																										
TGG	AAA	AGA	GCC	CGT	CAC	CTG	GTG	GCT	GGC	GCA	TGG	CTC	ATC	TCG	GCG	TTG	TTT	TCG	CTT	CCC	ATC	CTG	GTT	TTG	TAC	780
Trp	Lys	Arg	Ala	Arg	His	Leu	Val	Ala	Gly	Ala	Trp	Leu	Ile	Ser	Ala	Leu	Phe	Ser	Leu	Pro	Ile	Leu	Val	Leu	Tyr	260
GAG	GAG	AAG	CTC	ATC	CAA	GGA	CAT	CCG	CAA	TGC	TGG	ATT	GAG	TTG	GGT	TCA	CCG	ATC	GCC	TGG	CAG	GTA	TAC	ATG	AGC	858
Glu	Glu	Lys	Leu	Ile	Gln	Gly	His	Pro	Gln	Cys	Trp	Ile	Glu	Leu	Gly	Ser	Pro	Ile	Ala	Trp	Gln	Val	Tyr	Met	Ser	286
TMV																										
CTG	GTG	TCG	GCC	ACT	CTA	TTT	GCC	ATT	CCT	GCG	CTG	ATC	ATA	TCT	GCC	TGC	TAT	GCG	ATC	ATC	GTA	AAG	ACG	ATT	TGG	936
Leu	Val	Ser	Ala	Thr	Leu	Phe	Ala	Ile	Pro	Ala	Leu	Ile	Ile	Ser	Ala	Cys	Tyr	Ala	Ile	Ile	Val	Lys	Thr	Ile	Trp	312
GCA	AAG	GGT	TCC	ATT	TTT	GTA	CCC	ACG	GAA	CGT	GCT	GGT	TTT	GGA	GCT	GCA	CCT	GCC	AGG	AGG	GCC	AGC	TCG	AGG	GGC	1014
Ala	Lys	Gly	Ser	Ile	Phe	Val	Pro	Thr	Glu	Arg	Ala	Gly	Phe	Gly	Ala	Ala	Pro	Ala	Arg	Arg	Ala	Ser	Ser	Arg	Gly	338
TMVI																										
ATT	ATT	CCA	CGG	GCA	AAG	GTC	AAA	ACG	GTC	AAG	ATG	ACA	TTG	ACC	ATC	GTG	TTT	GTG	TTC	ATC	ATC	TGC	TGG	TCG	CCG	1092
Ile	Ile	Pro	Arg	Ala	Lys	Val	Lys	Thr	Val	Lys	Met	Thr	Leu	Thr	Ile	Val	Phe	Val	Phe	Ile	Ile	Cys	Trp	Ser	Pro	364
TAT	ATC	ATC	TTC	GAT	CTG	CTG	CAG	GTC	TTT	GGC	CAG	ATT	CCC	CAC	TCA	CAG	ACC	AAC	ATT	GCC	ATC	GCC	ACC	TTC	ATC	1170
Tyr	Ile	Ile	Phe	Asp	Leu	Leu	Gln	Val	Phe	Gly	Gln	Ile	Pro	His	Ser	Gln	Thr	Asn	Ile	Ala	Ile	Ala	Thr	Phe	Ile	390
TMVII																										
CAA	AGT	CTG	GCA	CCG	CTG	AAC	TCG	GCG	GCG	AAT	CCA	CTA	ATC	TAT	TGC	CTC	TTC	TCA	TCG	CAG	GTC	TTT	CGC	ACA	TTA	1248
Gln	Ser	Leu	Ala	Pro	Leu	Asn	Ser	Ala	Ala	Asn	Pro	Leu	Ile	Tyr	Cys	Leu	Phe	Ser	Ser	Gln	Val	Phe	Arg	Thr	Leu	416
AGT	CGC	TTT	CCG	CCT	TTT	AAG	TGG	TTC	ACA	TGC	TGC	TGC	AAG	TCA	TAC	CGC	AAC	AAC	TCG	CAG	CAA	AAC	CGC	TGC	CAC	1326
Ser	Arg	Phe	Pro	Pro	Phe	Lys	Trp	Phe	Thr	Cys	Cys	Cys	Lys	Ser	Tyr	Arg	Asn	Asn	Ser	Gln	Gln	Asn	Arg	Cys	His	442
ACG	GTT	GGT	CGT	CGG	CTT	CAC	AAC	AGT	TGC	GAT	TCG	ATG	AGG	ACA	CTG	ACC	ACT	TCG	TTG	ACG	GTT	TCC	CGA	AGG	TCC	1404
Thr	Val	Gly	Arg	Arg	Leu	His	Asn	Ser	Cys	Asp	Ser	Met	Arg	Thr	Leu	Thr	Thr	Ser	Leu	Thr	Val	Ser	Arg	Arg	Ser	468
ACC	AAC	AAG	ACG	AAC	GCC	CGT	GTG	GTA	ATC	TGC	GAA	CGT	CCC	ACC	AAG	GTG	GTT	ACC	GTG	CCA	GCC	ATG	TCG	GAG	GGA	1482
Thr	Asn	Lys	Thr	Asn	Ala	Arg	Val	Val	Ile	Cys	Glu	Arg	Pro	Thr	Lys	Val	Val	Thr	Val	Pro	Ala	Met	Ser	Glu	Arg	494
CGC	GGA	GTT	TCT	CTA	AAG	GGG	AAC	ACG	GAC	ATC	CTG	TGA	AGGCACCAGAGGCTTAGACAATAAGTTCGATAGTTATACCAGGCATGGGTA	1572												
Arg	Gly	Val	Ser	Leu	Lys	Gly	Asn	Thr	Asp	Ile	Leu	*	506													
AAGCCAAATGTATGTTTCTAGTCGTAAGTACGCCAAGTCTACGCTACGAATTTCAAACCTATTTCTACTTCGCAATCTATAGTTAAGATAACATTTTATACGGC																								1675		
ATTTAATTAACCTCTAAAATTAATAAATCTAAAGTGTCAAAATGGGTGATTGAATCTTTTCAAGATTTTGATTTTATATAAATTTAACTAAAAGAACCCATTATA																								1778		
ATTTGATATTTGGTACAAAGGTTTATATTTCCAAAAAATTTACCTGAAACATTTAATTAGGCATTTTTCGAATCAACTTTAACTGGAGAAATTAATTTAAGT																								1881		
ATATGGTATCCTAATGAAATAAGTGTTCGAAAATATTAAGATTTTGTCCACCACTCAGAACCTAGGTTTCAATCCATTAGTAGCTCGGTTTGTAGTTATAT																								1984		
TCTCTTAAACAGACTGATTTTACTTTGTATGTGAATCGCGCCATTATCAGAAATTTAAATGCAATTTCTTTCCAGTCCCTAAGTACGTTAAGGCCATTTAA																								2087		
ATTCCAATCATCCGTGTATATAGCCACTCGTAATGGTGGATAAACATGCCCAAGCATCTTAAACCAAAACCCATTAAAGCATAAGTAGACCATAAGAGTTAAT																								2190		
AGAGCATAACGTTGTAGTCAATTTGGTTTAAAGCTAATATTTATTTTCATAATATTGTGTGAAAGCAGCAATAAAGGCAACGACTTC(A) _n																								2274		

Fig. 1. cDNA and deduced amino acid sequence of the corrected gene CG6111. Nucleotides are numbered from 5'- to 3'-end and the amino acid residues are numbered starting with the first ATG codon in the open reading frame. The six introns are indicated by arrows and the exon nucleotides, bordering these introns, are highlighted in grey. The seven membrane spanning domains are boxed and labelled TMI-VII. The two putative glycosylation sites in the extracellular N terminus of the protein are indicated by filled triangles. The translation termination codon is indicated by an asterisk. The in-frame stop codon in the 5'-noncoding region is underlined. The putative polyadenylation signal in the 3'-noncoding region is underlined twice.

for crustacean cardioactive peptide, which is a neuropeptide that was first discovered in crustaceans, but that also occurs in identical form in insects [7–9].

Materials and methods

Primers were constructed based on the proposed exons of gene CG6111 (www.flybase.org) and used on cDNA from *D. melanogaster* third instar larvae (Canton S) as a template. The sense primer was 5'-ACCGTCATCGTTCTGGGCAATTCA-3' (corresponding to nucleotide positions 403–426 of Fig. 1) followed by the nested sense primer 5'-GCAAGTCGCGGATGAACTACTTCATTA-3' (corresponding to nucleotide positions 461–487 of Fig. 1) and the antisense primer was 5'-GATGTCCGTGTTCCCTTTAGAGA-3' (corresponding to nucleotide positions 1492–1515 of Fig. 1) followed by the nested antisense primer 5'-GTCCTCATCGAATCGCAACTGTTGTGA-3' (corresponding to nucleotide positions 1344–1369 of Fig. 1). The PCR program was 1 cycle of 95 °C for 3 min, 58 °C for 1 min, and 72 °C for 2 min followed by 35 cycles of 95 °C for 30 s, 58 °C for 1 min, 72 °C for 2 min, and a final extension step of 72 °C for 10 min. These PCRs were carried out using the Advantage2 PCR enzyme system (Clontech). For the amplification of the 3'-UTR the SMART RACE cDNA kit (Clontech) was used. The 3'-RACE reactions were made with the sense primer 5'-GCTTCACAACAGTTGCGATTCGATGAGGA CACTG ACCA-3' (corresponding to nucleotide positions 1341–1378 of Fig. 1) followed by the nested sense primers, 5'-GAAGGTCCACCAACAAGACG AACGCCGTGTGG TA-3' (corresponding to nucleotide positions 1397–1431 of Fig. 1) and 5'-GCCATGTCTCGGAGCGACGCGGA GTTCTCTAA-3' (corresponding to nucleotide positions 1468–1498 of Fig. 1). The PCR program was: 94 °C for 3 min, then touchdown PCR for 5 cycles, 94 °C for 30 s, 74 °C for 40 s, decreasing 2 °C for 5 cycles, and 72 °C for 3 min, followed by 25 cycles of 94 °C for 30 s, 62 °C for 40 s, 72 °C for 3 min, and a final extension step of 72 °C for 10 min. Due to the presence of CG-rich regions at the 5'-end of the receptor cDNA, instead of performing the 5'-RACE PCR, specific primers were used based on the genomic sequence. The sense primer was 5'-TTGCTAGCGCAATCATGCTGCACTTAAGGCTATTCG-3' (corresponding to nucleotide positions (–)8–22 of Fig. 1) and the antisense was 5'-GCCCCGTGGAATAATGCCCTCGAGCTGG-3' (corresponding to nucleotide positions 1001–1028 of Fig. 1). The PCR program was: 96 °C for 3 min followed by 35 cycles of 96 °C for 1 min, 55 °C for 30 s, and 72 °C for 3 min, using the Advantage2 PCR enzyme system (Clontech) and 5% formamide in the PCR mixture. All PCR products were cloned into pCR4-TOPO (Invitrogen), using the TOPO TA Cloning method (Invitrogen).

Northern blots were prepared, using the NorthernMax-Formaldehyde kit (Ambion) and BrightStar-Plus membranes (Ambion). A cDNA probe (corresponding to nucleotide positions 1505–2146 of Fig. 1) was labelled using the Strip-EZ DNA kit (Ambion). The *Drosophila* ribosomal protein 49 probe was generated as described in [10].

Chinese hamster ovary (CHO) cells were grown as described previously [6]. To amplify a full length cDNA coding for the receptor, the following primers were applied: the sense primer 5'-TTGCTAGCGCA ATCATGCTGCACTTAAGGCTATTCG-3' (corresponding to nucleotide positions 1–22 of Fig. 1) and the antisense primer 5'-TTG AATTCTCACAGGATGTCCGTGTTCCCTTTAGAG-3' (corresponding to nucleotide positions 1493–1521 of Fig. 1), using the *Pfu-Turbo* enzyme system (Stratagene). The *NheI* and *EcoRI* restriction sites which had been incorporated into the above primers facilitated the direct cloning into the pIRES2-EGFP vector (Invitrogen) after digestion of the PCR product with the above restriction enzymes. The insert was fully sequenced and the plasmid was transfected into CHO cells, as previously described [6]. The bioluminescence assay was performed as in [6,11].

DNA sequence compilation, nucleotide, and amino acid sequence comparisons were performed using the Lasergene DNA Software package (DNASTAR). The secondary structure of the receptor protein was analysed, using the TMHMM (v. 2.0) prediction server from the Center for Biological Sequence Analysis, BioCentrum-DTU (www.cbs.dtu.dk).

Results

The website of the *Drosophila* Genome Project consortium (www.flybase.org) contains the annotated gene CG6111, which is presumed to code for a G protein-coupled receptor. We constructed primers based on the six predicted exons of this gene and performed PCRs, using *Drosophila* cDNA as a template. These PCRs resulted in bands of the expected sizes and sequences, after which we performed 3'- and 5'-RACEs to obtain a more complete cDNA sequence. During this part of the work, it became clear that the gene CG6111 had not been correctly annotated and that there was an extra exon in front (5') of the annotated six exons (exon 1 in Fig. 1). Furthermore, this additional exon was annotated to be part of a different gene not related to G protein-coupled receptors (CG14547; www.flybase.org).

Table 1
Intron/exon boundaries of the CCAP receptor gene

Intron	5'-Donor	Intron size (bp)	3'-Acceptor	Intron phase
1	GAG gtgggtg... Glu	14352	...cttcag ACC Thr	3
2	G gtgagtt... Asp	62	...tttcag AT Asp	1
3	T gtaagat... Trp	190	...cttttag GG Trp	1
4	G gtaagat... Arg	83	...ttaacag AA Leu	1
5	AG gtatggt... Ser	91	...tttcag T Ser	2
6	GAG gtatgat... Glu	67	...ccttttag CGA Arg	3

Table 2

Nucleotide differences between the cDNA of Fig. 1 and the corresponding genomic sequences from the Berkeley “*Drosophila Genome Project*”

Position of the nucleotide in the cDNA	Type of nucleotide in the gene	Type of nucleotide in the cDNA	Change in amino acid
849	G	A	Val → Val
1134	A	C	Pro → Pro
1721	A	G	—
1731	A	T	—

The cDNA of the corrected gene CG6111 is shown in Fig. 1. It has a length of 2294 bp, a polyadenylation site at its 3'-end, and an in-frame stop codon preceding the first ATG (start) codon in its untranslated 5'-region. The cDNA codes for a protein of 506 amino acid residues, which has seven transmembrane domains, suggesting that it is a G protein-coupled receptor. Furthermore, the protein contains two potential N-glycosylation sites (following the NXS/T consensus sequence) in its extracellular N-terminal region (Fig. 1).

Comparison of the cDNA of Fig. 1 with the genomic sequence of the corrected CG6111 gene revealed that the receptor gene has seven exons and six introns (Fig. 1, Table 1). This comparison also showed that there are a few nucleotide differences between our cloned cDNA and the corresponding genomic sequences from the database (Table 2). These differences, however, did not lead to changes in amino acid residues (Table 2).

We stably transfected CHO cells with DNA corresponding to the coding region of the receptor. These cells did also stably express the promiscuous G protein, G-16 [11]. Two days before the bioassay, we transiently transfected these cells with DNA coding for the protein apoaquorin and 3 h before the assay, we added coelenterazine to the cell culture medium. Addition of receptor ligands and activation of the cloned receptor in these pretreated cells would lead to an IP_3/Ca^{2+} -mediated bioluminescence response that could easily be measured and quantified [2–6,11–13].

We tested a peptide library consisting of 22 *Drosophila* and other invertebrate neuropeptides on the pretreated CHO cells and found that the cloned receptor was activated by low concentrations of crustacean cardioactive peptide (CCAP) (EC_{50} , 5.4×10^{-10} M). No other *Drosophila* peptides, including Drm-AKH and Drm-corazonin, did activate the receptor (Fig. 2).

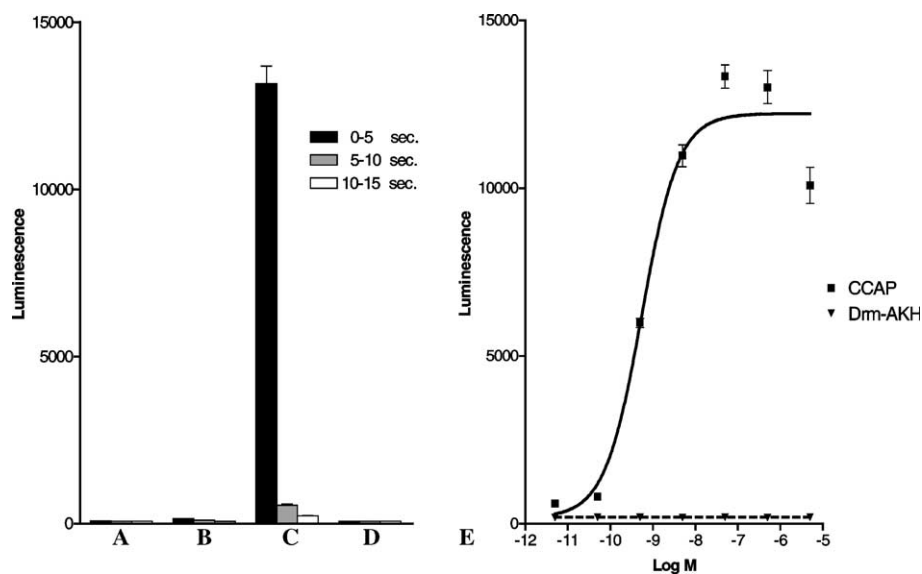


Fig. 2. Bioluminescence responses of nontransfected CHO cells and of a transfected and cloned cell line expressing the receptor of Fig. 1. The SEMs are given as vertical bars, which are sometimes smaller than the symbols. In these cases, only the symbols are given. (A) Responses of nontransfected cells to 10^{-7} M CCAP dissolved in PBS, containing 0.1% BSA. (B) Responses of the transfected cell line to PBS, containing 0.1% BSA. (C) Responses of the transfected cell line to 10^{-7} M CCAP dissolved in PBS, containing 0.1% BSA. (D) Responses of the transfected cell line to 10^{-5} M Drm-AKH. Drm-AKH was also tested on cells expressing the Drm-AKH receptor, where it gave a normal (high) bioluminescence response [6]. (E) Dose-response curve of CCAP for the activation of the receptor (filled squares). Drm-AKH (filled triangles) does not activate the receptor. The following peptides did not activate the receptor either (in concentrations up to 10^{-6} or 10^{-5} M): capa-1 and -2; *Drosophila* corazonin; *Drosophila* ecdysis triggering hormones-1 and -2; *Drosophila* myosuppressin; *Drosophila* myotropin; *Drosophila* pigment dispersing hormone; *Drosophila* pyrokinin-2; *Drosophila* small neuropeptide F-1; *Drosophila* tachykinin-3; Drostatin-A2, -B1, -C; FMRamide; *Heliothis zea* AKH; hug- γ ; Leucokinin III; Leucopyrokinin; and proctolin. For peptide structures see [8,9,20,24–26].

DCCAPR	MLHLRLFDSSLYYTLASASESSGLASSTSTERSFNGTQAGGVAAGGESLTPDVAAVNLTYFTPAISHV	70
DCR	-----MEDEWGSFDRLPSPVPSASMDLETENEVVSINWSTLANFTRLVAGAAPEIINYTLNMDVGVGMATD	65
DCCAPR	MLAPTTIATTTASATMVQIQTTAAPHSDLETGGNSTSSDPGEFDNLNSFYFYETEQFAVLWILFTVIVLG	140
ACCAPR	-----TEQFAVLWILFTVIVLG	17
DAKHR	-----MAKVAEENDHRDLSNWSNVNDTNGTIHLTKDMVFENDGHRLSITVYSILFVISTIG	55
DCR	ISNLSVSTTPLPAYAISNSSSLAHTNSRHEAPPMAEQVPEHVMHDAPQLSRSGLLKVYVLAVMALFSLLG	135
DCCAPR	NSAVLFVVMFINKNR-----KSRMNYFIKQLALADLCVGLLNVLTDIIWRITISWRAGNLACKAIRFSQ	203
ACCAPR	NSAVLVTLMLNRTR-----KSRMNYFIKQLALADLCVGLLNVLTDIIWRITVVRAGNAACKAIRFVQ	80
DAKHR	NSTVLYLLTKRRLRG-----PLRIDIMLHLAIADLMVLTLLMPMEIVWAWTVQWLSTDLMLCRMSEFR	119
DCR	NLLTIWNIIYKTRISRRNSRHTWSAIYSLMFHLSIADVLVTFWFCIIGEAANCYTQVWLANELTCKLVKLFQ	205
DCCAPR	VCVTYSSTYVLVAMSIDRYDAITHPMNFSSKSWKRARHLVAGAWLISALFSLPILVLY---EEKLIQGHQP	270
ACCAPR	VCVTYASTYVLVALSIDRYDAITHPMNFSSKSWKRARHLVAGAWLISALFSLPITYFY---EERLIQGMQ	147
DAKHR	VFGLYLSSYVMVCISLDYFAILLKPLKRS--YNGRIMLACAWLGSVVCSPQAFLFHLEHPAVTCYFQ	187
DCR	MFSLYLSTYVLVLIGVDRWIAVKYPMKSLNMAKRCHRLGGTYILSLVLSLPQFFIFHVARGPPEVEFYQ	275
DCCAPR	CWIE--LGSPIAWQVYMSLVSATLFAIPALIISACYAIIIVKTIWAKGSIFVPTER--AGFGAAPARRASS	336
ACCAPR	CWID--LVEAWRWQLYMCWVSGSLFVVPALIIISACYAVIVRTIWKGTILGPIIDRTHNGMADLATRRASS	215
DAKHR	CVIFSSFRSDFDEKLYQAASMSMYAFPLIMFIYCYGATYLEIYRK-----SQRVLKDVIAERFRSND	251
DCR	CVTHGFYTADWQEQMYATFTLVFTLELLPLCLLFGTYMSTFRTISSSEKMFQ-GSKLANYSTAKLPTQTNR	344
DCCAPR	RGIIIPRAKVTKMTLTIVFVFIICWSPYIIFDLLQVFGQIPHSQTNIAIATFIQSLAPLNSAANPLIYC	406
ACCAPR	RGIIIPRAKVTKMTITIVIVFVLCWSPYIIFDLLQVFGQIPATQTNVAIATFIQSLAPLNSAANPLIYC	285
DAKHR	-DVLRAKKRTLKMTITIVIVFVLCWSPYIISMWYLDKHSAGKINPLLRKALFIFASTNSCMNPLVYG	320
DCR	QRLIHKAKMKSRLISVVIIFLICWTPYVYMMIMFMF-LNPDKRLGDDLQDAIFFFGMSNSLVNPLIYG	413
DCCAPR	LFSSQVFR-----LSRFPFPKWF-----CCCKSYRNNSQ-QNRCHTV	444
ACCAPR	LFSTQVCRI-----IKRLPFRWLWSTK-----WCKPHDSGTLPLNGTIV	327
DAKHR	LYNIRGRMN-----NNNPSVNNRHTSLNRLDSSNQLMQKLTNNLLNNGRGQVM	370
DCR	AFHLCPGKGGKSSGGGNNNAYSLNRGDSQRTPSMLTAVTQVDGTGGSSRQMRARFQQSYRRSSSNGTAG	483
DCCAPR	G-----RRLHN--SCDSMRTLTSLTVSRSTNKTARVVICERPTK--VVTVPAMSERGVSLKGNID	504
ACCAPR	NGTA--HGRFHNHNSDSMRTLTSLTVSRSSCLRP-ARVVIVERPK-----TELAMSE-----	378
DAKHR	AAAVSATTKLANVVSLKGNANGNGSAAAAGTVPTPPLTVTIAPLATDDEANDDSCLSAVTIRCQDQSPI	440
DCR	PGAAPFKEQVGLLHVGPNGTPGGSVSSGETPQLIRKGSALLARQPS---CLREQEHQORLLHEKPSL	550
DCCAPR	IL-----	506
DAKHR	RQK-----	443
DCR	VLSYDSQRGGVGGVASGLLDNNERVSSV	579

Fig. 3. Amino acid sequence comparison between the *Drosophila* CCAP receptor (DCCAPR), the putative *Anopheles* CCAP receptor (ACCAPR) (corresponding to the annotated gene product with Accession No. [EAA01142.1](#)), the *Drosophila* AKH receptor (DAKHR), and the *Drosophila* corazonin receptor (DCR). Spaces are introduced to optimize alignment. Amino acid residues that are identical between DCCAPR and at least one of the other receptors are highlighted in grey. The seven membrane spanning domains (for DCCAPR) are indicated by TMI-VII. Common introns between the DCCAPR gene and one of the other receptor genes are indicated by vertical boxes.

Fig. 2E also shows that high concentrations of CCAP (10^{-5} M) partly inactivates the receptor, which is probably due to receptor desensitization. Furthermore, we noticed that CCAP sticks to the polyethylene tubing and other parts of our luminometer, which made it difficult to wash the peptide away (and get a zero background when new cells were tested) after a CCAP experiment. We solved this problem by dissolving the peptide in PBS + 0.1% bovine serum albumin (BSA).

A comparison of the *Drosophila* CCAP receptor with the other identified or annotated *Drosophila* neuropeptide receptors showed that the receptor was mostly related to the *Drosophila* AKH receptor (23% identical amino acid residues; 32% identities within the seven transmembrane domain; and 51% similar residues), followed by the *Drosophila* corazonin receptor (20% identical residues; 29% identical residues in the seven transmembrane domain; and 50% similar residues) (Fig. 3). However, the *Drosophila* CCAP receptor was

somewhat more related to the mammalian vasopressin V1b receptors than to any of the *Drosophila* receptors (26% identical residues; 35% identical residues in the seven transmembrane region; and 55% similar residues). None of the six introns in the CCAP receptor were shared by the *Drm*-AKH or *Drm*-corazonin receptors. All this showed that none of the *Drosophila* receptors were closely related to the *Drosophila* CCAP receptor.

A comparison of the *Drosophila* CCAP receptor with the sequences stored in the Genome Project database from the malaria mosquito *Anopheles gambiae* [14] revealed the existence of an *Anopheles* receptor that showed strong sequence similarities with the *Drosophila* CCAP receptor (70% identical amino acid residues; 80% identical residues in the transmembrane domain; and 89% similar residues). Furthermore, all six introns found in the *Drosophila* CCAP receptor gene did also occur (at identical positions and with

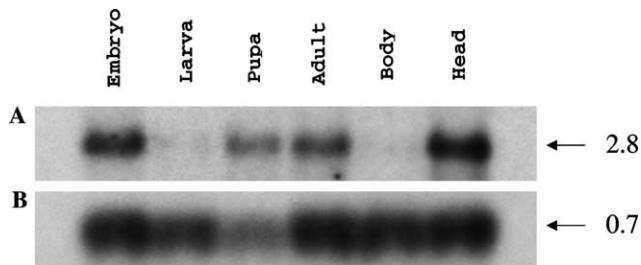


Fig. 4. Northern blots of different developmental stages from *Drosophila*. Each vertical lane contained about 5 μ g mRNA from embryos, larvae, pupae, adult flies, and body (thorax/abdomen) and head from adult flies. The numbers at the right give mRNA sizes in kilobase. (A) Hybridization with a cDNA probe specific for the *Drosophila* CCAP receptor mRNA. (B) The blot of (A) was stripped and hybridized with a cDNA probe specific for ribosomal protein-49. This blot gives the loading efficiency.

identical intron phasings) in the *Anopheles* gene, suggesting that this gene is coding for an *Anopheles* CCAP receptor (Fig. 3).

Northern blots showed that the receptor is expressed in all developmental stages from *Drosophila*, but only very weakly in larvae. In adult flies, the CCAP receptor is expressed mainly or exclusively in the head (Fig. 4).

Discussion

Very recently Park et al. [15] have also identified gene CG6111 as being a gene coding for *Drosophila* CCAP. There are, however, major differences between our findings presented here and those of Park et al. First, we find that the CCAP receptor has a much higher affinity for CCAP (EC_{50} , 5.4×10^{-10} M) than that Park et al. reported (EC_{50} , 1.3×10^{-7} M). The reason for this discrepancy might be that Park et al. used a different expression system (frog oocytes) than we did in our present study (CHO cells). Second, we have sequenced the complete coding region and non-coding 3'-end of the CCAP receptor cDNA, whereas Park and co-workers [15] have missed the whole last exon (exon 7, Fig. 1). This means that their expressed CCAP receptor lacks a portion of the C-terminus, which also might explain the low affinity of their receptor for its ligand CCAP. Third, we found that the CCAP receptor is specific for CCAP and that it does not crossreact with AKH or other *Drosophila* neuropeptides (Fig. 2). Park et al., however, find that their receptor is also activated by AKH in concentrations that are comparable to CCAP (EC_{50} , 2.4×10^{-7} M). One possible explanation for this amazing finding (the two peptides do not have similar structures) could be that their equipment (tubing and measuring chamber) still contained adsorbed CCAP after a CCAP measurement, which was then released during subsequent

measurements of AKH. As mentioned earlier, we had initially similar problems, which, however, could be overcome when CCAP was dissolved in PBS, containing 0.1% BSA. Therefore, we are convinced that the CCAP receptor is specific for CCAP and that it does not crossreact with AKH.

The *Drosophila* AKH and corazonin receptors are structurally related (30% identical amino acid residues; 56% similar residues) and, moreover, their genes have a common intron with identical intron phasings, showing that the two receptors also are clearly evolutionarily related [3]. As mentioned in the Results part, such a clear evolutionary relationship does not exist between the CCAP receptor and the AKH and corazonin receptors, although the CCAP receptor is structurally somewhat similar to the AKH/corazonin receptor group (Fig. 3). Functionally, however, the three receptors are more clearly related. In locusts it has been shown that CCAP releases AKH from the corpora cardiaca [16]. CCAP increases heartbeat in the moth *Manduca sexta* [17,18], as does AKH in the cockroach *Periplaneta americana* [19]. In probably all insects, AKH is involved in the mobilization of lipids from the fat body during energy-requiring activities such as flight and locomotion [20], which are situations where heartbeat is certainly also increased. Corazonin, again, increases heartbeat in cockroaches [21]. Whether all these actions do also occur in a single insect such as *Drosophila* is an open question. However, our Northern blots (Fig. 4) clearly showed that the *Drosophila* CCAP receptor is abundantly present in the head (brain and associated organs such as the corpora cardiaca), but virtually absent in the thorax and abdomen of *Drosophila*. This means that if CCAP would regulate the activity of the heart or other peripheral organs in *Drosophila*, its actions should be indirect, for example by releasing AKH or other hormones from the corpora cardiaca.

CCAP was originally isolated from the shore crab *Carcinus maenas*, because of its prominent cardioactive actions [7]. This cyclic peptide (PFCNAFTGCamide, where the Cs form a disulphide bridge) has later been identified in an identical form from various insects, including *Drosophila* [8,9,16–18]. In addition to its cardioexcitatory and AKH-releasing activities, CCAP also appears to be involved in insect molting, where it might steer the motor behaviour associated with ecdysis [22]. CCAP, therefore, might be a multifunctional peptide involved in various motor behaviours.

CCAP is remarkable in that its structure has remained unchanged during the last 500 million years of evolution (since the divergence of insects and crustaceans [23]). This would suggest that also its receptor has remained relatively invariable (at least in its peptide binding regions) and that it would be easy to identify the CCAP receptors from other arthropods, now that we

have identified the first specific invertebrate CCAP receptor from *Drosophila*. That this is a feasible option is illustrated by our identification of a likely CCAP receptor in the malaria mosquito *Anopheles gambiae* (Fig. 3). This receptor work on other model insects and crustaceans, together with the more genetically orientated work in *Drosophila* (e.g., receptor knock-outs), will certainly advance our knowledge on the actions of CCAP in invertebrates.

Acknowledgments

We thank Drs. S. Rees and J. Stables (Glaxo Wellcome, Stevenage, UK) for supplying the cell line CHO/G-16, Birgitte Paulsen for typing the manuscript, and Lundbeck Foundation, and Fabrikant Vilhelm Pedersen og Hustrus Mindelegat (Manufacturer Vilhelm Pedersen and Wife Memorial Legacy; this support was granted on recommendation from the Novo Nordisk Foundation) for financial support.

References

- [1] M.D. Adams et al., *Science* 287 (2000) 2185–2195.
- [2] G. Cazzamali, C.J.P. Grimmelikhuijzen, *Proc. Natl. Acad. Sci. USA* 99 (2002) 12073–12078.
- [3] G. Cazzamali, N.P.E. Saxild, C.J.P. Grimmelikhuijzen, *Biochem. Biophys. Res. Commun.* 298 (2002) 31–36.
- [4] A. Iversen, G. Cazzamali, M. Williamson, F. Hauser, C.J.P. Grimmelikhuijzen, *Biochem. Biophys. Res. Commun.* 299 (2002) 628–633.
- [5] A. Iversen, G. Cazzamali, M. Williamson, F. Hauser, C.J.P. Grimmelikhuijzen, *Biochem. Biophys. Res. Commun.* 299 (2002) 924–931.
- [6] F. Staubli, T.J.D. Jørgensen, G. Cazzamali, M. Williamson, C. Lenz, L. Søndergaard, P. Roepstorff, C.J.P. Grimmelikhuijzen, *Proc. Natl. Acad. Sci. USA* 99 (2002) 3446–3451.
- [7] J. Stangier, C. Hilbich, K. Beyreuther, R. Keller, *Proc. Natl. Acad. Sci. USA* 84 (1987) 575–579.
- [8] J. Vanden Broeck, *Peptides* 22 (2001) 241–254.
- [9] D.R. Nassel, *Prog. Neurobiol.* 68 (2002) 1–84.
- [10] F. Hauser, L. Søndergaard, C.J.P. Grimmelikhuijzen, *Biochem. Biophys. Res. Commun.* 249 (1998) 822–828.
- [11] J. Stables, A. Green, F. Marshall, N. Fraser, E. Knight, M. Sautel, G. Milligan, M. Lee, S. Rees, *Anal. Biochem.* 252 (1997) 115–126.
- [12] C. Lenz, M. Williamson, G.N. Hansen, C.J.P. Grimmelikhuijzen, *Biochem. Biophys. Res. Commun.* 286 (2001) 1117–1122.
- [13] T. Secher, C. Lenz, G. Cazzamali, G. Sørensen, M. Williamson, G.N. Hansen, P. Svane, C.J.P. Grimmelikhuijzen, *J. Biol. Chem.* 276 (2001) 47052–47060.
- [14] R.A. Holt et al., *Science* 298 (2002) 129–149.
- [15] Y. Park, Y.J. Kim, M.E. Adams, *Proc. Natl. Acad. Sci. USA* 99 (2002) 11423–11428.
- [16] D. Veeleart, P. Passier, B. Devreese, J. Vanden Broeck, J. Van Beeumen, H.G. Vullings, J.H. Diederens, L. Schoofs, A. De Loof, *Endocrinology* 138 (1997) 138–142.
- [17] H.K. Lehman, C.M. Murgic, T.A. Miller, T.D. Lee, J.G. Hildebrand, *Peptides* 14 (1993) 735–741.
- [18] K. Furuya, S. Liao, S.E. Reynolds, R.B. Ota, M. Hackett, D.A. Schooley, *Biol. Chem. Hoppe Seyler* 374 (1993) 1065–1074.
- [19] R.M. Scarborough, G.C. Jamieson, F. Kalish, S.J. Kramer, G.A. McEnroe, C.A. Miller, D.A. Schooley, *Proc. Natl. Acad. Sci. USA* 81 (1984) 5575–5579.
- [20] G. Gäde, K.H. Hoffmann, J.H. Spring, *Physiol. Rev.* 77 (1997) 963–1032.
- [21] J.A. Veenstra, *FEBS Lett.* 250 (1989) 231–234.
- [22] S.C. Gammie, J.W. Truman, *J. Exp. Biol.* 202 (1999) 343–352.
- [23] G.E. Budd, N.J. Butterfield, S. Jensen, *Science* 294 (2001) 2047a.
- [24] C. Lenz, M. Williamson, C.J.P. Grimmelikhuijzen, *Biochem. Biophys. Res. Commun.* 273 (2000) 1126–1131.
- [25] M. Williamson, C. Lenz, Å.M.E. Winther, D.R. Nässel, C.J.P. Grimmelikhuijzen, *Biochem. Biophys. Res. Commun.* 281 (2001) 544–550.
- [26] M. Williamson, C. Lenz, Å.M.E. Winther, D.R. Nässel, C.J.P. Grimmelikhuijzen, *Biochem. Biophys. Res. Commun.* 282 (2001) 124–130.