



ACADEMIC
PRESS

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Biochemical and Biophysical Research Communications 298 (2002) 31–36

BBRC

www.academicpress.com

Molecular cloning and functional expression of a *Drosophila* corazonin receptor[☆]

Giuseppe Cazzamali, Nicolaj P.E. Saxild, and Cornelis J. P. Grimmelikhuijzen*

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

Received 6 September 2002

Abstract

The insect adipokinetic hormones (AKHs) constitute a large family of neuropeptides that mobilize lipids and sugar from the insect fat body during energy-requiring activities such as flight. We have previously identified the first insect AKH receptors from the fruitfly *Drosophila melanogaster* and the silkworm *Bombyx mori* (Staubli et al., PNAS 2002, 99: 3446–3451). Here, we have cloned the cDNA of a *Drosophila* G protein-coupled receptor that was closely related to the first *Drosophila* AKH receptor both with respect to amino-acid sequence and gene structure. We have subsequently expressed this orphan receptor in Chinese hamster ovary cells and identified *Drosophila* corazonin as the endogenous ligand for the receptor. Corazonin increases heart beat in some insects, but its function in *Drosophila* is unknown. These results are intriguing, because not only are the *Drosophila* AKH and corazonin receptors structurally and evolutionarily related, but also are their preprohormones, which suggests a co-evolution of ligands and receptors. The *Drosophila* corazonin receptor is expressed in embryos, larvae, pupae, and adult flies. Furthermore, a receptor that is structurally very similar to the *Drosophila* corazonin receptor can be found in the genomic database from the malaria mosquito *Anopheles gambiae*.

© 2002 Elsevier Science (USA). All rights reserved.

Insects are the largest animal group on earth and are economically and ecologically extremely important, because most flowering plants depend on insects for their pollination. Despite the importance of insects, however, our knowledge of their endocrinology is still incomplete. During the last 20 years numerous insect peptide hormones have been isolated [1,2], but the identification of their receptors has remarkably lagged behind. This situation will certainly change after the recent publication of the genome sequence from the fruitfly *Drosophila melanogaster* [3], which has supplied us with the annotated gene sequences of 40–50 putative peptide hormone receptors [4]. These annotated sequences, however, are computer-predicted sequences that are sometimes incorrect. Furthermore, the predicted receptors are orphan receptors, i.e., their ligands are unknown. All this means that the predicted se-

quences have to be verified by cloning and that their cognate ligands have to be identified, using bioassays based on the heterologous expression of the receptors in a variety of cells [5–7].

The insect adipokinetic hormones (AKHs) are neuropeptides that steer lipid and sugar mobilization from the insect fat body during energy-requiring activities such as flight and locomotion [1,2,8,9]. We have recently identified the first insect AKH receptors, namely those from the fruitfly *D. melanogaster* and the silkworm *Bombyx mori* [5]. For the *Drosophila* receptor, this was done by expressing an orphan receptor [10] in Chinese hamster ovary (CHO) cells, and by using these transformed CHO cells as a bioassay to isolate the cognate ligand from a *Drosophila* extract (“the orphan receptor strategy”) [5,7]. For the *Bombyx* receptor, we similarly expressed an orphan receptor in CHO cells, but here we screened a peptide library to identify the ligand (“the reverse pharmacology strategy”) [5,7].

The AKHs constitute a large peptide family with about 80 different members isolated from more than 30 insect species [1]. Often one insect species produces two or three structurally different AKHs, suggesting that

[☆] The nucleotide sequence reported in this paper has been submitted to the GenBank Data Bank with Accession No. AF373862.

* Corresponding author. Fax: +45-35-32-12-00.

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

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

these AKHs have different functions, and that one insect might have several different AKH receptors [1,5].

The *Drosophila* Genome Project database (www.fly-base.org) contains the sequence of a second annotated G

protein-coupled receptor (CG10698), of which we assumed that it was a second *Drosophila* AKH receptor [5]. This assumption was based on its high percentage of amino-acid sequence identity with the first *Drosophila*

GCACGGAAATTTGAAAAATATGCAAGCACTTCGCTGACGCTTGGATTTCAGGATTTTCAAGTGGAAACAGAGTTTGTGGAGCAAGTGATAAG	-496
ATTGAGGAGTGGGAATTCACGCGAAAGGGCAATTAAGGCCAAGGAGTGCCAGTGCATTTTCGTGCTCCACAGAAAATACGATCCCAAGTGGAA	-397
TTTATGTAATTTAAATTTAACTCACTATTTAATACACAGTCTAAGTGTCATTTAAGTGTGATAACAGTTTAAACCAAGCAACAGCAAGATC	-298
	-199
TATTAATAGAGATATTACAAGCAATTTATTTATCCAACCTAAGCACTAATGATGCAATGCTGGTGACATATAAAACGAAAGTGGAAACGGAGTT	-100
CATTCTCTAAGTAACCAATTTAAGATGTTGTGCTCCCAAGATCCCTAAACCCGCTCCAAACCCCTTGAATAACTCATTAAAGCAAGATATAGACAAA	-1
ATG GAG GAC GAG TGG GGC TCC TTT GAT CGC CTG CCC AGT GTT CCG AGT GCC TCA ATG GAT TTG GAG ACG GAA AAC	75
Met Glu Asp Glu Trp Gly Ser Phe Asp Arg Leu Pro Ser Val Pro Ser Ala Ser Met Asp Leu Glu Thr Glu Asn	25
GAG GTG GTC AGC AAT TGG TCC ACA CTG GCC AAC TTC ACG CGA CTT GTG GCT GGT GCC GCT CCT GAA ATC ATC AAC	150
Glu Val Val Ser Asn Trp Ser Thr Leu Ala Asn Phe Thr Arg Leu Val Ala Gly Ala Ala Pro Glu Ile Ile Asn	50
TAT ACG CTG AAC ATG ATC GAC GTG GGT GTG GGC ATG GCC ACG GAT ATA TCC AAT CTG AGC GTC AGC ACT ACG CCC	225
Tyr Thr Leu Asn Met Ile Asp Val Gly Val Gly Met Ala Thr Asp Ile Ser Asn Leu Ser Val Ser Thr Pro	75
CTG CCC GCC TAC GCA ATC TCT AAT AGC TCC TCG CTG GCG CAC ACC AAT AGT CGC CAT GAA GCA CCA CCG ATG GCG	300
Leu Pro Ala Tyr Ala Ile Ser Asn Ser Ser Ser Leu Ala His Thr Asn Ser Arg His Glu Ala Pro Pro Met Ala	100
GAA CAG GTT CCG GAG CAC GTG ATG GAT CAC GCA CCT CAA CTA TCC CGA TCC GGG TTA CTA AAA GTG TAT GTC CTG	375
Tyr Glu Val Pro Glu His Val Met Asp His Ala Pro Gln Leu Ser Arg Ser Gly Leu Leu Lys Val Tyr Val Leu	125
GCG GTA ATG GCA CTG TTC TCT CTG CTG GGC AAC CTG CTG ACC ATC TGG AAT ATC TAC AAA ACC CGC ATC TCA AGA	450
Ala Val Met Ala Leu Phe Ser Leu Leu Gly Asn Leu Leu Thr Ile Trp Asn Ile Tyr Lys Thr Arg Ile Ser Arg	150
AGA AAC TCA CGG CAC ACG TGG AGT GCT ATC TAC TCA CTG ATG TTC CAT CTG TCC ATC GCC GAT GTC CTG GTC ACC	525
Arg Asn Ser Arg His Thr Trp Ser Ala Ile Tyr Ser Leu Met Phe His Leu Ser Ile Ala Asp Val Leu Val Thr	175
TGG TTC TGC ATC ATC GGG GAG GCC GCG TGG TGC TAC ACC CTC CAG TGG CTG GCC AAT GAG CTC ACC TGC AAG CTG	600
Trp Phe Cys Ile Ile Glu Glu Ala Ala Trp Cys Tyr Thr Val Gln Trp Leu Ala Asn Glu Leu Thr Cys Lys Leu	200
GTG AAG CTC TTC CAG ATG TTT AGC CTC TAC CTG AGC ACC TAT GTC CTG GTC CTC ATC GGA GTG GAC CGC TGG ATA	675
Val Lys Leu Phe Gln Met Phe Ser Leu Tyr Leu Ser Thr Tyr Val Leu Val Leu Ile Gly Val Asp Arg Trp Ile	225
GGG GTC AAG TAT CCG AAG TCG CTC AAC ATG GCC AAG AGG TGT CAT AGG CTA CTT GGC GGT ACT TAC ATC CTG	750
Ala Val Lys Tyr Pro Met Lys Ser Leu Asn Met Ala Lys Arg Cys His Arg Leu Leu Gly Gly Thr Tyr Ile Leu	250
TGG CTG GTG CTC AGC TTG CCA CAG TTC TTC ATG TTC CAT GTG GCG CGT GGC CCA TTG GTG GAG GAG TTC TAC CAG	825
Ser Leu Val Leu Ser Leu Pro Gln Phe Phe Ile Phe His Val Ala Arg Gly Pro Phe Val Glu Glu Phe Tyr Gln	275
TGC CTG ACC CAC GGA TTC TAC ACG GCG GAT TGG CAG GAG CAG ATG TAC GCC ACC TTG ACG CTG CTC TTC ACC TTC	900
Cys Val Thr His Gly Phe Tyr Thr Ala Asp Trp Gln Glu Gln Met Tyr Ala Thr Phe Thr Leu Val Phe Thr Phe	300
CTG CTG CCG CTG TGC ATA CTG TTT GGC ACC TAC ATG TCC ACC TTC CGC ACC ATT TCC AGC AGC GAA AAG ATG TTT	975
Leu Leu Pro Leu Ser Cys Ile Leu Phe Gly Thr Tyr Met Ser Thr Phe Arg Thr Ile Ser Ser Ser Glu Lys Met Phe	325
CAG GGA TCA AAG TCG GCC AAC TAC TCA ACG GCC AAA TTG CCC ACG CAG ACG AAT CGC CAG AGG CTG ATA CAC AAG	1050
Gln Gly Ser Lys Leu Ala Asn Tyr Ser Thr Ala Lys Leu Pro Thr Gln Thr Asn Arg Gln Arg Leu Ile His Lys	350
GCC AAG ATG AAG TCG CTT CGC ATA TCC GTG GTG ATC ATC ATA GCG TTC CTC ATC TGC TGG ACG CCC TAC TAC GTC	1125
Ala Lys Met Lys Ser Leu Arg Ile Ser Val Val Ile Ile Ala Phe Leu Ile Cys Trp Thr Pro Tyr Tyr Val	375
ATG ATG ATT ATG ATG TTC CTC AAT CCG GAC AAA AGC GTG GGC GAC GAT CTG CAG GAC CCC ATC TTC TTC TTC	1200
Met Met Ile Met Phe Met Phe Leu Asn Pro Asp Lys Arg Leu Gly Asp Asp Leu Gln Asp Ala Ile Phe Phe Phe	400
GGC ATG TCA AAC ACC CTG GTC AAC CCA CTC ATC TAC GGT GCC TTC CAC CTG TGT CCT GGC AAA GGG GGC AAG TCG	1275
Gly Met Ser Asn Ser Leu Val Asn Pro Leu Ile Tyr Gly Ala Phe His Leu Cys Pro Gly Lys Gly Gly Lys Ser	425
AGC GGC GGC GGC GGC AAC AAC AAC GCC TAC AGC TTA AAC AGC GGC GAC TCG CAG CGC ACT CCA TCC ATG CTA ACG	1350
Ser Gly Gly Gly Gly Gly Asn Asn Asn Ala Tyr Ser Leu Asn Arg Gly Asp Ser Gln Arg Thr Pro Ser Met Leu Thr	450
GCG GTG ACG CAG GTG GAC GGC ACA GGT GGC AGT TCC CGC CAG ATG CGG GCC TTC CGC CAG CAG AGC TAC TAC CGC	1425
Ala Val Thr Gln Val Asp Gly Thr Gly Gly Ser Ser Arg Gln Met Arg Ala Phe Arg Gln Gln Ser Tyr Tyr Arg	475
AGC TCC TCC AAC GGC ACA GCC GGA CCG GGT GCA GCT CCC TTT AAG GAG CAG GTT GGC CTG CTG CAC GTG GGT CCC	1500
Ser Ser Ser Asn Gln Thr Ala Gly Pro Gly Ala Ala Pro Phe Lys Glu Gln Val Gly Leu Leu His Val Gly Pro	500
GGC AAT GGC ACG CCC GGT GGC TCC GTC TCT AGC GGC GAA ACG CCG CAG TTG ATC CGT AAG GGA TCG GCT CTG TTG	1575
Gly Asn Gly Thr Pro Gly Gly Ser Val Ser Ser Gly Glu Thr Pro Gln Leu Ile Arg Lys Gly Ser Ala Leu Leu	525
GCC CGA CAA CCC AGC TGT CTG AGG GAG CAG GAG CAC CAG CAG CGT TTG CTG CTG CAC GAG AAG CCC TCG ACC CTG	1650
Ala Arg Gln Pro Ser Cys Leu Arg Glu Gln Glu His Gln Gln Arg Leu Leu Leu His Glu Lys Pro Ser Thr Leu	550
GTG CTC AGC TAC GAC AGC CAG CGG GGC GGA GTG GGC GTG GGC GTG GCC AGC GGT CTG CTG GAC AAC AAC GAG CGA	1725
Val Leu Ser Tyr Asp Ser Gln Arg Gly Gly Val Gly Val Gly Val Ala Ser Gly Leu Leu Asp Asn Asn Glu Arg	575
GTG TCG AGC GTG TGA GAGCAGCTGGCGATGGCGGTAGCGGCGCGGGTGGCGGTGGATGCACCGGACACAGTAGAGCAGCAGCAGCAG	1819
Val Ser Ser Val *	579
CATCAGCGGATGCGGATGCGAATGTGGCGATGGAGGACGCGTCTGCTGCTGCCATTGCGCTGAGCAGGAGCTGGTGGCGGCGAGTTGGAGCTCCAAAG	1918
TGCTTCCAGTGCAGTAGGCGGACTGCCAGTGGATTGATCCGGCTGGCTGAGTGGCGGAGCTGGGATTCAGCGATTCCCTCGGTCAACGTGGGGAG	2017
TGGCGCTGATGATTAGATCACCGGGGGCGTGCACATGCTCTAGATAGTGTCTAATCGAGAGTTAAGTCTTCAACACAATGGGTGTGCTCTAGAGATAGC	2116
GTAGGATTTAGAGCTAGTCCGATCAAGATGCTGCTACCTACCTGGGAGAACTCGTTATCACTTTGCAATGTTCTACTAATCTGACTTCAACACAGATA	2215
CAATATACAGAGATATATCTATCTACATCCCAATACCAATACCAATACCAATACCAATACCAATACCAATACCAATACCAATACCAATACCAAT	2314
AAATCTTGGCAATTCGGCGGGTGAATTCGCGATATAGCAACCAATATACCAATATACCAATATACCAATATACCAATATACCAATATACCAAT	2413
GACTGTAATTTGGACTATGATATATACGACTTCAATGAGTTTCTCTATTAGCAGGGCACTTTGTAATGTTAATTAACAGCGGTTTTCAGTCTTT	2512
CATTTGTGTTCCCGGAGTAACTCAAAATCGATTCTATTAATTTGTGTAATAAGAAATCAGCCAAACAGCGGTAATCTAATCTAATTTGTAATTA	2611
TTTACTCTATTATTTTGAGAGAAATATGTAAAAAAGATTTGTAAGTTTAAATATGATGATGATGTTCAACGACAC (A) _n	2695

Fig. 1. cDNA and deduced amino-acid sequence of the *Drosophila* receptor CG10698. 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 five introns are indicated by arrows and the exon nucleotides, bordering these introns, are highlighted in grey. The seven membrane spanning domains are boxed and labeled TM I–VII. The translation termination codon is indicated by an asterisk. In-frame stop codons in the 5'-noncoding region are underlined. The putative polyadenylation signal in the 3'-noncoding region is underlined twice. Five putative glycosylation sites in the extracellular N terminus are indicated by triangles.

AKH receptor and the presence of common intron positions with identical intron phasings [5]. In the present paper, we have cloned the receptor encoded by the CG10698 gene and identified its cognate ligand, using the “reverse pharmacology strategy.” This ligand is not a *Drosophila* AKH, but a related *Drosophila* neuropeptide, corazonin.

Materials and methods

Primers were constructed based on the proposed exons of the annotated gene, CG10698 (www.flybase.org), and used in PCR with cDNA from *D. melanogaster* third instar larvae (Canton S) as a template. The sense primer was: 5'-ATGGAGGACGAGTGGGGCTCC TTTG-3' (corresponding to nucleotide positions 1–25 of Fig. 1) followed by the nested sense primer 5'-AGTGCCCTCGATGGA TTTGGAGACGGA-3' (corresponding to nucleotide positions 46–71 of Fig. 1) and the antisense was: 5'-CACGCTCGACACTCGCTCG TTGTTG-3' (corresponding to nucleotide positions 1713–1737 of Fig. 1) followed by the nested antisense primer 5'-GCACCGTA-GATGAGTGGGTTGACCA-3' (corresponding to nucleotide positions 1217–1241 of Fig. 1). The PCR program was 95°C for 5 min, followed by 40 cycles of 95°C for 30 s, 56°C for 30 s, 72°C for 2 min, and a final extension step of 72°C for 10 min. 3'-RACE was made with the sense primer 5'-GGCAATGGGACGCCCGGTGGCTCCGTC TCTA-3' (corresponding to nucleotide positions 1501–1531 of Fig. 1) followed by the nested sense primer 5'-GGCG TGGCCAGCGGT CTGCTGGACAACAACGA-3' (corresponding to nucleotide positions 1691–1720 of Fig. 1). 5'-RACE was made with the antisense primer 5'-ACATCAGTGAGTAGATAGCACTCC ACGTGTGCCG TGA-3' (corresponding to nucleotide positions 457–493 of Fig. 1) followed by the nested antisense primer 5'-ACCTGTTCCGCC ATCGGTGGTGCTTCATGGCGA-3' (corresponding to nucleotide positions 276–311 of Fig. 1). All PCR products were cloned into pCR4-TOPO (Invitrogen), using the TOPO TA cloning method (Invitrogen).

Northern blots were carried out as in [6]. The cDNA probe used corresponded to nucleotide positions 1–1740 of Fig. 1. Cell culture, cell transfection, and the bioluminescence assay were performed as in [5]. To amplify a full-length cDNA coding for the receptor, we used the sense primer 5'-CGGGGTACCAAGATGGAGGACGAGTGGGG CTCCTTTG-3' (corresponding to nucleotide positions 1–25 of Fig. 1) and the antisense primer 5'-CCGCTCTCGAGTCACACGCTCGA CACTCGCTCGTT-3' (corresponding to nucleotide positions 1714–1740 of Fig. 1). The PCR product was cloned into pcDNA3.1 (Invitrogen), the insertion was fully sequenced, and one nucleotide substitution was corrected using the Quickchange XL Site-Directed Mutagenesis kit (Stratagene). Peptides were obtained from Bachem

(Bubendorf, Switzerland) or Genemed Synthesis (San Francisco) and sequence analyses were as in [6].

Results

To characterize the annotated *Drosophila* G protein-coupled receptor CG10698 (www.flybase.org) and identify its cognate ligand, we cloned its cDNA. This was done using PCR with primers based on the predicted exon sequences of gene CG10698 and cDNA from third instar larvae as a template. This PCR yielded a band of the expected size and sequence, after which we carried out 3'- and 5'-RACE to obtain the full-length cDNA sequence of the receptor (Fig. 1).

The cDNA of Fig. 1 is 3226 nucleotides long. It contains a polyadenylation signal and a polyA⁺ tail in its untranslated 3' end and several in-frame stop codons in its untranslated 5' region. The cDNA codes for a protein of 579 amino-acid residues, containing seven transmembrane domains. Furthermore, the extracellular N terminus contains five potential glycosylation sites, following the N-X-T/S consensus sequence.

Comparison of the cDNA with the genomic sequence of gene CG10698 revealed the existence of five introns, of which four were lying in the coding region of the gene (Fig. 1, Table 1). This comparison also revealed the presence of a large number of nucleotide differences between our cloned cDNA and the genomic sequence from the database. However, most differences did not lead to a difference in the amino-acid sequence of the protein, with the exception of two residues (Val → Ile at amino-acid residue 49 and Ala → Glu at residue 513 of Fig. 1) (Table 2). Finally, this comparison also showed that the intron/exon organization of gene CG10698 had not been correctly predicted by the *Drosophila* Genome Project consortium (an extra intron was predicted in the coding region, which turned out to be an exon; www.flybase.org).

We stably expressed the cDNA of the coding region of our cloned receptor (Fig. 1) in Chinese hamster ovary (CHO) cells, which also stably expressed the promiscuous G protein, G-16 [11]. Two days before the assay, we

Table 1
Intron/exon boundaries of the corazonin receptor gene

Intron	5'-Donor	Intron size (bp)	3'-Acceptor	Intron phase
1	AAG gtaagtt...	3041	...cttgacag CTT	—
2	CAG gtaagtt... Gln	4581	...gttccag TTC Phe	3
3	A gtgagtt... Ser	64	...cgtgcag GC Ser	1
4	AGG gtgagtt... Arg	900	...cttgacag CTG Leu	3
5	AG gtgagtg... Arg	631	...tccgcag G Arg	2

Table 2

Nucleotide differences between the corazonin receptor 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
–434	G	C	—
–385	C	A	—
–348	C	A	—
–343	A	C	—
–323	T	C	—
–229	A	T	—
–208	G	A	—
–33	C	T	—
–24	T	C	—
30	T	C	Arg → Arg
54	G	A	Ser → Ser
145	G	A	Val → Ile
159	C	G	Leu → Leu
205	T	C	Leu → Leu
213	G	C	Val → Val
219	C	T	Thr → Thr
246	C	T	Ser → Ser
258	A	G	Ser → Ser
357	G	A	Leu → Leu
360	G	A	Leu → Leu
537	T	C	Ile → Ile
792	A	G	Val → Val
819	T	C	Phe → Phe
918	C	A	Ile → Ile
975	C	T	Phe → Phe
1020	A	G	Thr → Thr
1098	T	C	Phe → Phe
1311	G	A	Leu → Leu
1538/1539	CC	AA	Ala → Glu
1816	A	G	—
1827	C	G	—
2171	GG	Absent	—
2172	C	Absent	—
2173	C	Absent	—
2174	TTG	Absent	—
2176	TGTT	Absent	—
2179	C	T	—
2182	A	C	—
2576	T	G	—
2578	T	C	—

transiently transfected these cells with DNA, coding for apoaeguorin, and three hours before the assay, we added coelenterazine to the cell culture medium. Activation of the receptor in these pretreated cells would result in a Ca^{2+} -mediated bioluminescence response, which could be easily measured and quantified [5,6,11].

We applied the "reverse pharmacology strategy" [7] in our search to find the cognate ligand for our cloned orphan receptor (Fig. 1), i.e., we tested a peptide library of *Drosophila* and other insect and invertebrate neuropeptides on the transfected CHO cells. No effects were found for most of these peptides with the exception of *Drosophila* corazonin, which gave a prominent bioluminescence response with an EC_{50} of 1.8×10^{-8} M

(Figs. 2A and B). This identified our receptor as a *Drosophila* corazonin receptor.

We carried out Northern blots of various developmental stages of *Drosophila*, using a probe coding for the *Drosophila* corazonin receptor. These blots showed that the corazonin receptor starts to be expressed in the 8–16 h old embryos and that it continues to be expressed in all developmental stages. In adult flies, both the head and the body (thorax/abdomen) express the corazonin receptor (Fig. 3). The size of the transcripts (3.2 kb) in

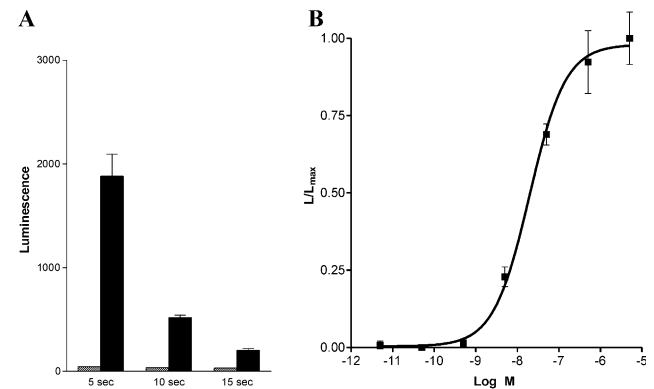


Fig. 2. Bioluminescence responses of nontransfected CHO cells (highlighted in grey) and a CHO cell line permanently transfected with DNA coding for the orphan receptor of Fig. 1 (highlighted in black). The SEMs are given as vertical bars. (A) Bioluminescence responses 0–5, 5–10, and 10–15 s after addition of 10^{-6} M *Drosophila* corazonin. (B) The dose-response curve of *Drosophila* corazonin. The following peptides did not give responses in concentrations up to 10^{-5} M: crustacean cardioactive peptide; *Drosophila* adipokinetic hormone; *Drosophila* myosuppressin; *Drosophila* small neuropeptide F-1; *Drosophila* tachykinin-3; drostatins-A4, -B2, -C; ecdysis-triggering-hormone-2; FMRFamide; *Heliothis* adipokinetic hormone; leucokinin III; leucomyosuppressin; leucopyrokinin; *Manduca* adipokinetic hormone; *Drosophila* pigment dispersing hormone; proctolin. For peptide structures see [1,17–20].

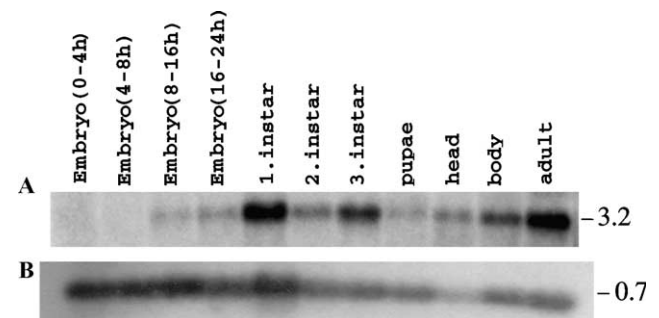


Fig. 3. Northern blots of different developmental stages from *Drosophila*. Each vertical lane contained 2.5 µg mRNA from various embryonic stages, 1st, 2nd, and 3rd instar larvae, pupae, head, and body (thorax/abdomen) of adult flies (mixed sexes), and intact adult flies (mixed sexes). The numbers at the right give the estimated mRNA sizes in kb. (A) Hybridization with a cDNA probe corresponding to nucleotide positions 1–1740 of Fig. 1. (B) The Northern blot of A was stripped and hybridized with a cDNA probe corresponding to ribosomal protein 49. This blot gives the loading efficiency.

DCR	MEDEWGSFDRLPSVPSASMDLETENEVSNWSTLANFTRLVAGAAPEIINYTLNMDVGVGMATDISNLS	70
DAKHR	-----MAKVAEEND	9
RGRHR	-----MANNASLEQ	9
TM I		
DCR	VSTTPLPAYAISNSSSLAHITNSRHEAPPMAEQVPEHVMHAPQLSRSGLLKVVYLAVMAFLFSLGLNLLTI	140
ACR	-----HAPTLKSGSVIRVIVLSAMALVSLGLNVAITM	31
DAKHR	HRDLSNWSNVNDTNGTIHLT-----KDMVFNDGHRLSITVYSILFVISTIGNSTVL	60
RGRHR	DQNHCSAINNSIPLTQG-----KLPTTLTSGKIRVTVTFFLFLSTAFNASFL	57
TM II		
DCR	WNIYKTRISRRNSRHTWSAIYSLMFHLSIADVLVTWFCIIIEAAWCYTVQWLANELTCKLVKLFQMFSLY	210
ACR	WNIQKNRKSRRTVRHNWSAIYSLIFHLSIADVLVTGFCIIIEAAWYTVQWVAGNLCKLVKLFQMFSLY	101
DAKHR	YLLTK--RRLRGPLR----IDIMLMHLAIADLMVTLMLPMEIVWAWTVQWLSITDLCKRMLMSFRVFGLY	124
RGRHR	VKLQR-WTQKRKKGKLSRMKVLKHLTLANLLETILVMPDGMWNIIVQWYAGEFLCKVLKSLKLFMSY	126
TM III		
DCR	LSTYVVLVIGVDRWIAVKYPMKSLNMAKRCHRLGGTYILSLVLSLPIFFIF--HVARGPFVEEFYQCV	277
ACR	LSTYVVLVIGVDRWIAVKYPMKSLNMAKRCHRLGGTYILSLVLSLPIFFIF--RVAKGPFVEEDFYQCV	168
DAKHR	LSSVVMVCIISLDRIYFAILKPLK--SYNRRIMLACAWLGSVVCISIPQAFIF--HLEEHFAVTGYFQCV	189
RGRHR	APAFMMVVISLDRSLAVTQPLAVQSKSLERSMTSLAWILSIVFAGIQLYIFRMIYLDAGSGPAVFSQCV	196
TM IV		
DCR	THGFYTADWQEQMYATFTLVFTFLPLCILFGTYMSTFTTISSEKMFQ--GSKLANYSTAKLPTQTNRQR	346
ACR	THGFYTDRWQEQLYTFTLVFMFIIPLILIGTYLSTFMTTISSEKIFRIDTSAVDRTTYRRSDTNRQR	238
DAKHR	IFSSFRSDFDEKLYQAASCMYAFPLIMFIYCYG----AIY-LEIYRKSQRVLKDVIAERFRSND--	252
RGRHR	THCSFPQWHEAFYNFTTFSCFLIIPLLIML-----ICNAKIIIFALTRVTHQDPRKLQNLQSKNN-	256
TM V		
DCR	LIHKAKMKSRLISVVIITAFILCWTPTYVMMIMFMFLNPK--RLGDDLDQDAIFFFGMSNSLVNPLIYGAF	415
ACR	LIHKAKMKSRLISVVIIVAFVVCWTPYYIMMLIFMFLNPK--RFGEDLQSGIFFFGMSNSLVNPLIYGAF	307
DAKHR	VLSRAKRTLKMTITIVIVFIICTWPTYTISMWYWLDKHSAGKINPLRKALFIHASTNCCMNPVLYGGLY	322
RGRHR	-IPRARLRLKMTIVAFGTSEVICWTPYYVLGIWYWFDEPEMLNRVSEPVNHFFFLFAFLNCPCFDPLIYGFT	325
TM VI		
DCR	HLCPGKGGKSSGGGNNNAYSLSNRGDSQRTFPMSTLAVTQVDTGTGSSSRQMRAFRQSSYYSSSSNGTAGPG	485
ACR	HLVP-----IRQRN	317
DAKHR	NIRG-----RMNNNPSVNNRHTLSNRLDSSNQLMQQLTNNLLNCGQVMAAAVSA	376
RGRHR	SL-----	327
TM VII		
DCR	AAPFKEQVGLLHVGPNGTTPGGSVSSGETEQILRKGSALLARQPSCLREQEHQRLLLHEKPSSTLVLSYD	555
DAKHR	TTKLANVSLKGNANGSAAAGTVPITTEPLTVTIAPLATDDEANDSCLSAVTIRCQDQSPIRQK---	443
DCR	SQRGGVGVGVASGLLDNNERVSSV-----	579

Fig. 4. Amino-acid sequence comparison between the *Drosophila* corazonin receptor (DCR) of Fig. 1, a partial sequence of the putative *Anopheles* corazonin receptor (ACR; see text), the *Drosophila* adipokinetic hormone receptor (DAKHR) [5,10], and the rat GnRH receptor (RGRHR) [21,22]. Spaces are introduced to optimize alignment. Amino-acid residues that are identical between DCR and at least one of the other receptors are highlighted in grey. The seven membrane spanning domains are indicated by TM I–VII. Common introns in the genes are indicated by vertical boxes.

these developmental stages fits very well with the size of our cloned cDNA (Fig. 1).

A comparison of the amino-acid sequence of the *Drosophila* corazonin receptor with that of other proteins from the GenBank database showed that the receptor had 30% amino-acid residue identities with the *Drosophila* AKH receptor (56% similar residues) and 28% amino-acid residue identities with the rat GnRH receptor (49% similar residues). Furthermore, by blasting the *Drosophila* receptor protein sequence into the database of the malaria mosquito *Anopheles gambiae* Genome Project, we could identify a partial sequence on clone AAAB01008987.1 that had 71% amino-acid residue identities (83% similar residues) with the *Drosophila* corazonin receptor sequence. In addition, all three introns known to be present in this coding region of the *Drosophila* corazonin gene did also occur at the same positions and with the same phasings in the *Anopheles* gene (Fig. 4). This *Anopheles* gene sequence (for which there exists no CG number or annotation), therefore, most likely represents an *Anopheles* corazonin receptor gene.

Discussion

We have earlier described that the *Drosophila* AKH and the invertebrate GnRH receptors are structurally

and evolutionarily related [5,10]. In the present paper, we have identified the *Drosophila* corazonin receptor and found a putative corazonin receptor from *Anopheles*. These two receptors are structurally highly related and, in addition, have three introns at the same positions and with the same intron phasings, thereby also showing a close evolutionary relationship (Fig. 4). It is interesting that one of these introns (intron 2, Table 1) is also shared by the *Drosophila* AKH and rat GnRH receptor genes (Fig. 4). Furthermore, the three types of receptors share a considerable number of identical and conserved amino-acid residues (Fig. 4). These data together, then, suggest that the insect corazonin, the insect AKH, and the vertebrate GnRH receptors are all evolutionarily related.

It has earlier been recognized that the peptides corazonin and AKH themselves also are structurally related [12]. Furthermore, their preprohormones have a similar organization [13]. This led Veenstra to conclude that the two neuropeptide precursor genes must have a common evolutionary origin [13]. Because the corazonin and AKH receptors also appear to have a common evolutionary origin, these results suggest a co-evolution of receptors and ligands. Such a co-evolution between the corazonin and AKH receptors and their ligands has also recently been proposed by other researchers [14]. Furthermore, several other examples of receptor and

ligand co-evolution in invertebrates have recently been described [14,15].

The insect neuropeptide corazonin (<ETFQYS RGWTNamide) has originally been isolated from cockroaches, because of its strong cardio-excitatory actions on isolated cockroach hearts [12]. From other insects, such as locusts, His-7-corazonin (<ETFQYSH GWTNamide) has been purified, which induces, in locusts, a strong darkening of the exoskeleton, a phenomenon that is associated with the well-known swarm formation of locusts and their migration [16]. In *Drosophila*, a corazonin identical to that from cockroaches has been found [13], but whether it stimulates heart beat in *Drosophila* is unknown. The molecular identification of the *Drosophila* corazonin receptor makes it now possible (by in situ hybridization or immunocytochemistry) to identify all corazonin target cells in the fly. This, together with the development of knock-out (P-element insertion) receptor mutants from *Drosophila*, will further our knowledge of the actions of corazonin in insects.

Acknowledgments

We thank Drs. S. Rees and J. Stables (Glaxo Wellcome, Stevenage, UK) for supplying cell line CHO/G16, Birgitte Paulsen for typing the manuscript, Michael Williamson for critically reading the manuscript, and Lundbeck Foundation, the Danish Natural Science Research Council (equipment grant), and Novo Nordisk Foundation for financial support.

References

- [1] G. Gäde, K.H. Hoffmann, J.H. Spring, *Physiol. Rev.* 77 (1997) 963–1032.
- [2] L. Schoofs, D. Veelaert, J. Vanden Broeck, A. De Loof, *Peptides* 18 (1997) 145–156.
- [3] M.D. Adams, et al., *Science* 287 (2000) 2185–2195.
- [4] R.S. Hewes, P.H. Taghert, *Genome Res.* 11 (2001) 1126–1142.
- [5] 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.
- [6] G. Cazzamali, C.J.P. Grimmelikhuijzen, *Proc. Natl. Acad. Sci. USA* 99 (2002) 12073–12078.
- [7] O. Civelli, H.P. Nothacker, Y. Saito, Z. Wang, S.H. Lin, R.K. Reinscheid, *Trends Neurosci.* 24 (2001) 230–237.
- [8] J.V. Stone, W. Mordue, K.E. Batley, H.R. Morris, *Nature* 263 (1976) 207–221.
- [9] 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.
- [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] J.A. Veenstra, *FEBS Lett.* 250 (1989) 231–234.
- [13] J.A. Veenstra, *Biochem. Biophys. Res. Commun.* 204 (1994) 292–296.
- [14] Y. Park, Y.J. Kim, M.E. Adams, *Proc. Natl. Acad. Sci. USA* 99 (2002) 11423–11428.
- [15] R.E. Van Kesteren, C.P. Tensen, A.B. Smit, J. van Minnen, L.F. Kolakowski, W. Meyerhof, D. Richter, H. van Heerikhuizen, E. Vreugdenhil, W.P. Geraerts, *J. Biol. Chem.* 271 (1996) 3619–3626.
- [16] A.I. Tawfik, S. Tanaka, A. De Loof, L. Schoofs, G. Baggerman, E. Waelkens, R. Derua, Y. Milner, Y. Yerushalmi, M.P. Pender, *Proc. Natl. Acad. Sci. USA* 99 (1999) 7083–7087.
- [17] J. Vanden Broeck, *Peptides* 22 (2001) 241–254.
- [18] C. Lenz, M. Williamson, C.J.P. Grimmelikhuijzen, *Biochem. Biophys. Res. Commun.* 273 (2000) 1126–1131.
- [19] M. Williamson, C. Lenz, Å.M.E. Winther, D.R. Nässel, C.J.P. Grimmelikhuijzen, *Biochem. Biophys. Res. Commun.* 281 (2001) 544–550.
- [20] M. Williamson, C. Lenz, Å.M.E. Winther, D.R. Nässel, C.J.P. Grimmelikhuijzen, *Biochem. Biophys. Res. Commun.* 282 (2001) 124–130.
- [21] K.A. Eidne, R.E. Sellar, G. Couper, L. Anderson, P.L. Taylor, *Mol. Cell. Endocrinol.* 90 (1992) R5–R9.
- [22] U.B. Kaiser, D. Zhao, G.R. Cardona, W.W. Chin, *Biochem. Biophys. Res. Commun.* 189 (1992) 1645–1652.