The Remarkable Conservation of Corticotropin-Releasing Hormone (CRH)-Binding Protein in the Honeybee (Apis mellifera) Dates the CRH System to a Common Ancestor of Insects and Vertebrates

Mark O. Huising and Gert Flik

Department of Animal Physiology (M.O.H., G.F.), Institute for Neuroscience, Radboud University Nijmegen, 6525 ED Nijmegen, The Netherlands; and Department of Cell Biology and Immunology (M.O.H.), Wageningen University, 6709 PG Wageningen, The Netherlands

CRH-binding protein (CRH-BP) is a key factor in the regulation of CRH signaling; it modulates the bioactivity and bioavailability of CRH and its related peptides. The conservation of CRH-BP throughout vertebrates was only recently demonstrated. Here we report the presence of CRH-BP in the honeybee (Apis mellifera) and other insects. Honeybee CRH-BP resembles previously characterized vertebrate CRH-BP sequences with respect to conserved cysteine residues, gene organization, and overall sequence identity. Phylogenetic analyses confirm the unambiguous orthology of insect and vertebrate CRH-BP sequences. Soon after their discovery, it was noted that insect diuretic hormone-I (DH-I) and its receptor share similarities with the vertebrate CRH family and their receptors. Despite these similarities, demonstration of common ancestry of DH-I and the vertebrate CRH family is still speculative: the mature neuropeptides are short, and their genes differ substantially with regard to the number of coding exons. Moreover, DH and CRH receptors belong to the much larger family of G protein-coupled receptors. In contrast, the unique and conspicuous features of CRH-BP greatly facilitate the establishment of orthology over much larger evolutionary distances. The identification of CRH-BP in insects clearly indicates that this gene predates vertebrates by at least several hundred million years. Moreover, our findings imply that a CRH system is shared by insects and vertebrates alike and, consequently, that it has been present at least since the common ancestor to both phylogenetic lines of proto- and deuterostomians. (Endocrinology 146: 2165–2170, 2005)
DH receptors with the vertebrate CRH family and CRH receptors, establishing orthology is complicated for three major reasons. First, the mature neuropeptides are short in length; second, DH-I and the vertebrate CRH family members differ vastly in gene structure; and third, the DH and CRH receptors are only a few members of the much larger GPCR family. In contrast, the evolutionary well-conserved CRH-BP does not bear appreciable sequence similarity to any other protein and appears to constitute an autonomous protein family. Therefore, CRH-BP is far better suited to establish the age of the CRH system.

We here report the presence of CRH-BP in insects. We cloned the complete coding sequence of the CRH-BP gene from the honeybee (Apis mellifera). Inspection of both the honeybee gene, as well as the corresponding protein, reveals striking similarities to vertebrate CRH-BP sequences. For completeness, we confirm that the honeybee, like many other insect species, possesses a DH-I sequence. Collectively, our findings strongly support the notion that the endocrine CRH signaling system, including its binding protein, is shared by insects and vertebrates and has been present since the common ancestor to both phyla.

Materials and Methods

Animals

Honeybees (Apis mellifera) were obtained from the Dutch beekeeping expertise center ‘Het Bijenhus’ in Wageningen where they were housed according to standard beekeeping practice. Animals were rapidly sedated on ice preceding dissection.

RNA isolation and first-strand cDNA synthesis

Organs for RNA isolation were flash-frozen in liquid nitrogen. RNA isolation was conducted with Trizol (Invitrogen, Carlsbad, CA) according to the manufacturer’s protocol. Total RNA was precipitated in ethanol, washed, and dissolved in water. All reagents for cDNA synthesis were obtained from Invitrogen, and cDNA synthesis was carried out as previously described (15).

Cloning and sequencing

PCR was carried out with bBP.fw1 (AATGACAAATGGAGGAGTCTCCTG) and bBP.rv1 (TTCATACCGATATTTTTACCACA) primers based on a honeybee expressed sequence tag (B514351). The sequence of the remaining encoding N-terminal part of the sequence as well as a short stretch of 5’ untranslated region was obtained by PCR on cDNA from the head of a single bee with bBP.fw4 (GGATTCTTGAGGTTTCATTAGAA) and bBP.rv2 primers. Similarly, the C-terminal part of the sequence as well as a partial 3’ untranslated region were obtained by PCR with bBP.fw2 (TCACTCTTACCCATATACCC) and bBP.rv7 (GATAAATTTATGAAAGACATCTAG) primers. For the assessment of CRH-BP gene expression we used the following primers: bBP.fw3 (CTTGAGATCCTGCTTCACAGG) and bBP.rv3 (GACGCCGACATAAGTGCAATT). Honeybee actin (XM_393368) and 40S ribosomal protein S11 (XM_394541; not shown) were used as internal reference genes, and results were very similar after comparison with either gene. Reference gene primers were actin.fw1 (CCTACCCCATCACCATGAA), actin.rv1 (GAAGACGAAGGACCACCCCA), 40S.fw1 (CCTACAAGACCCGACTCAT), and 40S.rv1 (AAGAATGGCTCTCTAAATGAAATGTT). The mature honeybee DH-I sequence was cloned by PCR with bDH-I.fw1 (GGATTCTTGAGGTTTCATTAGAA) and bDH-I.rv2 (CTTTTCATATGCTTCCAAAAG) primers based on a honeybee genomic sequence retrieved from the Baylor College of Medicine honeybee genome project (assembly Amel 1.2). All oligonucleotides were obtained from Eurogentec (Seraing, Belgium). PCRs were performed with 0.2 μl Taq DNA polymerase (Goldstar, Eurogentec) supplemented with 1.5 mm MgCl₂, 200 nm dNTPs, and 400 nm of each primer in a final volume of 25 μl. Cycling conditions were 94 C for 2 min, 30–35 cycles of 30 sec at 94 C, 30 sec at 55 C, and 1 min at 72 C followed by 10 min at 72 C, and PCRs were carried out on a GeneAmp PCR System 9700 (PE Applied Biosystems, Foster City, CA). PCR products were ligated in the pGEM-T-easy vector (Promega, Madison, WI) and cloned in JM-109 cells according to the manufacturer’s protocol. Plasmid DNA was isolated with the Qiaprep Spin Miniprep kit (QIAGEN, Chatsworth, CA), and sequence reactions were carried out with the ABI Prism Bigdye terminator cycle sequencing ready reaction kit according to the manufacturer’s protocol and analyzed with an ABI 377 sequencer.

Bioinformatics

Multiple sequence alignments were carried out using ClustalW 1.82. The organization of the honeybee CRH-BP gene structure was carried out by comparison of the complete cDNA sequence with the honeybee genome sequence at the Ensemble site (http://www.ensembl.org/). Other nonvertebrate CRH-BP sequences were retrieved via BLAST searches. Phylogenetic trees were constructed on the basis of amino acid differences (p-distance) by the neighbor-joining method using MEGA version 3.0 (21). Reliability of the tree was assessed by bootstrapping, using 1000 bootstrap replications.

Results

The honeybee ortholog of vertebrate CRH-BP sequences was amplified from cDNA of the head of a single worker bee. The complete coding sequence measures 999 nucleotides and encodes a 332-amino-acid protein, which is 10 amino acids longer than most vertebrate CRH-BPs. Honeybee CRH-BP shares highest identity (up to 33%) with two automatically annotated dipteran CRH-BP sequences of the malaria mosquito (Anopheles gambiensis) and the fruit fly (Drosophila melanogaster) (Table 1). Amino acid identity with the more distantly related CRH-BP sequences of various vertebrate species is slightly lower at 25–29%. Eight of the 10 cysteine residues that characterize vertebrate CRH-BP sequences are conserved and identically spaced in honeybee CRH-BP, but the final pair of cysteines is absent (Fig. 1). Furthermore, throughout the alignment, several short stretches of amino acids are identical in all CRH-BP sequences. When honeybee CRH-BP is subjected to a BLAST search to identify the sequences in the Swissprot database that are most similar, the only significant hits are other CRH-BP sequences (Table 2). This once again illustrates the uniqueness of CRH-BP and provides further testimony of the unambiguous orthology of honeybee and vertebrate CRH-BP sequences.

The honeybee CRH-BP gene consists of seven exons, as is the case for all vertebrate CRH-BP genes that have been characterized (Fig. 2). Furthermore, exon sizes correspond well to the lengths of each of the seven vertebrate exons, with honeybee exons two, three, and four each extending merely one triplet over the sizes of their corresponding vertebrate exons. Also the distribution of the conserved cysteine residues over the exons is highly similar in honeybee and vertebrates, and all honeybee introns contain well-recognized 5’ donor (GT) and 3’ acceptor (AG) splice sites.

Phylogenetic analyses corroborate the notion that all CRH-BP sequences conform to the accepted patterns of animal evolution, with sequences from more distantly related species clustering less proximately (Fig. 3). Inclusion of the human IGF-BP family as an outgroup results in a phylogenetic tree where all CRH-BP sequences form a stable cluster, separate from the IGF-BP sequences (not shown).
Because of its CRH motif, insect DH-I would be a likely candidate to bind to insect CRH-BP. Despite its identification in many different insect species, DH-I has not yet been reported in the honeybee. To confirm the presence of DH-I in the honeybee, we amplified the DNA sequence encoding the honeybee mature DH-I peptide. This peptide resembles previously identified DH-I peptides from other insect species (Fig. 4). Moreover, a honeybee DH-receptor sequence has been automatically annotated (XP_397268) from the honeybee genome.

To learn whether insect CRH-BP, like its vertebrate orthologs, is largely centrally expressed, we established the expression pattern of CRH-BP mRNA in the three major body regions of the honeybee. CRH-BP gene expression occurs in the head of the honeybee (Fig. 5) and to a lesser extent in the abdomen but is absent from the thorax.

### TABLE 1. Percentages of amino acid sequence identity for CRH-BP sequences of various animal species

<table>
<thead>
<tr>
<th></th>
<th>Honeybee</th>
<th>Malaria mosquito</th>
<th>Fruit fly</th>
<th>Sea squirt</th>
<th>Carp 1</th>
<th>Carp 2</th>
<th>Pufferfish</th>
<th>Xenopus</th>
<th>Chicken</th>
<th>Sheep</th>
<th>Rat</th>
<th>Mouse</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Honeybee</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malaria mosquito</td>
<td>33.1</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fruit fly</td>
<td>29.8</td>
<td>46.3</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sea squirt</td>
<td>26.8</td>
<td>28.0</td>
<td>28.6</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carp 1</td>
<td>29.0</td>
<td>22.1</td>
<td>24.0</td>
<td>33.9</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carp 2</td>
<td>28.3</td>
<td>23.3</td>
<td>24.6</td>
<td>33.9</td>
<td>97.8</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pufferfish</td>
<td>26.6</td>
<td>23.7</td>
<td>23.3</td>
<td>33.3</td>
<td>67.8</td>
<td>68.4</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xenopus</td>
<td>24.9</td>
<td>21.5</td>
<td>25.6</td>
<td>31.6</td>
<td>58.9</td>
<td>59.2</td>
<td>54.7</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chicken</td>
<td>26.1</td>
<td>21.5</td>
<td>24.3</td>
<td>34.5</td>
<td>59.8</td>
<td>59.5</td>
<td>57.2</td>
<td>72.0</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheep</td>
<td>26.5</td>
<td>24.3</td>
<td>23.6</td>
<td>33.9</td>
<td>57.0</td>
<td>57.0</td>
<td>55.3</td>
<td>64.5</td>
<td>70.8</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rat</td>
<td>26.7</td>
<td>21.5</td>
<td>23.6</td>
<td>33.9</td>
<td>58.6</td>
<td>55.0</td>
<td>63.9</td>
<td>71.1</td>
<td>77.3</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouse</td>
<td>26.7</td>
<td>21.8</td>
<td>24.3</td>
<td>32.1</td>
<td>60.8</td>
<td>60.4</td>
<td>55.9</td>
<td>67.6</td>
<td>73.0</td>
<td>79.8</td>
<td>93.8</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td>28.9</td>
<td>24.3</td>
<td>24.6</td>
<td>31.6</td>
<td>61.7</td>
<td>62.0</td>
<td>58.1</td>
<td>68.2</td>
<td>73.9</td>
<td>85.1</td>
<td>84.5</td>
<td>87.3</td>
<td>100</td>
</tr>
</tbody>
</table>

* Partial sequence only, potentially influencing the percentages amino acid identity with other sequences.

Fig. 1. Amino acid alignment of honeybee CRH-BP with CRH-BP sequences of various vertebrate species. Cysteine residues that are involved in the formation of disulfide bonds are shaded. Asterisks indicate amino acid identity, whereas colons and dots indicate decreasing degrees of amino acid similarity. Accession numbers are as follows: honeybee (Apis mellifera), AJ780964; carp (Cyprinus carpio), AJ490886; Xenopus (Xenopus laevis), Q61653; chicken (Gallus gallus), BU565515 and BU567671; mouse (Mus musculus), Q660571; and human (Homo sapiens), P24387.
TABLE 2. List of BLAST hits after comparison of honeybee CRH-BP to the Swissprot database

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Species</th>
<th>Description</th>
<th>E value</th>
</tr>
</thead>
<tbody>
<tr>
<td>P24388</td>
<td>Rat</td>
<td>CRH-BP</td>
<td>8 × 10⁻²⁹</td>
</tr>
<tr>
<td>Q28557</td>
<td>Sheep</td>
<td>CRH-BP</td>
<td>4 × 10⁻²⁷</td>
</tr>
<tr>
<td>Q66571</td>
<td>Mouse</td>
<td>CRH-BP</td>
<td>3 × 10⁻²⁶</td>
</tr>
<tr>
<td>P24387</td>
<td>Human</td>
<td>CRH-BP</td>
<td>1 × 10⁻²⁵</td>
</tr>
<tr>
<td>Q91653</td>
<td>Xenopus</td>
<td>γ-chain</td>
<td>2 × 10⁻²⁵</td>
</tr>
<tr>
<td>P22482</td>
<td>Bacillus</td>
<td>ATP synthase</td>
<td>1.2</td>
</tr>
<tr>
<td>Q9Y6W3</td>
<td>Human</td>
<td>Calpain 7</td>
<td>6.4</td>
</tr>
<tr>
<td>P43153</td>
<td>Clostridium</td>
<td>Microbial collagenase precursor</td>
<td>8.1</td>
</tr>
</tbody>
</table>

**Discussion**

Here we describe the identification of the complete cDNA sequence of CRH-BP from the honeybee. The conservation of unique features such as key cysteine residues and gene structure provide testimony to its bona fide orthology with vertebrate CRH-BP sequences. The discovery of CRH-BP in the honeybee substantiates that the CRH system predates vertebrates and is likely to share ancestry with insect DH-I and its receptor. Overall amino acid identity of honeybee CRH-BP with various vertebrate CRH-BP sequences is moderate at around 25–29%, which is not surprising as the evolutionary distance between insects and vertebrates is estimated at between 700 and 993 million years (22–24). Nonetheless, the high similarity in gene structure, stable clustering in phylogenetic analyses, as well as the conservation in presence and spacing of the first eight cysteine residues all point clearly to the unambiguous orthology of vertebrate and insect CRH-BPs. The final two C-terminal cysteine residues are missing from the honeybee CRH-BP sequence. The simultaneous disappearance of this pair of cysteines is in line with the observation that they form an intrachain disulfide bridge (1). Furthermore, the obvious sequence identity between the location of both missing cysteine residues, complemented by the presence of this C-terminal cysteine pair in the predicted CRH-BP sequences of Drosophila melanogaster and Anopheles gambiae, indicates a loss of these two cysteines in the honeybee after its divergence from both dipteran species. Other than the cysteine residues, several short amino acid stretches are identical in all sequences, which suggests that these residues are structurally imperative or indispensable for ligand binding.

Establishment of orthology for the vertebrate CRH family members with insect DH-I is not straightforward, because the CRH motif is not very stringent and the mature neuropeptides are short (46 amino acids or fewer), which impairs phylogenetic analyses. Furthermore, the genes encoding all four vertebrate CRH family members (CRH, urotensin-I/urocortin-I, urocortin-II, and urocortin-III) possess two exons and are encoded completely by the second exon, whereas the coding region of tobacco hornworm (Manduca sexta) DH-I is divided over four exons (25). And although insect diuretic receptors and vertebrate CRH receptors both belong to the class B (secretin-like) family of GPCRs, this family also includes receptors for secretin, vasoactive intestinal peptide, PTH and its related peptide, GHRH, calcitonin, and others (26), which complicates establishment of one-to-one orthology. In contrast, the uniqueness of CRH-BP greatly facilitates establishment of orthology over large evolutionary distances.

CRH-BP takes its name from the modulation of CRH bioactivity, either antagonistically via abrogated CRH signaling (8, 9) or agonistically via extension of protein half-life (6). But CRH-BP also has the potential to bind to and modulate signaling of other members of the CRH family. In fact, several reports indicate that CRH-BP has a similar or higher affinity for urocortin-I and urotensin-I compared with CRH (4, 27). Insect DH-I, with its CRH family motif, is the most likely candidate to bind to CRH-BP. Given the colocalization of CRH and CRH-BP in the pars intermedia of carp, it is obvious to assume that CRH-BP colocalizes with DH-I in the insect corpora cardiaca. Although we were unable to demonstrate so in the honeybee, this is indeed the case in the locust Schistocerca gregaria (De Loof, A., and M. O. Huisng, unpublished observation). Furthermore, the gene expression pattern of honeybee CRH-BP is consistent with these findings.

Although the pars intercerebralis-corpus cardiacum axis is considered the analog of the vertebrate hypothalamo-pituitary axis in general, the insect corpora cardiaca bear a morphological resemblance to the pituitary gland of fish in particular. Because fish lack a median eminence, their pituitary gland is directly contacted by hypothalamic neurons. Furthermore, the nerve
terminals in the pituitary pars intermedia of fish contain such an abundance of several neuropeptides, including CRH and CRH-BP (15), that it is considered a neurohemal site, analogous to the insect corpora cardiaca. A second neurohemal organ in fish, the caudal neurosecretory system, also releases CRH as well as urotensin-I (28). Interestingly, the latter peptide has a
role in osmoregulation (29), analogous to insect DH-I that acts distally on the Malpighian tubules within the insect abdomen to promote active cation transport, thereby increasing primary urine production (30).

In summary, we have demonstrated that CRH-BP is well conserved and is clearly identifiable in insect species. It follows that CRH-BP has been present since the common ancestor to insects and vertebrates. More importantly, the unequivocal orthology of insect and vertebrate CRH-BPs adds substantial weight to the supposition that the vertebrate CRH system and the insect DH system stem from a common ancestor.

Acknowledgments

We gratefully thank Mr. R. ten Kleijn of the Dutch beekeeping expertise center ‘Het Bijnheus’, for supplying us with honeybees and advice. We also thank Prof. A. De Loof and Mr. M. Timmermans for useful comments about the manuscript.

Received November 23, 2004. Accepted February 7, 2005. Address all correspondence and requests for reprints to: Dr. G. Flik, Department of Animal Physiology, Institute for Neuroscience, Radboud University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands. E-mail: G.Flik@science.ru.nl.

The honeybee CRH-BP and DH-I sequences have been submitted to the European Molecular Biology Laboratory database under accession numbers AJ780964 and AJ876408.

24. Gu X 1998 Early metazoan divergence was about 830 million years ago. J Mol Evol 47:369–371


date of 9/13/2005.

Endocrinology is published monthly by The Endocrine Society (http://www.endo-society.org), the foremost professional society serving the endocrine community.