The Journal of Experimental Biology 207, 4573-4586
Published by The Company of Biologists 2004
doi:10.1242/jeb.01317

Functional characterisation of the *Anopheles* leucokinins and their cognate G-protein coupled receptor

Jonathan C. Radford, Selim Terhzaz, Pablo Cabrero, Shireen-A. Davies and Julian A. T. Dow*

Institute of Biomedical and Life Sciences, Division of Molecular Genetics, University of Glasgow, Glasgow G11 6NU, UK

*Author for correspondence (e-mail: j.a.t.dow@bio.gla.ac.uk)

Accepted 29 September 2004

Summary

Identification of the *Anopheles gambiae* leucokinin gene from the completed *A. gambiae* genome revealed that this insect species contains three leucokinin peptides, named *Anopheles* leucokinin I-III. These peptides are similar to those identified in two other mosquito species, *Aedes aegypti* and *Culex salinarius*. Additionally, *Anopheles* leucokinin I displays sequence similarity to *Drosophila melanogaster* leucokinin.

Using a combination of computational and molecular approaches, a full-length cDNA for a candidate leucokinin-like receptor was isolated from *A. stephensi*, a close relative of *A. gambiae*. Alignment of the known leucokinin receptors – all G protein-coupled receptors (GPCRs) – with this receptor, identified some key conserved regions within the receptors, notably transmembrane (TM) domains I, II, III, VI and VII.

The *Anopheles* leucokinins and receptor were shown to be a functional receptor-ligand pair. All three *Anopheles* leucokinins caused a dose-dependent rise in intracellular calcium ($[Ca^{2+}]_i$) when applied to S2 cells co-expressing the receptor and an aequorin transgene, with a potency order of I>II>III.

***Drosophila* leucokinin was also found to activate the *Anopheles* receptor with a similar EC_{50} value to *Anopheles* leucokinin I. However, when the *Anopheles* peptides were applied to the *Drosophila* receptor, only *Anopheles* leucokinin I and II elicited a rise in $[Ca^{2+}]_i$. This suggests that the *Anopheles* receptor has a broader specificity for leucokinin ligands than the *Drosophila* receptor.**

Antisera raised against the *Anopheles* receptor identified a doublet of approx. 65 and 72 kDa on western blots, consistent with the presence of four *N*-glycosylation sites within the receptor sequence, and the known glycosylation of the receptor in *Drosophila*. In *Anopheles* tubules, as in *Drosophila*, the receptor was localised to the stellate cells.

Thus we provide the first identification of *Anopheles* mosquito leucokinins (*Anopheles* leucokinins) and a cognate leucokinin receptor, characterise their interaction and show that Dipteran leucokinin signalling is closely conserved between *Drosophila* and *Anopheles*.

Key words: mosquito, *Anopheles gambiae, Anopheles stephensi*, aequorin, calcium, leucokinin.

Introduction

The Malpighian (renal) tubules of insects are mission-critical organs (Beyenbach, 2003b; Dow and Davies, 2001) and as such are an excellent target tissue for the development of novel insecticides against agriculturally and biomedically relevant pest species. Not only do they regulate water and ion homeostasis, but they have a major role in detoxification. Thus, modulation of fluid transport by Malpighian tubules is an effective method for novel pest control strategies.

Insect leucokinins elicit potent diuretic effects on the Malpighian tubules of various insect species (Gade, 2004) and have also been implicated in a number of other physiological functions, such as the contraction of the hindgut (Holman et al., 1999; Howarth et al., 2002; Veenstra et al., 1997) – hence the term 'myokinin'. In addition, recent studies have suggested the involvement of leucokinins in dietary regulation and energy mobilisation (Nachman et al., 2002; Seinsche et al., 2000). As such, the insect leucokinins have attracted a great deal of interest as lead molecules for novel pesticides, including the development of peptidase-resistant analogues of this family of peptides (Teal et al., 1999). Given the diverse roles of insect leucokinins, elucidation of the mode of action of these peptides via their cognate G protein-coupled receptors (GPCRs) is of importance. Furthermore, as leucokinins have only been found in invertebrates, it is likely that careful design of leucokinin antagonist or agonist analogues will avoid interactions with mammalian species.

While identification of leucokinins and their cognate receptors has been successfully undertaken in some insects (Holman et al., 1984; Holman et al., 1999; Veenstra et al., 1997), including the genetically tractable Dipteran, *Drosophila melanogaster* (Radford et al., 2002; Terhzaz et al., 1999), less

progress has been made in studies of leucokinin signalling in biomedically relevant insects.

The malaria mosquito, *Anopheles gambiae*, is one such insect. Initial attempts to curb the spread of malaria involved the use of larvicides and insecticides, against the mosquito vectors, and also the use of chloroquine, which halts the progression of the disease in patients. Despite these efforts, resistance has evolved in both the mosquitoes and in the malaria parasites. Thus malaria, as other vector-borne diseases, is now classed as a re-emerging disease (Gubler, 1998). However, the sequencing of the *A. gambiae* genome (Holt et al., 2002) provides a fresh direction for anti-malarial research. The action of leucokinins on the Malpighian tubules of the yellow fever mosquito, *Aedes aegypti*, has already been studied in great detail (Beyenbach, 2003b; Veenstra et al., 1997). Thus in *Anopheles*, it is likely that leucokinins will also play an important role in the regulation of water and ion homeostasis. An initial survey of the completed *Anopheles* genome identified a single leucokinin-like gene (Riehle et al., 2002). Approximately 37 neurohormone receptor-like-encoding sequences were also identified in a survey of the GPCR repertoire of the *Anopheles* genome (Hill et al., 2002). It is likely that one of these will represent a receptor for *Anopheles* leucokinins.

In this study we identified three leucokinin peptides (*Anopheles* leucokinins I-III) from the *A. gambiae* genome and demonstrated effects on calcium signalling via a putative cognate GPCR-coupled receptor cloned from its close relative, *A. stephensi*. All *Anopheles* leucokinins increase intracellular calcium in a dose-dependent manner. Furthermore, the *Anopheles* leucokinin receptor is responsive to *D. melanogaster Drosophila* leucokinin, while the *D. melanogaster* leukokinin-like receptor (LKR) is only sensitive to *Anopheles* leucokinins I and II.

Materials and methods

Bioinformatics

Identification of *Anopheles gambiae* Miles leucokinins: TBLASTN analysis was carried out on the completed *A. gambiae* genome using the predicted protein product of the *Aedes aegypti* L. preproleucokinin transcript (Veenstra et al., 1997) (GenBank accession no. AAC47656). The BLOSUM62 matrix (default settings) was used for all BLAST analysis. A 333 bp sequence was identified, which potentially encoded a protein with homology to the *Aedes* preproleucokinin (E-value 2×10^{-29}). This region plus the surrounding 20 kb of genomic sequence either side was then analysed with the Softberry FgenesH gene prediction program (www.softberry.com).

Identification of *A. gambiae* leucokinin receptor: TBLASTN analysis was carried out on the completed *Anopheles gambiae* genome using the predicted protein product of the *Drosophila* leucokinin receptor, CG10626 (Radford et al., 2002). The protein sequences of the *Lymnaea stagnalis* L. (Cox et al., 1997) and *Boophilus microplus* Canestrini (Holmes et al., 2003) leucokinin-like receptors were also used to confirm the sequence match. The BLOSUM62 matrix (default settings) was used for all BLAST analysis as above. This identified a sequence within the *Anopheles* genome that encoded a putative protein of 377 amino acids (GenBank accession no. agCP10499, E-value 2×10^{-59}). No other good sequence matches were identified in the completed genome sequence. However, this sequence only represented a portion of a presumed GCPR, encompassing only the strictly conserved TM domains. Efforts were made to identify a full-length transcript by the use of the FgenesH gene prediction program, although subsequent attempts to amplify the putative open reading frame (ORF) region by polymerase chain reaction (PCR) from *Anopheles* cDNA failed. Therefore, a RACE approach was undertaken in order to identify the correct cDNA sequence.

Insects

A. stephensi Liston and its close relative *A. gambiae* are malaria-carrying anopheline mosquitoes. For reasons of availability, *A. stephensi* was used as a source of cDNA in this study. Non-infective, sugar-water-fed adults were a kind gift from Dr L. Ranford-Cartwright, University of Glasgow, UK. Female animals were used upon receipt. If mosquitoes were not used immediately, they were maintained over a 12 h:12 h L:D photoperiod at 55% humidity at 22°C, on 5% sucrose (v/v) solution *ad libitum* for a maximum of 3 days before use in experiments.

RT-PCR of putative leucokinin receptor

For cDNA preparations, total RNA was extracted (Sigma Tri-reagent, Gillingham, Dorset, UK) from whole *A. stephensi* and reverse transcribed with Superscript II (Invitrogen). 1 µl of the reverse transcription reaction was used as a template for PCR containing the gene-specific primer pairs given below. Additionally, to control against genomic contamination in cDNA preps, primers were used that had been designed around intron/exon boundaries of the predicted *A. gambiae* leucokinin receptor gene. Use of such primers verified the quality of the cDNA used in PCR reactions. Further controls were performed, which included non-reverse transcribed template (i.e. no cDNA). The primers used were: GGAATCTGCCC-GAGTTTATGTG and GTTCTTCAGCATCGTAATGTCGC. PCR cycle conditions for reactions were as follows: 93°C 3 min; 36 cycles of (93°C 30 s, 59°C 30 s, 72°C 1 min); 72°C (1 min). PCR products obtained from such RT-PCR experiments were cloned into pCRII-TOPO™ using the Invitrogen Topoisomerase (TOPO TA Cloning) system. Cloned plasmids were purified using Qiagen kits and sequenced to confirm their identity.

5′-RACE and 3′-RACE of putative leucokinin receptor

Poly(A)$^+$ RNA was purified from whole fly total RNA using the magnetic Dynabeads® mRNA purification kit (Dynal® Bromborough, UK) according to the manufacturers' instructions. The RACE procedure was carried out using the SMART™ RACE cDNA Amplification kit (Clontech, Oxford,

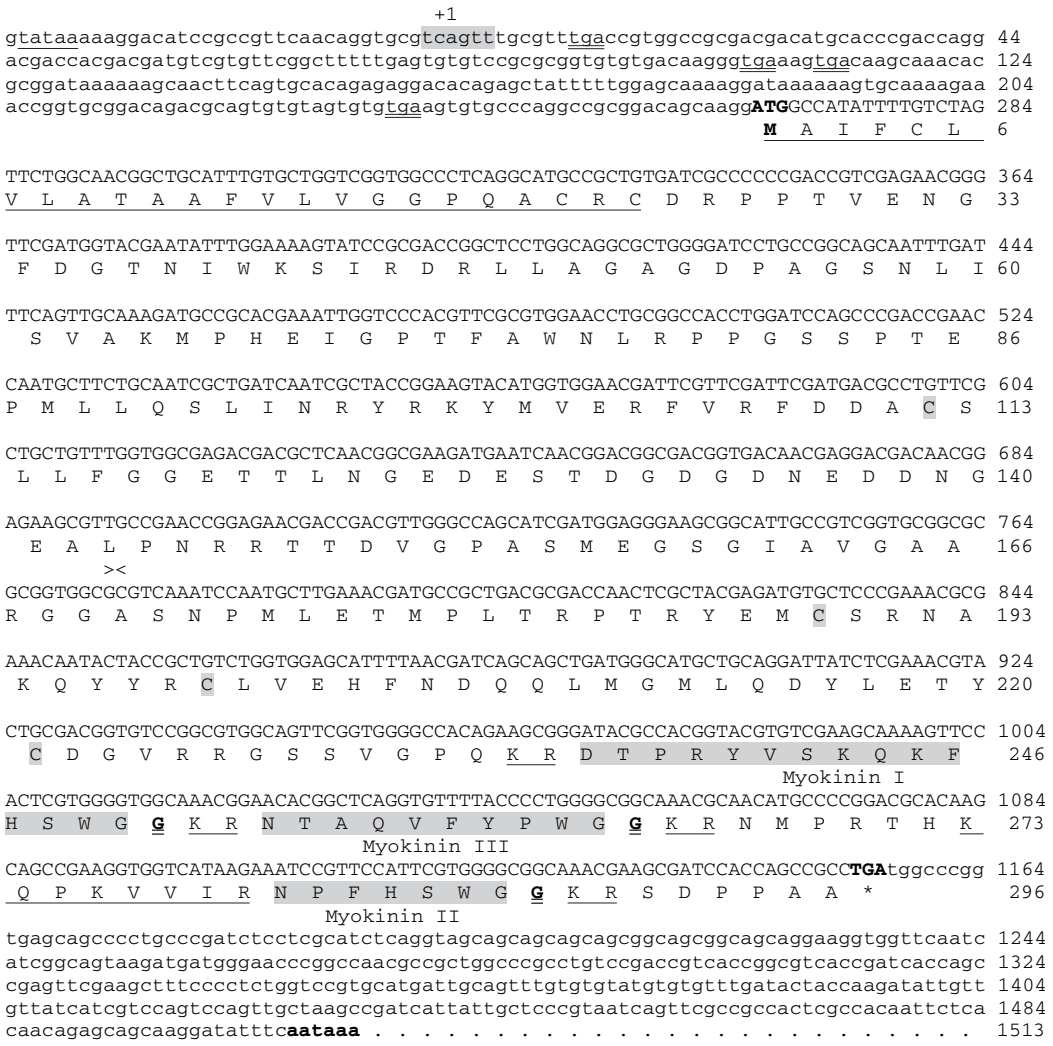


Fig. 1. The cDNA and protein sequences of the putative *Anopheles gambiae* preproleucokinin gene. Nucleotide and amino acid residue numbers are indicated at the end of each line, amino acids are centred on their codons. The start codon, following stop codon and putative polyadenylation signal are indicated in bold type. Upstream in-frame stop codons are double underlined. A putative TATA-box sequence is underlined, the initiation of transcription consensus sequence is indicated by a shaded box, and the position of the single intron is indicated by ><. Within the putative preproleucokinin the location of the three *A. gambiae* leucokinin peptides are indicated by shaded boxes and a possible signal peptide is underlined. Proteolytic processing sites thought to be used are underlined. The Gly (G) residues that are presumed to be processed to C-terminal amides in the mature *A. gambiae* leucokinins are indicated in bold and double underlined. Four Cys (C) residues within the preproleucokinin are also indicated by shaded boxes.

UK). This kit provides a method for performing both 5′- and 3′-RACE. 5′- and 3′-RACE-ready cDNAs are generated as separate cDNA samples, using 1 μg poly(A)$^+$ mRNA as starting material for each of the 5′- and 3′-RACE-ready cDNAs.

SMART™ RACE PCR reactions were carried out according to the manufacturers' instructions using Advantage 2 Polymerase Mix (Clontech). Both 5′- and 3′-RACE reactions were set up according to the protocol, using 200 nmol l^{-1} gene-specific primer and 2.5 μl RACE-ready cDNA in the appropriate reaction mix. Gene-specific primers were carefully designed in such a way that they had the following characteristics: 23–28 nucleotides, 50–70% GC, T_m≥70°C. To perform 5′-RACE PCR, an antisense primer was designed, and for 3′-RACE PCR a sense primer was designed. Primers were situated as close as possible to the end of known cDNA sequence in order to keep the size of RACE products to a minimum.

Designing primers with a T_m≥70°C allowed the use of touchdown PCR to improve the specificity of the amplification. This method uses an annealing temperature during the initial PCR cycles that is higher than the T_m of the universal primer, allowing only gene-specific synthesis during these cycles. Cycling was performed in thin-walled dome-topped 0.2 ml PCR tubes in a Hybaid PCR Express-Gradient thermocycler. This was performed as follows: 94°C, 3 min; 5 cycles of 94°C 5 s, 72°C 3 min; 5 cycles of 94°C 5 s, 70°C 10 s, 72°C, 3 min; 20–25 cycles of 94°C 5 s, 68°C 10 s, 72°C 3 min. Note that the

Drosophila leucokinin	**NS**VVLGK**KQ**R**FHSWG**amide
Aedes leucokinin I	**NS**-KYVS**KQ**K**F**Y**SWG**amide
Anopheles leucokinin I	DTPRYVS**KQ**K**FHSWG**amide
Culex leucokinin III	TKYVS**KQ**-**F**F**SWG**amide
Aedes leucokinin II	NP**FH**A**WG**amide
Anopheles leucokinin II	NP**FHSWG**amide
Culex leucokinin I	NP**FHSWG**amide
Aedes leucokinin III	NNPN-**F**YP**WG**amide
Anopheles leucokinin III	NTA**Q**V**F**YP**WG**amide
Culex leucokinin II	NNANV**F**YP**WG**amide

Fig. 2. Comparison of the fruit fly and mosquito leucokinin peptides. The sequences of the *Drosophila melanogaster* (*Drosophila* leucokinin), *Aedes aegypti* (*Aedes* leucokinin), *Culex salinarius* (*Culex* leucokinin) and *Anopheles gambiae* (*Anopheles* leucokinin) leucokinin peptides are compared. Residues conserved with *Drosophila* leucokinin are indicated in bold. Analysis of the *D. pseudoobscura* genome sequence suggests that a single leucokinin identical to that of *D. melanogaster* is encoded.

extension time is dependent on the length of the fragment being amplified; 3 min is suitable for cDNA fragments of 2–4 kb.

RACE products were then separated by agarose gel electrophoresis under standard conditions and individual products gel-purified. RACE products were then directly cloned into pCRII-TOPO™ vector and individual clones analysed by restriction enzyme digestion and automated sequencing.

Expression of Anopheles *leucokinin receptor in S2 cells*

The ORF of the *A. stephensi* leucokinin-like receptor was amplified using the primers GCCCAGAAGAAATCATGC-AAGCAACAG and GCAAAACAGCTCACAGTTAATA-CACATTGCTCG, and *A. stephensi* whole fly cDNA as template (see Fig. 3). This was cloned into the pMT/V5-His TOPO® vector, and the correct orientation determined by restriction enzyme digestion. Constructs were then sequenced to confirm error free cloning of the ORF. The amplification product included the native stop codon to prevent inclusion of the C-terminal V5-His peptide in the expressed protein. S2 cells, cultured under standard conditions (Radford et al., 2002) were transiently transfected with the *apoaequorin* ORF (Radford et al., 2002) and the *A. stephensi* leucokinin-like receptor ORF constructs, and expression induced using Cu^{2+} (Radford et al., 2002).

Peptide synthesis

The three putative *Anopheles* leucokinins identified in this work were synthesised as C-terminally amidated peptides (Research Genetics/Invitrogen Inc.). Peptides were dissolved in H_2O to a concentration of 1 mmol l^{-1} and then diluted to the required working concentration in Schneider's medium supplemented with 10% foetal calf serum (FCS; Invitrogen Inc.).

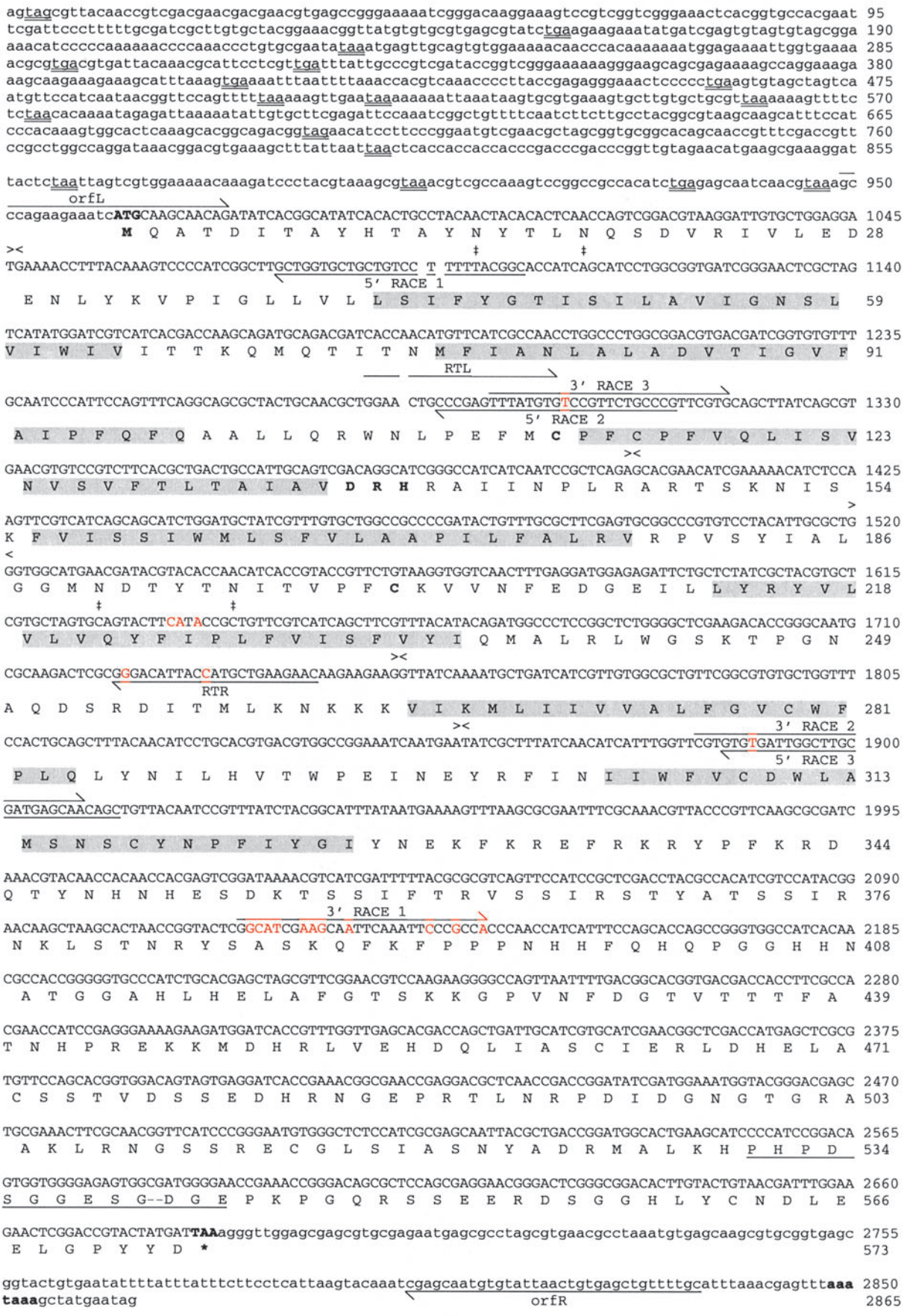
Measurements of intracellular Ca^{2+} using aequorin

Transfected S2 cells were harvested and incubated with 2.5 μmol l^{-1} coelenterazine in the dark at room temperature (RT) for 1–2 h (Radford et al., 2002). 25,000 cells were then placed in 135 μl Schneider's medium supplemented with 10% FCS in a well of a white polystyrene 96-well plate (Berthold Technologies, Redbourn, UK). Bioluminescence recordings were carried out using a Mithras LB940 automated 96-well plate reader (Berthold Technologies) and MikroWin software. 15 μl of each of the *Anopheles* leucokinin peptides was applied at the required concentration. At the end of each recording samples were disrupted by the addition of 100 μl lysis solution, and the Ca^{2+} concentrations calculated as previously described (Rosay et al., 1997).

Generation of antibodies against Anopheles *leucokinin receptor and immunolocalisation of the receptor*

Rabbit anti-peptide antibodies were raised against the epitope PHPDSGGESGGDGE (residues 531–543; Genosphere Technologies, Paris, France). An N-terminal cysteine residue was incorporated to permit conjugation to bovine serum albumin (BSA). The antiserum to *Anopheles* leucokinin receptor showed some background immunoreactivity and, therefore, was purified on a HiTrap Protein A HP column (Amersham Pharmacia Biotech, Little Chalfont, UK) according to the manufacturer's instructions. The protocol used for immunohistology was as described previously (Radford et al., 2002). Briefly, the IgG to *Anopheles* leucokinin receptor was diluted 1:1000 or the pre-immune serum diluted 1:500. Primary antibody incubations were performed overnight. A Texas Red-conjugated affinity-purified goat anti-rabbit antibody (Jackson Immunologicals, Westgrove, PA, USA) was used at a dilution of 1:1000 for visualization of the primary antiserum. Prior to mounting on slides, tubules were stained with 1 μg ml^{-1} of 4′,6′-diamidino-2-phenylindole hydrochloride (DAPI; Sigma-Aldrich, Gillingham, UK). Slides were viewed using a Zeiss 510 META confocal microscope and images were processed with a Zeiss LSM 5 Image Browser.

Fig. 3. The cDNA and putative protein sequence of the *Anopheles stephensi* leucokinin receptor gene. Nucleotide and amino acid residue numbers are indicated at the end of each line, amino acids are centred on their codons. The polyadenylation signal, and start and stop codons are marked in bold type. Upstream in frame stop codons are double underlined, and the positions of predicted introns in the cDNA sequence are indicated by ><. Predicted TM regions are indicated by a shaded box. The peptide epitope used to raise the anti-CG10626 antibody (Radford et al., 2002) is not present within the *A. stephensi* receptor sequence. ‡Potential *N*-glycosylation sites in the protein sequence. The conserved GPCR Asp-Arg-Tyr/His ($DR^{Y}/_{H}$) motif just after TM III and conserved Cys (C) residues in extracellular loops 1 and 2 are also marked in bold type. The position of binding sites for primers used in the analysis and construction of the cDNA sequence are indicated. Nucleotides that are different between the cDNA and primer sequences are indicated in *red* type.


orfL
5′ RACE 1
RTL
3′ RACE 3
5′ RACE 2
RTR
3′ RACE 2
5′ RACE 3
3′ RACE 1
orfR

Western blot analysis

Protein samples were prepared from tubule or head tissues by homogenization in ice-cold Tris lysis buffer (20 mmol l^{-1} Tris, pH 7.5, 250 mmol l^{-1} sucrose, 2 mmol l^{-1} EDTA, 100 mmol l^{-1} NaCl, 50 mmol l^{-1} β-mercaptoethanol, 2% (w/v) SDS) with protease inhibitor cocktail (P-8340, Sigma). Samples were centrifuged for 10 min at 13 000 ***g*** at 4°C to remove debris. Supernatants were removed to a clean tube and assayed for protein concentration (Lowry protein assay). 15 μg of each sample were run on SDS-PAGE and blotted according to standard methods. The filter was blocked for 3 h in PBS with 0.1% Tween 20 and 10% non-fat dry milk and washed in PBS/Tween 20 once for 5 min. The filter was incubated for 3 h at RT with IgG to *Anopheles* leucokinin receptor, diluted 1:1000 (or the pre-immune serum diluted 1:500) in PBS/Tween 20/milk, washed in PBS/Tween 20 three times for

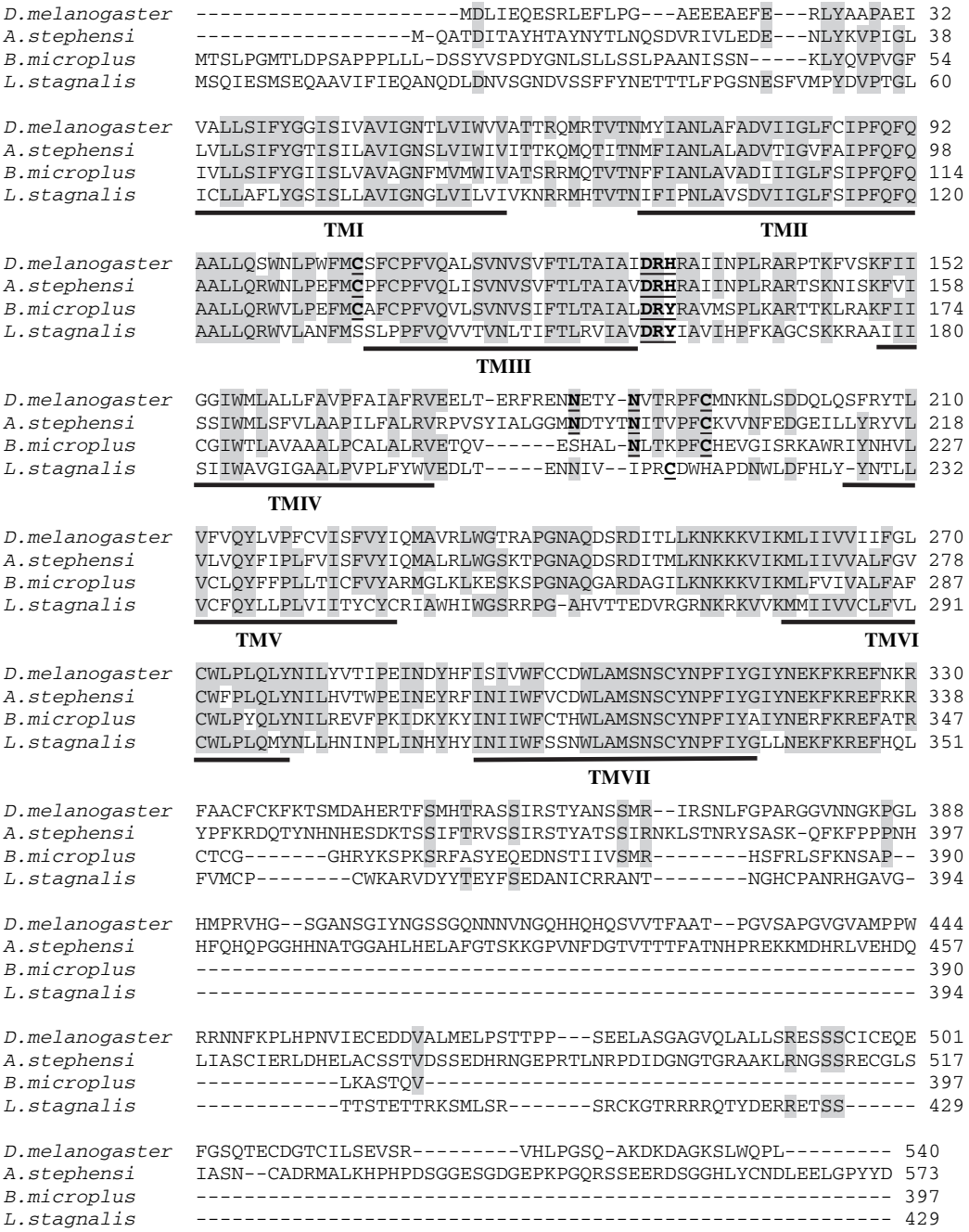


Fig. 4. Alignment of the known leucokinin receptor protein sequences. Amino acid residue numbers are indicated at the end of each line. The predicted TM domains are underlined. TM domains were predicted using the TMHMM 2.0 prediction program. Identical residues with a threshold limit of 75% are indicated by a shaded box. Sequence alignments were performed using CLUSTAL X, and annotated using BioEdit. Key conserved residues, such as the Asp-Arg-Tyr/His ($DR^{Y}/_{H}$) triplet motif after TM III, Cys (C) residues in the second and third extracellular loops and potential *N*-glycosylation sites are in bold type and underlined.

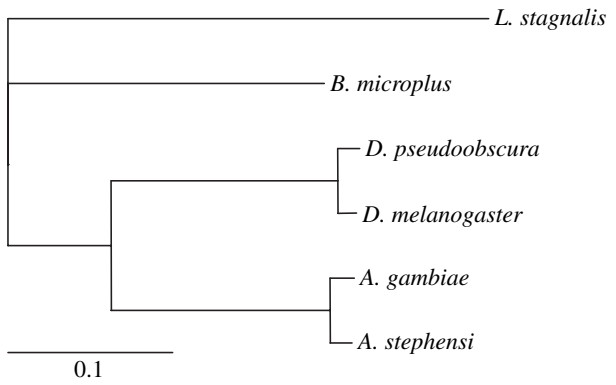


Fig. 5. Dendrogram of the known leucokinin receptors. A CLUSTAL X protein alignment was performed using the putative TM spanning regions of the known leucokinin receptors. From this a phylogram was produced using the TREEVIEW program. TM domains were predicted using the TMHMM 2.0 program. The scale bar gives an approximation of the number of substitutions per site. The *Lymnaea stagnalis* (pond snail) receptor is used as an outgroup.

10 min, and incubated for 1 h with secondary antibody (1:5000 horseradish peroxidase-labelled anti-rabbit IgG antibody; Amersham Biosciences) diluted in PBS/Tween 20/milk. The filter was then washed in PBS for 1 h and protein bands visualized using enhanced chemiluminescence (ECL, Amersham Biosciences).

Statistics

Where appropriate, statistical significance was assessed using Student's *t*-test for unpaired samples, taking $P<0.05$ as the critical value.

Results

Identification of the A. gambiae *leucokinins*

The Softberry FgenesH gene prediction program predicted a gene contained within a 2179 bp genomic region, including a single intron of 629 bp. Further inspection revealed a transcript of 1545 bp in length, containing an ORF of 891 bp, a 265 bp 5′-UTR and a 3′-UTR of approximately 385 bp (Fig. 1). This would appear to be a full mature transcript, as it contains a perfect consensus sequence for the initiation of transcription (–2 to +4, TCAGTT), and a polyadenylation signal consensus sequence in the 3′-UTR (1508–1513, AATAAA). There is also a TATA box motif upstream of the initiation of transcription site (–30 to –34, TATAA). There is a single presumed ATG start codon at positions 266–268, preceded by four in-frame stop codons.

The ORF of the *A. gambiae* leucokinin gene encodes a predicted protein of 296 amino acids (Fig. 1). Analysis using the PROSITE program (Bairoch et al., 1997) identified a putative 24-amino-acid signal peptide. Within the remaining protein sequence there are three leucokinin peptides predicted, which have been named according to their similarity to the three known *Aedes aegypti* leucokinins. The *Anopheles* leucokinins I and III (*Anopheles* leucokinins) are flanked by dibasic proteolytic cleavage sites, and in all three peptides a C-terminal Gly is present, which is predicted to be processed into a C-terminal amide group in the mature peptides. A different proteolytic cleavage site is present at the N terminus of *Anopheles* leucokinin II, consisting of a single Arg with a Lys at residue –8. Four Cys residues are also present in the protein region before to the leucokinin peptide sequences. The position of the four Cys residues is identical in *Aedes*, suggesting that these residues may play an important role in the function of the precursor protein, perhaps in the formation of disulphide bridges (Veenstra et al., 1997). It has been proposed that these residues are responsible for paraldehyde-fuchsin staining observed in the leucokinin-immunoreactive neuroendocrine cells of the abdominal ganglion in hemimetabolous insects (Veenstra et al., 1997). Owing to the conservation of this staining between insect species, it was suggested that this

Table 1. *Percentage identity and similarity of the protein sequences of the leucokinin receptors*

(A) TM domains

Species	*A. gambiae*	*D. melanogaster*	*D. pseudoobscura*	*B. microplus*	*L. stagnalis*
A. stephensi	97/99	70/82	70/82	60/75	47/66
A. gambiae	–	71/82	70/82	60/76	47/65
D. melanogaster	–	–	97/99	59/75	48/67
D. pseudoobscura	–	–	–	59/75	48/67
B. microplus	–	–	–	–	51/65

(B) Entire proteins

Species	*D. melanogaster*	*B. microplus*	*L. stagnalis*
A. stephensi	45/56	33/44	26/38
D. melanogaster	–	34/43	28/41
B. microplus	–	–	38/51

Percentage identity and similarity were calculated using the BioEdit software package based on sequence alignments carried out using CLUSTAL X, and were scored on the BLOSUM62 matrix.

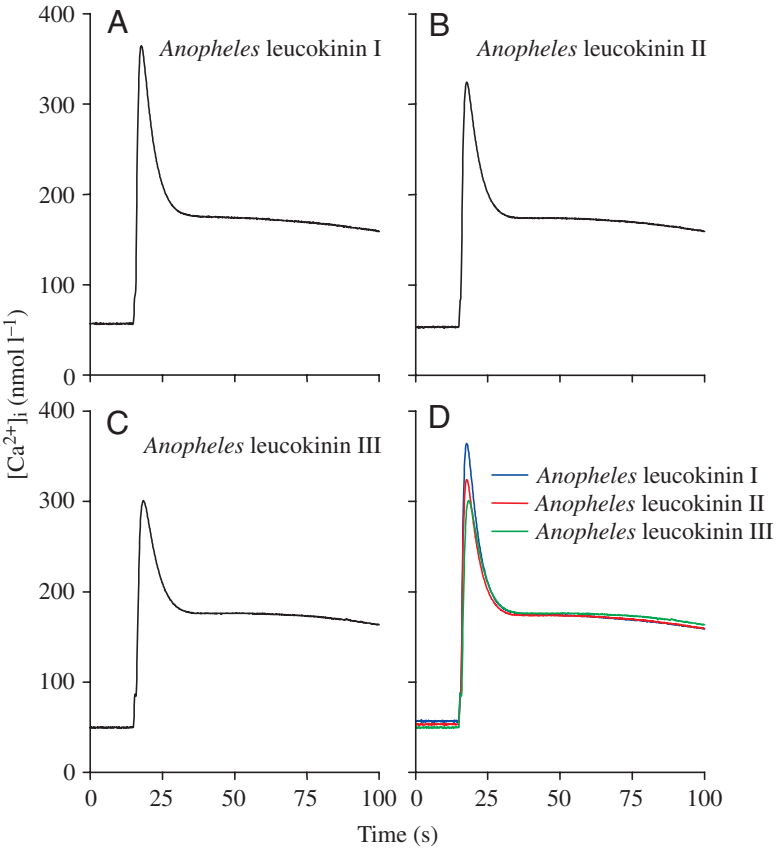


Fig. 6. Real-time measurement of the $[Ca^{2+}]_i$ response by *A. stephensi* leucokinin receptor. S2 cells were co-transfected with the *A. stephensi leucokinin receptor* ORF and *apoaequorin* ORF constructs, and expression induced. Peptide was injected at 15 s. Samples were stimulated with either *Anopheles* leucokinin I (A), II (B) or III (C) peptide at a concentration of 10^{-7} mol l^{-1}. (D) Comparison of the responses to the three *Anopheles* leucokinin peptides applied at a concentration of 10^{-7} mol l^{-1}. Data are expressed as $[Ca^{2+}]_i$ (nmol l^{-1}) against time (s); measurements were taken at 0.1 s intervals. The traces shown are average responses ± S.E.M. (N=8). Error bars are negligible for all panels.

region of the precursor protein would be conserved between species. However, alignment of the *Anopheles* and *Aedes* leucokinin precursors revealed very little other sequence conservation within this region.

Using the same technique, a leucokinin-like precursor gene was also identified in the available sequence of the *Drosophila pseudoobscura* genome. The *D. pseudoobscura* sequence encodes a putative protein of 176 amino acids, containing only a single leucokinin peptide sequence, identical to *Drosophila* leucokinin from *D. melanogaster*. A full-length mature transcript could not be reliably identified for *D. pseudoobscura*, although alignment of the protein with that from *D. melanogaster* (data not shown) suggested that the entire ORF had been identified.

Comparison of the sequence of the Anopheles *leucokinin peptides*

Anopheles leucokinin I is 15 residues in length, equal to *Drosophila* leucokinin, the longest leucokinin known to date (Fig. 2). It is also similar in sequence, being identical to *Drosophila* leucokinin in the bioactive C-terminal pentapeptide –Phe–His–Ser–Trp–Gly–amide. Unsurprisingly, *Anopheles* leucokinin is most similar across its entire length to *Aedes* leucokinin I, being identical at 10 residues. Significant similarity can also be seen to another mosquito leucokinin, culekinin III. Similar to *Aedes aegypti* and *Culex salinarius*, there are three leucokinins present in *Anopheles*. The shortest of the three, *Anopheles* leucokinin II (7 residues), is identical to culekinin I, with only one residue different from *Aedes* leucokinin II (Fig. 2). In addition, the C-terminal pentapeptide is identical to *Drosophila* leucokinin and *Anopheles* leucokinin I. The third peptide, *Anopheles* leucokinin III, is 10 residues in length, 1 longer than *Aedes* leucokinin III, but equal to culekinin II (Fig. 2). The C-terminal core is more divergent in *Anopheles* leucokinin III, although it retains the essential –Phe–X_1–X_2–Trp–Gly–amide motif, the His being replaced by a Tyr and the Ser being replaced by a Pro. This is identical to the C-terminal cores of *Aedes* leucokinin III and culekinin II, although less similarity is seen in the more N-terminal residues.

Insect leucokinin gene families

This characterisation of a leucokinin peptide-encoding gene within the *A. gambiae* genome shows that, as in the yellow fever mosquito, *Aedes aegypti* (Veenstra et al., 1997), three leucokinin peptides are encoded by a single transcript. Three leucokinins have also been identified in *Culex salinarius* (Meola et al., 1998), suggesting that this may be a conserved feature among mosquitoes. By contrast, only one leucokinin is found in *Drosophila melanogaster* (Terhzaz et al., 1999), *Drosophila pseudoobscura* (this work) and *Musca domestica* (Coast et al., 2002). This division reflects Dipteran taxonomy: *Anopheles*, *Aedes* and *Culex* are all members of the family Culicidae, whereas *Drosophila* and *Musca* are both Schizophora. The two groups thus diverge at the suborder level (Nematocera and Brachycera, respectively); so it is possible that both at the gene organisation level and their proposed modes of action (Beyenbach, 1998; Dow and Davies, 2003), the leucokinins may plausibly differ between mosquitoes and *Drosophila*.

In non-Dipteran insects, the numbers of known leucokinins vary widely: three leucokinins have been isolated from the moth *Helicoverpa zea*; eight from *Leucophaea maderae* and five from *Acheta domesticus* and *Periplaneta Americana* (Torfs et al., 1999). It would be interesting to determine whether the leucokinin peptides from non-mosquito species are also contained within one precursor protein.

The putative Anopheles *leucokinin receptor is a GPCR*

Having identified a putative leucokinin receptor from the *A. gambiae* genome, RT-PCR primers were designed within

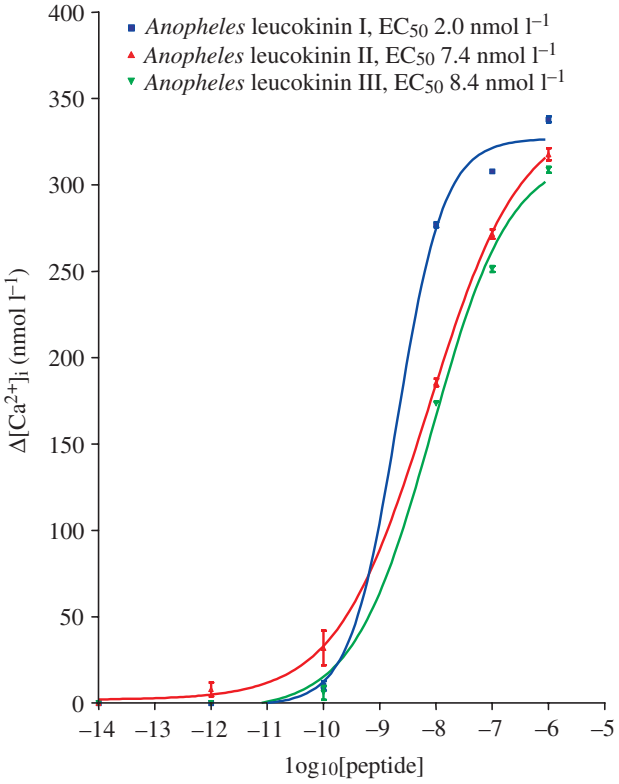


Fig. 7. Dose–response curves for the action of the *Anopheles* leucokinins on the *A. stephensi* leucokinin receptor. S2 cells were co-transfected with the *A. stephensi leucokinin receptor* ORF and *apoaequorin* ORF constructs, and expression induced. Peptide-stimulated $[Ca^{2+}]_i$ increases were measured in S2 cell aequorin-based assays, at different concentrations of peptide as indicated. Values were expressed as maximal $[Ca^{2+}]$–background $[Ca^{2+}]$ (nmol l^{-1}; mean ± S.E.M., N=5–8). Where error bars are not visible they are too small to reproduce.

regions of high sequence similarity to *CG10626* and PCR carried out on *A. stephensi* cDNA. Successful PCR showed bands of expected product sizes for *A. gambiae* cDNA and genomic DNA (470 bp and 632 bp, respectively; data not shown). Cloning and sequencing of these bands confirmed that the identified *Anopheles* sequence is expressed. Primers were then designed within the identified gene sequence to carry out 5′-RACE and 3′-RACE analysis to determine the full transcript sequence of the putative leucokinin receptor gene. Three primers were designed for each direction and RACE-ready cDNAs were prepared from *A. stephensi* whole fly poly(A)$^+$ mRNA. Discrete products were amplified using the AnLKR 5′ RACE 1 (~1.2 kb), 2 (~1.4 kb) and 3 (~1.9 kb), and 3′ RACE 2 (~1.1 kb) and 3 (~1.8 kb) primers (data not shown). Amplified products were gel extracted, cloned into the pCRII-TOPO® vector and sequenced in full. From this information the full transcript of the putative *A. stephensi* leucokinin receptor gene was assembled.

Analysis of the sequence identified from 5′-RACE and 3′-RACE experiments suggests that the sequence of the full mature transcript for this gene has been identified. It is a 2684 bp transcript containing the coding sequence of 1722 bp, a 962 bp 5′-UTR and a 181 bp 3′-UTR. The sequence (bases 1–4) contains a portion of a consensus sequence for the initiation of transcription. The lack of known genomic sequence for *A. stephensi* precluded the analysis of upstream sequence in order to identify further regulatory sequences. However, no downstream promoter element (DPE)-dependent promoter sequences could be identified in the transcript sequence, as has been found in the *Drosophila* genomic sequence for LKR (Radford et

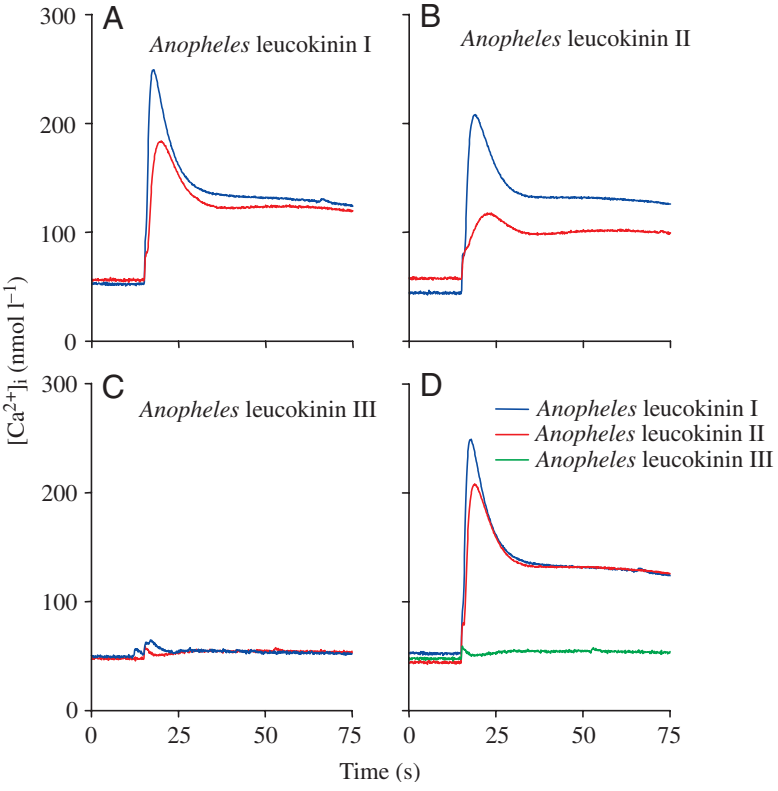


Fig. 8. Real-time measurement of the $[Ca^{2+}]_i$ response in S2 cells expressing *D. melanogaster* LKR, *CG10626*. S2 cells were co-transfected with the *D. melanogaster* LKR, *CG10626* ORF (Radford et al., 2002) and *apoaequorin* ORF constructs, and expression induced. Data are expressed as $[Ca^{2+}]_i$ (nmol l^{-1}) against time (s); measurements were taken at 0.1 s intervals. The traces shown are average responses (N=5). Peptide was injected at 15 s. (A–C) Samples were stimulated with either *Anopheles* leucokinin I (A), II (B) or III (C) peptide at a concentration of 10^{-6} mol l^{-1} (blue) or 10^{-7} mol l^{-1} (red). (D) Comparison of the responses to the three *Anopheles* leucokinin peptides applied at a concentration of 10^{-6} mol l^{-1}. Error bars are negligible for all panels.

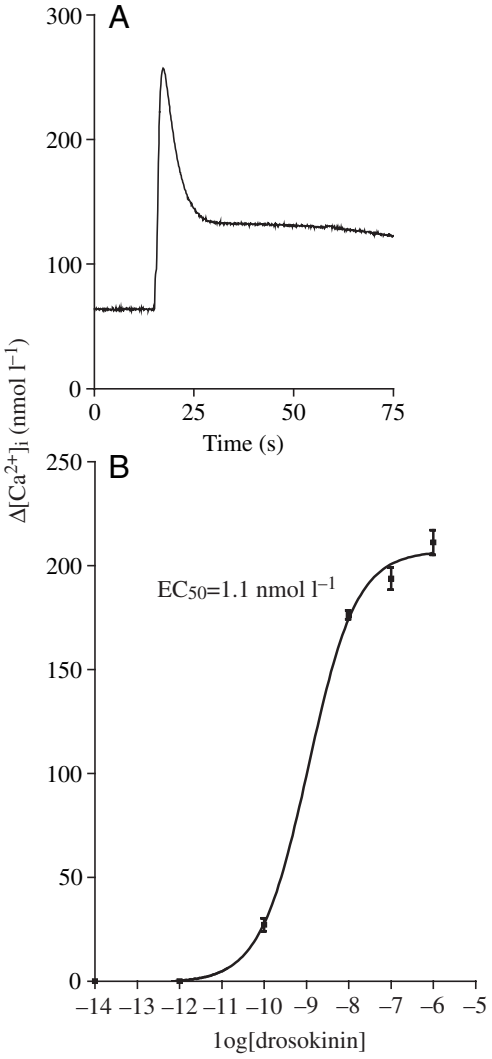


Fig. 9. Cross-specific activation of the *A. stephensi* leucokinin receptor with *Drosophila* leucokinin (drosokinin). (A) Real-time measurements of $[Ca^{2+}]_i$ in S2 cells expressing the *A. stephensi* leucokinin receptor and apoaequorin in response to *Drosophila* leucokinin. Data are expressed as $[Ca^{2+}]_i$ (nmol l^{-1}) against time (s); measurements were taken at 0.1 s intervals. The trace shown is an average response (N=5). Peptide was injected at 15 s. (B) Dose–response curve. Values were expressed as maximal $[Ca^{2+}]$–background $[Ca^{2+}]$ (nmol l^{-1}; mean ± S.E.M., N=5). Where error bars are not visible they are too small to reproduce.

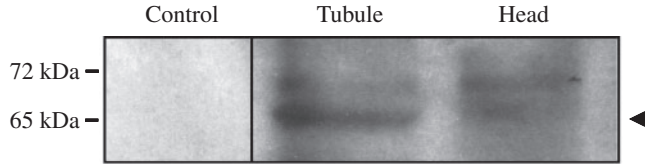


Fig. 10. Western blot analysis of the *Anopheles* leucokinin receptor. Western blot of adult Malpighian tubule and head proteins using *Anopheles* leucokinin receptor purified IgG. The antibody recognises both a protein of the expected size (65 kDa) and a heavier band, of approximately 72 kDa. The control lane is blotted with pre-immune serum.

al., 2002). A single ATG start codon was identified, beginning at base 963 and terminating at a stop codon at position 2682. Upstream of this presumed start codon there are 17 in-frame stop codons. A polyadenylation signal is also present within the 3′-UTR of the sequence (AATAAA, 18 bp from the polyadenylation site). By alignment of the *A. stephensi* transcript with the *A. gambiae* genome it is likely that at least six introns are contained within this gene. There is less conservation of the nucleotide sequence within the 5′-UTR, and so the presence of additional introns within this region cannot be ruled out. Again, by inference from the *A. gambiae* genome, the transcript is thought to be contained within approximately a 7.5 kb genomic region.

The ORF of the *A. stephensi* leucokinin-like receptor transcript encodes a 574 amino acid protein, which has an estimated molecular mass of 65 kDa. Analysis using the TMHMM program (Krogh et al., 2001) suggests that this predicted protein exhibits the conserved 7 TM domain structure, consistent with it being a functional GPCR (Fig. 3). Other conserved GPCR motifs are also present, such as a triplet motif Asp-Tyr-His at residues 136–138, just downstream of the putative third TM domain. Also two conserved Cys residues, Cys^{112} and Cys^{201}, located in the first and second extracellular loops respectively, are predicted to form a disulphide bond. There are also four potential *N*-glycosylation sites within the protein sequence, Asn^{14} and Asn^{18} in the N-terminal region, and Asn^{190} and Asn^{195} in the putative second extracellular loop. Interestingly, a difference exists between the C-terminal domain of the *A. stephensi* leucokinin receptor and that of the *D. melanogaster* LKR: it does not contain the epitope used to raise the anti-CG10626 (*D. melanogaster* LKR) antibody (Radford et al., 2002).

Alignment and comparison of the known leucokinin receptors

The protein sequences of the known leucokinin-like receptors, the *Drosophila* LKR (CG10626; Radford et al., 2002), the lymnokinin receptor (GenBank accession AAD11810; Cox et al., 1997), the *B. microplus* receptor (AAF72891; Holmes et al., 2003) and the putative *A. stephensi* receptor (Fig. 3) were aligned using the CLUSTAL X program (Thompson et al., 1994). The sequence alignment was annotated using BioEdit (Hall, 1999) (Fig. 4). The alignment demonstrates that there is considerable similarity between the four protein sequences, particularly within the TM domain-containing regions, with the N- and C-terminal regions being more divergent (Fig. 4). However, the sequence similarity is not as high within TM domains IV and V. The size and spacing of the TM domains is also consistent between the proteins, with only the second extracellular loop being variable in size. Interestingly, the first extracellular loop also appears highly conserved within these proteins, suggesting possible involvement in ligand binding. Several key residues are also conserved. A typical GPCR triplet motif is present immediately after the third TM domain as either an

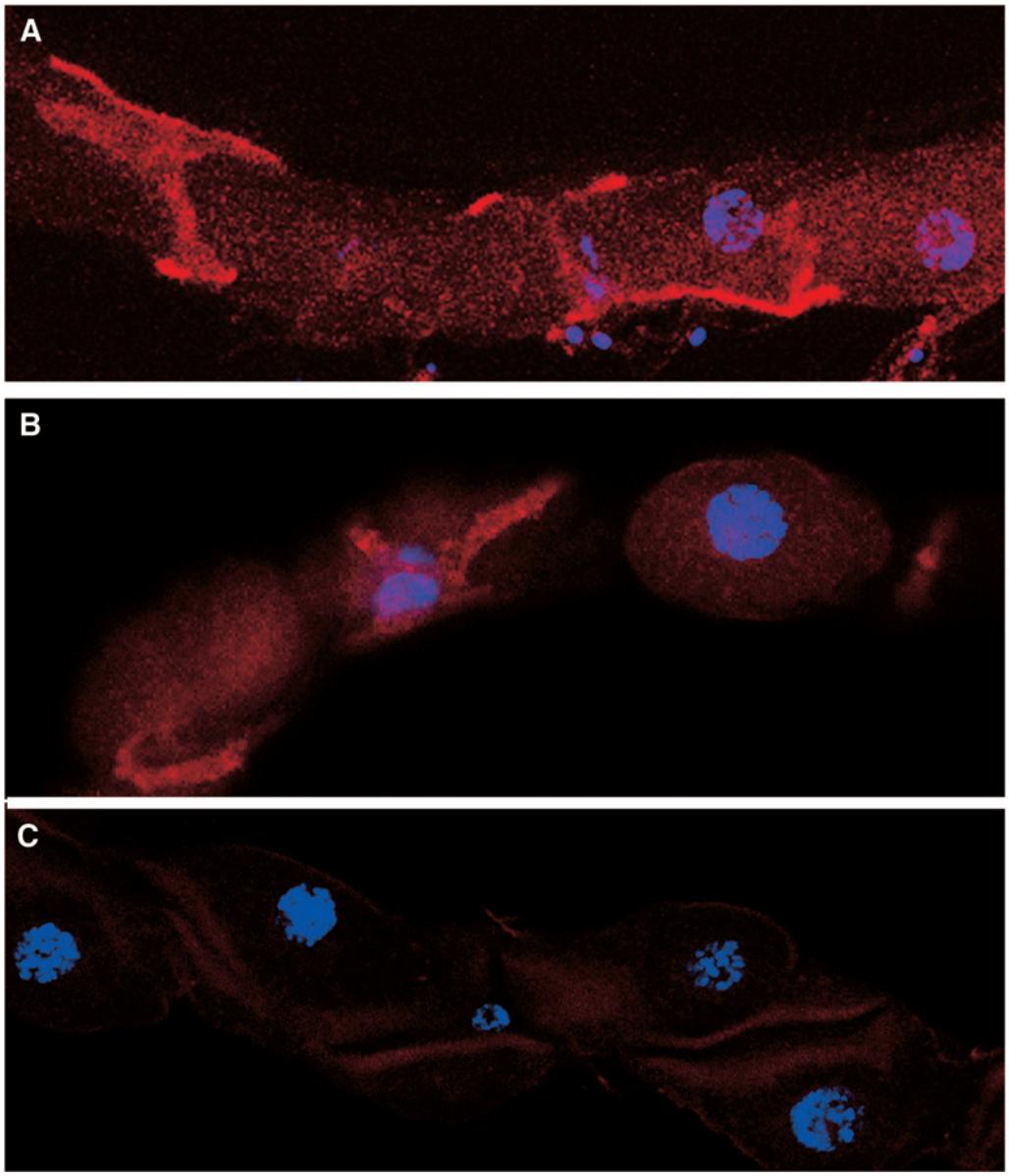


Fig. 11. *Anopheles* leucokinin receptor is expressed in stellate cells of the Malpighian tubule. Tubules were stained with anti-*Anopheles* leucokinin peptide purified IgG, raised as described in the text. Texas Red secondary antibodies were used to visualize the primary antiserum and DAPI was used to stain nuclei (blue). (A,B), Fluorescence microscopy of immunostained whole-mount tubules revealed staining in the secondary stellate cell type (arrows), concentrated in the basolateral membrane. (C) Tubules were processed as for A and B, but with pre-immune serum, confirming the specificity of the antibody. Only low-level non-specific staining of apical microvilli was observed. All images were captured on a Zeiss 510 Meta confocal microscope using a 63× objective and the approximate scale can be determined from the tubule diameter, which can be taken to be 35 μm.

Asp–Arg–Tyr or Asp–Arg–His sequence. Cys residues, thought to form a disulphide bridge in GPCRs, are also present in the first and second extracellular loops of all but the lymnokinin receptor. Similarly putative *N*-glycosylation sites in the second extracellular loop are present in all proteins except the lymnokinin receptor. Although there is a great deal of sequence diversity within the C-terminal domains, several Ser and Thr residues appear conserved, representing possible sites of phosphorylation.

TBLASTN analysis using the putative *A. stephensi* receptor was also used to identify a similar sequence in the *A. gambiae* genome sequence. In addition, *Drosophila* LKR was used to identify a similar sequence within the partially sequenced *D. pseudoobscura* genome. Without experimental confirmation the C-terminal domains could not be reliably predicted for the *A. gambiae* and *D. pseudoobscura* proteins. Therefore, the putative seven TM domain-containing regions of all the protein sequences were determined using the TMHMM program, and then aligned as before. From the resulting output dnd file a dendrogram was created using the TREEVIEW program

(Page, 1996) (Fig. 5). The percentage identity and similarity of each were also calculated using BioEdit and were scored on the BLOSUM62 matrix (Table 1). This was carried out for the TM domain-containing regions and for the four known full-length proteins.

The dendrogram of the TM domain regions of the leucokinin-like receptors reflects the phylogeny of the species concerned (Fig. 5). The *D. melanogaster* and *D. pseudoobscura* sequences are closely related, as are the *A. stephensi* and *A. gambiae* sequences. These four sequences are more closely related to each other than to the *Boophilus* sequence, with the molluscan *Lymnaea* sequence being the least similar. This ancestral relationship is verified by the identity and similarity values for each sequence comparison (Table 1).

The Anopheles *leucokinins act on the* Anopheles *leucokinin receptor to raise intracellular calcium*

Having identified both leucokinins and a leucokinin-like receptor within *Anopheles*, it was important to establish that they are a functional receptor–ligand pairing. S2 cells were transiently transfected with the *apoaequorin* ORF and the *A. stephensi leucokinin-like receptor* ORF constructs, and their expression induced, as previously described (Radford et al., 2002). The S2 cells were then subsequently assayed for agonist-dependent activation by monitoring $[Ca^{2+}]_i$ levels. An agonist-dependent response in $[Ca^{2+}]_i$ level was observed for each of the three *Anopheles* leucokinin peptides, with an order of potency of I>II>III for this particular concentration (Fig. 6). $[Ca^{2+}]_i$ levels increased from basal levels of 50 nmol l^{-1} to a peak concentration of 365 nmol l^{-1}, 325 nmol l^{-1} and 300 nmol l^{-1}, respectively, upon addition of *Anopheles* leucokinin I, II or III, representing a 6- to 7.3-fold increase. The $[Ca^{2+}]_i$ responses were biphasic in nature, with a primary Ca^{2+} spike followed by a secondary wave that peaked at approximately 175 nmol l^{-1}, for all three peptides, 20–30 s post-stimulation.

Dose–response curves were then generated for the action of each *Anopheles* leucokinin on the *A. stephensi* leucokinin receptor. The receptor responds to all three *Anopheles* leucokinins in a dose-dependent manner (Fig. 7). *Anopheles* leucokinin I appears to be slightly more potent at stimulating the receptor, with an EC_{50} value of 2.0 nmol l^{-1}, compared to values of 7.4 nmol l^{-1} and 8.4 nmol l^{-1} for the action of *Anopheles* leucokinin II and III, respectively. The EC_{50} values for the actions of these peptides on the *A. stephensi* receptor are considerably higher than the value for the action of *Drosophila* leucokinin on the *Drosophila* LKR, 56.5 pmol l^{-1} (Radford et al., 2002). Similar EC_{50} values were determined for the effect of the eight known leucokinins on *Leucophaea maderae* hindgut contraction (Cook et al., 1989; Cook et al., 1990). The existence of a higher affinity leucokinin receptor within *Anopheles* cannot be ruled out, although it is likely that this sequence would also have been identified from the genomic sequence. Nonetheless, it should be remembered that the action of *Aedes* leucokinins on *Aedes* tubules were consistent with the existence of more than one receptor (Veenstra et al., 1997); and that the broad concentration range of *Drosophila* leucokinin on *Drosophila* tubule was also taken as suggestive of multiple receptor classes (Terhzaz et al., 1999).

Cross-specific leucokinin signalling

The effects of the *Anopheles* leucokinins on S2 cells expressing the *Drosophila* LKR, *CG10626* (Radford et al., 2002) were also established. As this was assessing cross-specific activity, relatively high concentrations of peptide were used (10^{-6} mol l^{-1} and 10^{-7} mol l^{-1}). The application of both *Anopheles* leucokinin I (15 amino acids) and *Anopheles* leucokinin II (7 amino acids) produced a concentration-dependent increase in $[Ca^{2+}]_i$ in the S2 cells (Fig. 8). *Anopheles* leucokinin III (10 amino acids) did not produce any observable $[Ca^{2+}]_i$ response at either concentration tested. This is probably because *Anopheles* leucokinin III possesses a C-terminal pentapeptide, which diverges from the –Phe–His–Ser–Trp–Gly–amide present in *Drosophila* leucokinin. Both *Anopheles* leucokinin I and II contain a C-terminal pentapeptide identical to that of *Drosophila* leucokinin. The fact that only the *Anopheles* receptor responds to the divergent *Anopheles* leucokinin III peptide suggests that the *Anopheles* receptor has a broader specificity than the *Drosophila* receptor. The only extracellular regions of these proteins that are considerably different in sequence are the short N-terminal domain and the second extracellular loop. It is tempting to speculate that differences in these regions may define the specificity of the receptor-ligand interaction. Although responses to *Anopheles* leucokinin I and II were seen at both 10^{-6} mol l^{-1} and 10^{-7} mol l^{-1}, the response to 10^{-6} mol l^{-1} was significantly larger. At this concentration, $[Ca^{2+}]_i$ levels were seen to increase from basal levels of 50–60 nmol l^{-1} to a peak concentration of 250 nmol l^{-1} (*Anopheles* leucokinin I) and 208 nmol l^{-1} (*Anopheles* leucokinin II), approximately a four- and fivefold increase, respectively. It was not possible to determine whether these were maximal responses because the high concentrations required meant that dose–response curves could not be generated. For both peptides the $[Ca^{2+}]_i$ response was biphasic in nature, with a primary Ca^{2+} spike and evidence of a sustained secondary wave that peaked at approximately 130 nmol l^{-1} 20–30 s post-stimulation. Although the primary $[Ca^{2+}]_i$ responses to 10^{-7} mol l^{-1} *Anopheles* leucokinin I and II were different, the secondary responses were similar.

The effects of *Drosophila* leucokinin (Terhzaz et al., 1999) on S2 cells expressing the *A. stephensi* leucokinin receptor were also ascertained. *Drosophila* leucokinin was found to stimulate the *A. stephensi* receptor in a similar manner to the *Anopheles* leucokinins, displaying an EC_{50} value of 1.1 nmol l^{-1}, very similar to that of *Anopheles* leucokinin I (Fig. 9).

The Anopheles *leucokinin receptor is expressed in stellate cells of the Malpighian tubule*

Antisera against the *Anopheles* leucokinin receptor

identified a band of the predicted size of 65 kDa on western blots, together with a heavier band of approximately 72 kDa (Fig. 10). A similar doublet was observed in *Drosophila*, and was shown to be due to *N*-glycosylation of the receptor (Radford et al., 2002). Consistent with this, four potential *N*-glycosylation sites are present within the receptor sequence.

Immunocytochemistry of adult *Anopheles* tubule revealed staining specific to the stellate cells (Fig. 11), as has previously been reported for the *Drosophila* leucokinin receptor.

Discussion

This paper identifies and characterises a cognate pairing of the *Anopheles* leucokinins and their receptor in a genus containing major human and animal disease vectors. By comparison with the abundant knowledge of the leucokinin family in insects, it is now possible to distinguish significant differences in the numbers of leucokinins and their potencies across the Order Diptera. This may help to explain a radical difference between the diuretic actions of leucokinins in Diptera. In the *Drosophila* tubule, the *Drosophila* leucokinin receptor is found only in the type II (stellate) cells, and (using a transgenic aequorin calcium reporter) the peptide is known to raise calcium only in these cells (Radford et al., 2002). By contrast, in the *Aedes* tubule, the leucokinins are thought to act on principal cells to regulate paracellular permeability (Beyenbach, 2003a; Yu and Beyenbach, 2004). Given the relatively divergent taxonomy of these two insects, and the differences in prepropeptide structure and in receptor C-terminal sequence, the functional differences might not be so surprising. However, in *Anopheles*, which is phylogenetically much closer to *Aedes*, all aspects of leucokinin signalling, including the receptor localisation to stellate cells in the tubule, appear much closer to *Drosophila* than to *Aedes*. It will thus be of great interest to locate the homologous receptor within the *Aedes* tubule.

This work was supported by the Biotechnology and Biological Sciences Research Council, UK and by a Wellcome Trust PhD scholarship to J.C.R.

References

Bairoch, A., Bucher, P. and Hofmann, K. (1997). The PROSITE database, its status in 1997. *Nucleic Acids Res.* **25**, 217-221.

Beyenbach, K. W. (1998). Reversible paracellular permeability changes in insect Malpighian tubules induced by the neuropeptide leucokinin. *FASEB J.* **12**, 5869.

Beyenbach, K. W. (2003a). Regulation of tight junction permeability with switch-like speed. *Curr. Opin. Nephrol. Hypertens.* **12**, 543-550.

Beyenbach, K. W. (2003b). Transport mechanisms of diuresis in Malpighian tubules of insects. *J. Exp. Biol.* **206**, 3845-3856.

Coast, G. M., Zabrocki, J. and Nachman, R. J. (2002). Diuretic and myotropic activities of N-terminal truncated analogs of *Musca domestica* kinin neuropeptide. *Peptides* **23**, 701-708.

Cook, B. J., Holman, G. M., Wagner, R. M. and Nachman, R. J. (1989). Pharmacological actions of a new class of neuropeptides, the leucokinins I-IV, on the visceral muscles of *Leucophaea maderae*. *Comp. Biochem. Physiol.* **93C**, 257-262.

Cook, B. J., Holman, G. M., Wagner, R. M. and Nachman, R. J. (1990). Comparative pharmacological actions of leucokinins V-VIII on the visceral muscles of *Leucophaea maderae*. *Comp. Biochem. Physiol.* **95C**, 19-24.

Cox, K. J., Tensen, C. P., Van der Schors, R. C., Li, K. W., van Heerikhuizen, H., Vreugdenhil, E., Geraerts, W. P. and Burke, J. F. (1997). Cloning, characterization, and expression of a G-protein-coupled receptor from *Lymnaea stagnalis* and identification of a leucokinin-like peptide, PSFHSWSamide, as its endogenous ligand. *J. Neurosci.* **17**, 1197-1205.

Dow, J. A. T. and Davies, S. A. (2001). The *Drosophila melanogaster* Malpighian tubule. *Adv. Insect Physiol.* **28**, 1-83.

Dow, J. A. T. and Davies, S. A. (2003). Integrative physiology and functional genomics of epithelial function in a genetic model organism. *Physiol. Rev.* **83**, 687-729.

Gade, G. (2004). Regulation of intermediary metabolism and water balance of insects by neuropeptides. *Annu. Rev. Entomol.* **49**, 93-113.

Gubler, D. J. (1998). Resurgent vector-borne diseases as a global health problem. *Emerg. Infect Dis.* **4**, 442-450.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* **41**, 95-98.

Hill, C. A., Fox, A. N., Pitts, R. J., Kent, L. B., Tan, P. L., Chrystal, M. A., Cravchik, A., Collins, F. H., Robertson, H. M. and Zwiebel, L. J. (2002). G protein-coupled receptors in *Anopheles gambiae*. *Science* **298**, 176-178.

Holman, G. M., Cook, B. J. and Wagner, R. M. (1984). Isolation and partial characterization of five myotropic peptides present in head extracts of the cockroach, *Leucophaea maderae*. *Comp. Biochem. Physiol.* **77C**, 1-5.

Holman, G. M., Nachman, R. J. and Coast, G. M. (1999). Isolation, characterization and biological activity of a diuretic myokinin neuropeptide from the housefly, *Musca domestica*. *Peptides* **20**, 1-10.

Holmes, S. P., Barhoumi, R., Nachman, R. J. and Pietrantonio, P. V. (2003). Functional analysis of a G protein-coupled receptor from the southern cattle tick *Boophilus microplus* (Acari: Ixodidae) identifies it as the first arthropod myokinin receptor. *Insect Mol. Biol.* **12**, 27-38.

Holt, R. A., Subramanian, G. M., Halpern, A., Sutton, G. G., Charlab, R., Nusskern, D. R., Wincker, P., Clark, A. G., Ribeiro, J. M., Wides, R. et al. (2002). The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298**, 129-149.

Howarth, C. J., Prince, R. I., Dyker, H., Losel, P. M., Seinsche, A. and Osborne, R. H. (2002). Structure-activity relationship of contractile effects induced by helicokinins in the isolated gut of the lepidopteran caterpillar *Spodoptera frugiperda*. *J. Insect Physiol.* **48**, 75-82.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567-580.

Meola, S. M., Clottens, F. L., Holman, G. M., Nachman, R. J., Nichols, R., Schoofs, L., Wright, M. S., Olson, J. K., Hayes, T. K. and Pendleton, M. W. (1998). Isolation and immunocytochemical characterization of three tachykinin-related peptides from the mosquito, *Culex salinarius*. *Neurochem. Res.* **23**, 189-202.

Nachman, R. J., Strey, A., Isaac, E., Pryor, N., Lopez, J. D., Deng, J. G. and Coast, G. M. (2002). Enhanced in vivo activity of peptidase-resistant analogs of the insect kinin neuropeptide family. *Peptides* **23**, 735-745.

Page, R. D. M. (1996). TREEVIEW: An application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**, 357-358.

Radford, J. C., Davies, S. A. and Dow, J. A. T. (2002). Systematic GPCR analysis in *Drosophila melanogaster* identifies a leucokinin receptor with novel roles. *J. Biol. Chem.* **277**, 38810-38817.

Riehle, M. A., Garcynski, S. F., Crim, J. W., Hill, C. A. and Brown, M. R. (2002). Neropeptides and peptide hormones in *Anopheles gambiae*. *Science* **298**, 172-175.

Rosay, P., Davies, S. A., Yu, Y., Sozen, A., Kaiser, K. and Dow, J. A. T. (1997). Cell-type specific calcium signalling in a *Drosophila* epithelium. *J. Cell Sci.* **110**, 1683-1692.

Seinsche, A., Dyker, H., Losel, P., Backhaus, D. and Scherkenbeck, J. (2000). Effect of helicokinins and ACE inhibitors on water balance and development of *Heliothis virescens* larvae. *J. Insect Physiol.* **46**, 1423-1431.

Teal, P. E., Meredith, J. A. and Nachman, R. J. (1999). Development of amphiphylic mimics of insect neuropeptides for pest control. *Ann. NY Acad. Sci.* **897**, 348-360.

Terhzaz, S., O'Connell, F. C., Pollock, V. P., Kean, L., Davies, S. A., Veenstra, J. A. and Dow, J. A. T. (1999). Isolation and characterization of a leucokinin-like peptide of *Drosophila melanogaster*. *J. Exp. Biol.* **202**, 3667-3676.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680.

Torfs, P., Nieto, J., Veelaert, D., Boon, D., van de Water, G., Waelkens, E., Derua, R., Calderon, J., de Loof, A. and Schoofs, L. (1999). The kinin peptide family in invertebrates. *Ann. NY Acad. Sci.* **897**, 361-373.

Veenstra, J. A., Pattillo, J. M. and Petzel, D. H. (1997). A single cDNA encodes all three *Aedes* leucokinins, which stimulate both fluid secretion by the Malpighian tubules and hindgut contractions. *J. Biol. Chem.* **272**, 10402-10407.

Yu, M. J. and Beyenbach, K. W. (2004). Effects of leucokinin-VIII on Aedes Malpighian tubule segments lacking stellate cells. *J. Exp. Biol.* **207**, 519-526.