Insect Molecular Biology (2000) **9**(1), 75–84

Cloning and analysis of a cecropin gene from the malaria vector mosquito, *Anopheles gambiae*

J. Vizioli,[1] P. Bulet,[1] M. Charlet,[1] C. Lowenberger,[2] C. Blass,[3] H.-M. Müller,[3] G. Dimopoulos,[3] J. Hoffmann,[1] F.C. Kafatos[3] and A. Richman[4]

[1]*UPR CNRS 9022, Institut de Biologie Moléculaire et Cellulaire, 15, rue René Descartes, Strasbourg Cedex, France,* [2]*AHABS University of Wisconsin-Madison, Madison WI, USA,* [3]*European Molecular Biology Laboratory, Heidelberg, Germany, and* [4]*Department of Entomology, University of Maryland, College Park MD, USA*

Abstract

Parasites of the genus *Plasmodium* are transmitted to mammalian hosts by anopheline mosquitoes. Within the insect vector, parasite growth and development are potentially limited by antimicrobial defence molecules. Here, we describe the isolation of cDNA and genomic clones encoding a cecropin antibacterial peptide from the malaria vector mosquito *Anopheles gambiae*. The locus was mapped to polytene division 1C of the X chromosome. Cecropin RNA was induced by infection with bacteria and *Plasmodium*. RNA levels varied in different body parts of the adult mosquito. During development, cecropin expression was limited to the early pupal stage. The peptide was purified from both adult mosquitoes and cell culture supernatants. *Anopheles gambiae* synthetic cecropins displayed activity against Gram-negative and Gram-positive bacteria, filamentous fungi and yeasts.

Keywords: *Anopheles gambiae*, cecropin, malaria, antimicrobial peptide, innate immunity.

Introduction

Insect vectors of pathogens provide a physiological environment permissive for growth and differentiation of the disease-causing organism. In the case of the mosquito vectors of malaria, however, even efficient vector species present multiple barriers to unrestricted parasite growth (Warburg & Miller, 1991; Beier, 1998). Such barriers may be due to numerous factors, including anatomical features of the mosquito host, or physiological incompatibilities between insect and parasite. An intriguing question is to what extent the innate immune system of the mosquito may act to restrict parasite growth or development. Insects respond to bacterial or fungal infection by rapidly synthesizing a battery of potent antimicrobial peptide factors (Hetru *et al.*, 1998). The cloning of genes encoding these peptides in model insect species, particularly the fruit fly *Drosophila melanogaster*, has provided powerful tools with which to explore the mechanisms involved in the elicitation of the insect innate immune response (Hoffmann *et al.*, 1996). Recently, progress has been made in applying basic knowledge of invertebrate immunity to dipteran insects of medical importance (Richman & Kafatos, 1996), in particular the mosquito *Anopheles gambiae* (the most important African vector of the human malaria parasite *Plasmodium falciparum*) and the yellow fever mosquito *Aedes aegypti* (also a vector of protozoan and metazoan parasites). Initial studies of humoral immunity in both insect species led to the purification of defensin antimicrobial peptides and the cloning of defensin-encoding cDNAs (Chalk *et al.*, 1994; Lowenberger *et al.*, 1995; Cho *et al.*, 1996; Richman *et al.*, 1996). Both *An. gambiae* and *Ae. aegypti* respond to bacterial infection through rapid induction of defensin RNA and protein (Lowenberger *et al.*, 1995; Richman *et al.*, 1996). In *An. gambiae*, further studies have shown that humoral immune mechanisms are activated in multiple host mosquito tissues and at multiple time points during the course of infection by the rodent malaria parasite, *P. berghei* (Richman *et al.*, 1997; Dimopoulos *et al.*, 1998). The fact that immune-competent mosquitoes nevertheless provide a physiological milieu at least partially permissive for *Plasmodium* growth and differentiation represents an intriguing biological phenomenon of great significance for human health.

The extent to which endogenous humoral effector molecules may act to limit parasite development or growth in insects is unknown. Defensins have been shown to have

Received 22 June 1999; accepted 26 August 1999. Correspondence: Adam Richman, Department of Entomology, University of Maryland, College Park, MD 20742, USA. e-mail: ar133@umail.umd.edu

effects on certain stages of *Plasmodium* either *in vitro* or when injected into the haemolymph of infected mosquitoes (Shahabuddin *et al*., 1998). However, a microorganism or parasite invading a dipteran insect will probably encounter multiple humoral defence factors which may act synergistically. Here, we describe the identification and initial analysis of a second humoral factor from *An. gambiae*, a homologue of insect cecropin.

Cecropins have been isolated from both insects and mammals (Boman, 1994). Recently, cecropin homologues have been isolated from the mosquitoes *Ae. albopictus* (Sun *et al*., 1998) and *Ae. aegypti* (Lowenberger *et al*., 1999a). They are small (3–5 kDa), basic, helical peptides which were initially described as broadly active against Gram-negative, and a few species of Gram-positive bacteria, targeting the microbial membrane (Shai, 1998). Subsequently, cecropins have been shown to have varying degrees of activity against certain fungi (DeLucca *et al*., 1997) as well as metazoan (Chalk *et al*., 1995) and protozoan parasites including *Plasmodium* (Gwadz *et al*., 1989; Durvasula *et al*., 1997). The results reported here show that the *An. gambiae* cecropin is active against numerous Gram-negative and Gram-positive bacteria, as well as several species of filamentous fungi and yeasts. In *An. gambiae*, cecropin RNA is upregulated in response to both bacterial and *Plasmodium* infections. The gene is expressed in midgut and other tissues and can be induced in haemocyte-like cell lines challenged with killed bacteria. The peptide is produced in bacteria-inoculated adult mosquitoes, and is also detected in cell culture supernatants. Cecropin therefore represents a valuable molecular marker that, along with defensin, will be useful in assaying both activation of the mosquito acute phase response to *Plasmodium* challenge, and the ability of endogenous defence factors to limit parasite growth.

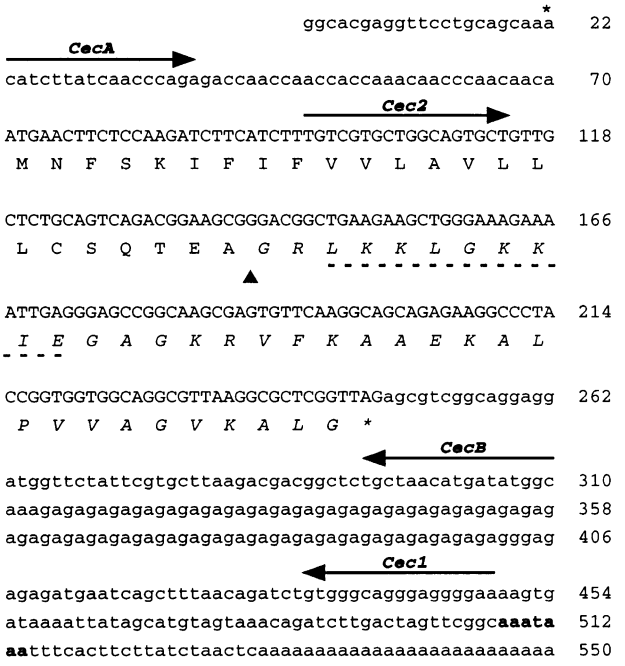

Figure 1. Sequence of *An. gambiae* cecropin cDNA and deduced amino acid sequence of cecropin precursor protein. The nucleotide sequence represents the consensus of partial clones generated by PCR and one clone isolated from a cDNA library by hybridization screening (the 5′ end of which, at nucleotide 22, is marked by an asterix). Arrows indicate location and orientation of specific PCR primers, while the dashed line indicates the amino acid sequence corresponding to the degenerate PCR primer. A potential polyadenylation signal is indicated in boldface. The arrowhead indicates the putative signal peptidase cleavage site. The mature cecropin sequence is indicated in italics.

Results

Cloning of cecropin cDNA and mapping of the cecropin locus

A polymerase chain reaction (PCR)-based screen of *Anopheles gambiae* cDNA was used to clone DNA fragments encoding cecropin (see Experimental procedures). A fragment of one clone was then used to probe a cDNA library prepared from *An. gambiae* haemocyte-like cells, yielding a clone 469 bp in length. The 5′ end of this clone corresponded to the nucleotide indicated by an asterix in Fig. 1. cDNA clones obtained from PCR-based screening of a larval cDNA library yielded the additional 5′ sequences shown in Fig. 1. Variability among clones was detected in non-protein-coding regions only. Among six sequenced clones, two displayed the sequence shown in Fig. 1, while four showed various deletions or additions between nucleotides (nt) 8 and 24. A GA simple sequence repeat (microsatellite) of variable length was found in the 3′ untranslated region; the longest repeat (89 nt) is shown in Fig. 1. All clones encoded identical polypeptide sequences of fifty-eight amino acids. Processing of a twenty-three-amino acid putative signal peptide (arrowhead in Fig. 1) would yield a thirty-five-amino acid peptide, exhibiting a high degree of similarity to cecropin proteins of other insect species (see Discussion). The cDNA encoded a C-terminal glycine residue indicating potential amidation (Bradbury & Smyth, 1991), leading to a thirty-four-amino acid mature peptide. Interestingly, a cecropin cDNA fragment was cloned independently, using messenger RNA (mRNA) differential display analysis of naïve vs. bacteria-infected mosquitoes. A differentially expressed cDNA of approximately 360 bp (Fig. 2) was sequenced and corresponded to the 3′ region of the cecropin sequence shown in Fig. 1, beginning with nt 187.

Characterization of the cecropin locus was initiated through the identification of a genomic clone in a bacterial artificial chromosome (BAC) library, using a PCR-based screening procedure (see Experimental procedures). Identification of a BAC clone (designated 17J21) containing the cecropin gene, was verified by Southern hybridization

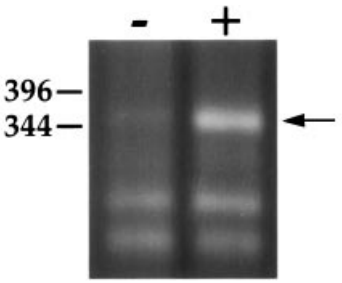

Figure 2. Differential display PCR. Amplified cDNA from naïve (–) and bacteria-infected (+) strain 4a r/r female mosquitoes. The band corresponding to a partial cDNA sequence of cecropin is indicated (arrow). Numbers indicate positions of molecular size markers (Gibco-BRL) in base pairs.

Figure 3. *In situ* hybridization of a cecropin-containing BAC clone to polytene chromosomes. The BAC clone 17J21 (see Experimental procedures) localizes specifically to band 1C of the X chromosome of *An. gambiae* (arrow). An identical localization was determined using as probe a cecropin cDNA clone.

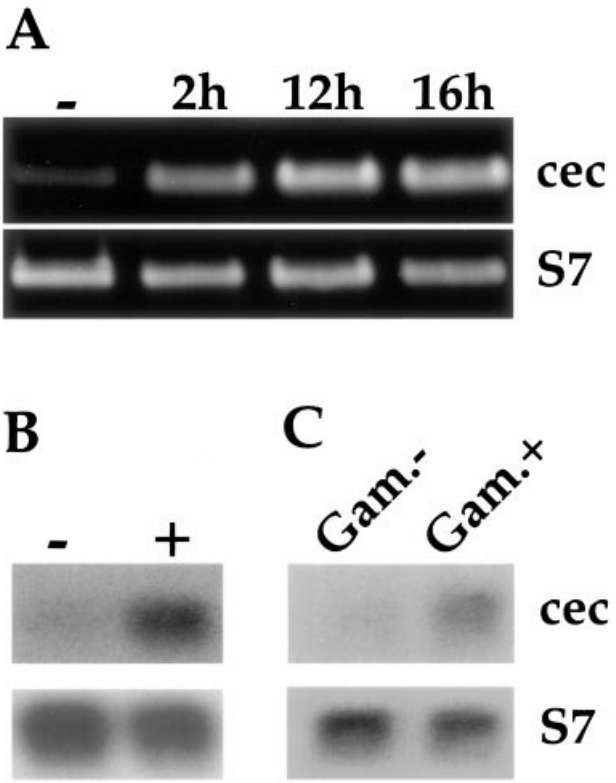

Figure 4. Analysis of inducible cecropin RNA expression. The transcript corresponding to the ribosomal protein S7 was used as a constitutive control. (A) RT-PCR analysis of cecropin (cec) expression in naïve (–) or bacteria-challenged female mosquitoes of the G3 strain. Time points in hours after challenge are indicated. (B) Northern analysis of total RNA from uninduced (–) or bacteria induced (+) cells of the 4a-2 line. (C) Northern analysis of total RNA extracted from L3–5 strain female mosquitoes 24 h after feeding upon mice infected with either the non-gametocyte producing Anka 2.33 strain of *P. berghei* (Gam.–) or the wild-type strain Anka 2.34 (Gam.+).

analysis using as probe a fragment of the cecropin cDNA (data not shown). The cytogenetic location of the cecropin gene identified in this study was determined by *in situ* hybridization of both the genomic BAC clone and a cDNA clone to ovarian polytene chromosomes. Using both probes, a signal was specifically detected on the X chromosome at division 1C (Fig. 3).

Antibacterial and antifungal assays

Peptides corresponding to mature *An. gambiae* cecropin were chemically synthesized in both C-terminally amidated and glycine-extended (non-amidated) forms. These peptides were tested for activity against a panel of Gram-negative and Gram-positive bacteria, several species of filamentous fungi, and yeasts. The synthetic amidated cecropin A of *Drosophila melanogaster* was used for comparison. The results revealed interesting differences that are considered in the Discussion (Table 1).

Expression of the cecropin gene

Insect cecropin expression is induced upon infection challenge with bacteria (Kylsten *et al*., 1990). Accordingly, the *An. gambiae* cecropin gene was tested for inducible expression at the RNA level by reverse transcription PCR (RT-PCR), under different induction conditions. Animals were tested for cecropin expression at 2, 12 and 16 h following inoculation with bacteria. As shown in Fig. 4(A) cecropin RNA levels are significantly elevated (approximately fourfold over non-infected control mosquitoes) within 2 h of infection. The RNA levels continue to increase, though moderately, between 2 and 16 h postinfection, reaching at this time a level approximately sixfold over control. The response to bacterial challenge was further investigated in haemocyte-like cell lines. Cells were exposed to heat-killed *Micrococcus luteus*, then analysed by RNA blot hybridization using a cecropin probe. As indicated in Fig. 4(B), cecropin RNA levels are strongly upregulated in these cells upon exposure to bacteria.

In order to ascertain whether cecropin gene expression is upregulated during *Plasmodium* infection, female mosquitoes were fed upon mice infected by wild-type *P. berghei*. As a control, mosquitoes were fed in parallel upon a strain of *P. berghei* which does not produce gametocyte-stage parasites in the mouse and is therefore not infectious to the mosquito. Twenty-four hours after feeding, wild-type

Table 1. Antimicrobial activity spectrum of synthetic cecropins from *Anopheles gambiae* and *Drosophila melanogaster*. Minimal inhibitory concentration (MIC) values are expressed as the interval between the highest concentration allowing growth and the lowest inhibitory concentration.

	MIC (μM)		
Microorganism	*Anopheles* cecropin A amidated	*Anopheles* cecropin A Gly-extended	*Drosophila* cecropin A amidated
Gram-negative bacteria			
Alcaligenes faecalis	1–2.5	2.5–5	10–25
Escherichia coli 1106	0.5–1	0.5–1	0.25–0.5
Escherichia coli D22	0.1–0.25	0.1–0.25	0.1–0.25
Escherichia coli D31	0.1–0.25	0.1–0.25	0.1–0.25
Escherichia coli SBS363	0.05–0.1	0.05–0.1	0.1–0.25
Enterobacter cloacae β12	0.5–1	0.5–1	0.5–1
Erwinia carotovora	0.5–1	0.5–1	0.5–1
Klebsiella pneumoniae	1–2.5	1–2.5	1–2.5
Pseudomonas aeruginosa	1–2.5	1–2.5	1–2.5
Salmonella typhimurium	0.1–0.25	0.25–0.5	0.5–1
Serratia marcescens Db11	*	*	*
Xanthomonas campestris	0.5–1	0.5–1	0.5–1
Gram-positive bacteria			
Aerococcus viridans	0.5–1	0.5–1	5–10
Bacillus cereus	50–100	*	*
Bacillus megaterium	0.5–1	0.5–1	1–2.5
Bacillus thuringiensis	*	*	*
Enterococcus faecalis	*	*	*
Lactobacillus sp.	*	*	*
Listeria monocytogenes	50–100	*	*
Micrococcus luteus	0.5–1	1–2.5	5–10
Staphylococcus aureus	50–100	*	*
Staphylococcus epidermidis	10–25	*	*
Staphylococcus haemolyticus	1–2.5	2.5–5	*
Staphylococcus saprophyticus	1–2.5	10–25	*
Streptococcus pyogenes	0.1–0.25	0.1–0.25	*
Fungi			
Aspergillus fumigatus	5–10	10–25	*
Beauveria bassiana	*	*	*
Botrytis cinerea	5–10	5–10	5–10
Fusarium culmorum	0.5–1	0.5–1	1–2.5
Fusarium oxysporum	1–2.5	1–2.5	1–2.5
Neurospora crassa	1–2.5	2.5–5	2.5–10
Yeasts			
Candida albicans	1–2.5	1–2.5	*
Candida glabrata	*	*	*
Cryptococcus neoformans	5–10	25–50	*
Saccharomyces cerevisiae	25–50	50–100	*

*Not detectable at 100 μM (the highest concentration tested).

parasites differentiate into motile ookinetes which invade the midgut epithelial cell layer (Richman *et al.*, 1997). As shown in Fig. 4(C), this process is associated with a significant upregulation of cecropin RNA levels in the infected mosquito.

Cecropin RNA levels were also monitored during different developmental stages. As shown in Fig. 5(A), significant expression in preadult stages is limited to young pupae. In this experiment a low 'basal' level of cecropin RNA is also detectable in 4-day-old adult females. In order to examine the tissue sites of cecropin gene expression, uninfected female mosquitoes were dissected into distinct body parts: head and thorax, abdominal carcass (composed primarily of abdominal epithelium, fat body, and body wall musculature), midgut and remaining tissues (consisting mainly of malpighian tubules, hindgut, ovaries, genitalia). These tissues were separately analysed by RT-PCR analysis. As shown in Fig. 5(B), cecropin gene expression is relatively abundant in head/thorax, is expressed at somewhat lower levels in the abdominal carcass and midgut, and is negligible in the remaining tissues. Midgut expression was further investigated by analysing separately the morphologically distinct anterior and posterior midgut regions (the latter representing the site of

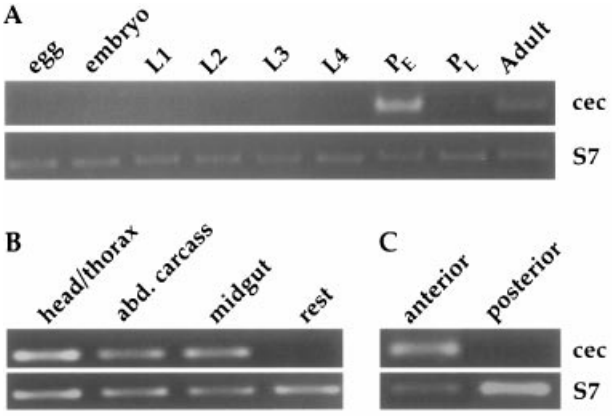

Figure 5. RT-PCR analysis of cecropin RNA levels in different developmental stages and in adult body parts. (A) Developmental analysis of cecropin RNA levels in freshly laid eggs, 18-h embryos, larval instars (L1–L4), early pupae (P_E), late pupae (P_L) and 4-day-old adults. (B) Female mosquitoes were dissected and the indicated body parts or organs analysed. 'abd. carcass': abdominal cuticle plus associated epidermis, fat body and musculature; 'rest': hindgut, malpighian tubules, ovaries and genitalia. (C) Adult female anterior and posterior midgut tissues, analysed separately. Naïve mosquitoes of strains 4a r/r (panel A) and L3–5 (panels B and C) were used for analysis. S7 and cecropin specific products were generated using twenty-four and twenty-five amplification cycles, respectively, in B and C. In panel A, S7 and cecropin were amplified twenty-five and thirty-one cycles, respectively.

bloodmeal digestion and invasion by *Plasmodium* ookinetes). As was previously described for *An. gambiae* defensin gene expression (Dimopoulos *et al.*, 1997), cecropin RNA was exclusively detected by RT-PCR in the anterior midgut (Fig. 5C).

Purification of native cecropin peptides

Culture supernatants of *An. gambiae* cell lines 4a-3A and 4a-3B as well as an extract prepared from bacteria infected mosquitoes were analysed by reversed-phase high-performance liquid chromatography (RP-HPLC) in order to detect amidated and/or glycine-extended native cecropin peptides. The cecropin-containing fractions were identified on the basis of comparative RP-HPLC analysis, loading the synthetic peptides as references. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) analysis of selected fractions revealed two molecules corresponding in mass to the amidated and glycine-extended forms of cecropin (3530.2 MH^+ and 3588.1 MH^+, respectively) in both mosquito extract (Fig. 6A) and cell supernatant 4a-3A fractions (Fig. 6B). In the cell line 4a-3B supernatant only the putative glycine-extended form (3587.7 MH^+) was detected (data not shown). Anti-bacterial activity against *Escherichia coli* strain SBS363 was also detected in cecropin-containing fractions from cell culture supernatants. Partial N-terminal sequence analysis of the putative glycine-extended and amidated forms (purified from 4a-3B and 4a-3A culture supernatants, respectively) corresponded to the expected N-terminal sequence of mature cecropin (GRLKKLGKKI).

Discussion

Cecropin structure and activity

The *An. gambiae* cDNA identified in this study encodes a protein of fifty-eight amino acids showing a high degree of similarity to insect cecropins, particularly to the recently

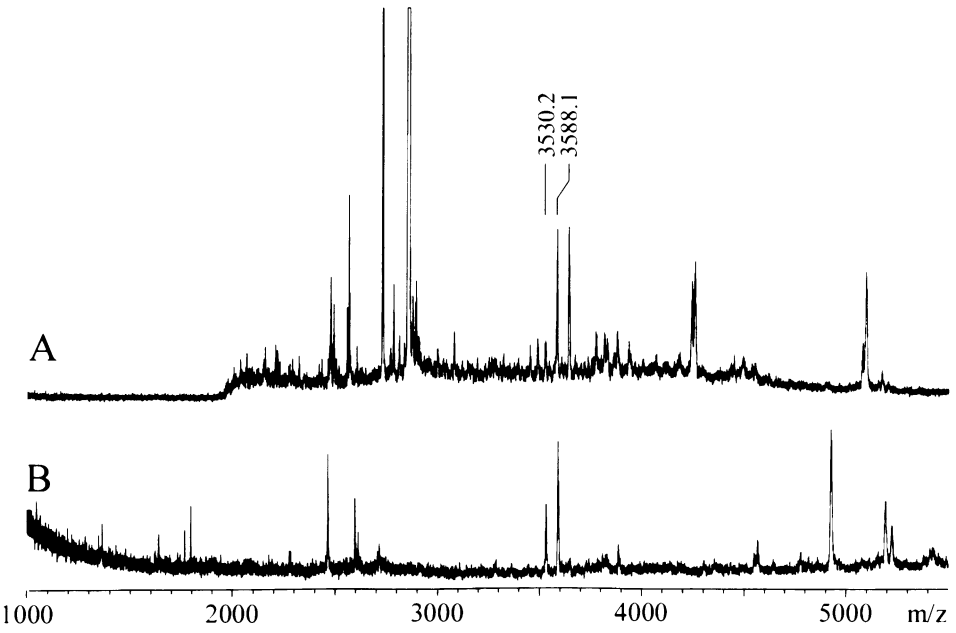

Figure 6. MALDI-TOF-MS analysis of the putative cecropin-containing reversed-phase HPLC fractions. Samples were isolated from bacteria-challenged mosquitoes (A) and cell line 4a-3A supernatant (B). Values indicate the peaks with m/z ratio corresponding to the putative amidated (3530.2 MH^+) and glycine-extended (3588.1 MH^+) cecropin isoforms.

identified cecropins of the mosquitoes *Ae. albopictus* (Sun *et al.*, 1998) and *Ae. aegypti* (Lowenberger *et al.*, 1999a). Sequence comparisons and phylogenetic analysis of *Anopheles* and *Aedes* cecropins have been reported by Lowenberger *et al.* (1999a). All cDNA clones obtained in this study encoded an identical cecropin peptide. Further analysis will be required to establish whether additional cecropin genes are present in *An. gambiae*. The peptide sequence reported here will therefore be designated AngCecA for *An. gambiae* cecropin A.

A potential signal sequence of twenty-three amino acids (Nielsen *et al.*, 1997) was identified, ending with the residues QTEA, a motif found at the ends of numerous cecropin signal peptides of dipteran insects (Lowenberger *et al.*, 1999a). A single cleavage reaction at Ala_{23} therefore would generate the thirty-five-amino acid AngCecA peptide (calculated molecular mass of 3588.2 MH^+, including the C-terminal glycine residue), displaying the general structural features of cecropins. Sequence analysis of native AngCecA confirmed the location of the cleavage site. The N-terminal portion of the molecule is highly basic, while the C-terminal domain is predominantly hydrophobic. Nuclear magnetic resonance (NMR) analysis of cecropin A from the lepidopteran *Hyalophora cecropia* revealed that the peptide consists of an N-terminal amphipathic α-helix, a 'hinge' region of three amino acids, and a C-terminal α-helical domain (Holak *et al.*, 1988; Boman, 1994). Structure prediction (Kneller *et al.*, 1990) for AngCecA supports an overall α-helical structure with a strikingly amphipathic N-terminal α-helix.

With the exception of the *Aedes* cecropins (Sun *et al.*, 1998; Lowenberger *et al.*, 1999a) and *Bombyx mori* cecropin D (Hara *et al.*, 1994), all other insect cecropins so far characterized are C-terminally amidated. This post-translational modification has been considered necessary for the full antimicrobial activity of the molecule (Li *et al.*, 1988; Hara *et al.*, 1994), and may protect the peptide from carboxypeptidase digestion (Callaway *et al.*, 1993). The presence of a glycine residue at the end of the deduced amino acid sequence of the *An. gambiae* cecropin suggests that C-terminal amidation (via terminal glycine removal) (Bradbury & Smyth, 1991) may occur in this insect to produce a fully active peptide. Biochemical analysis of whole mosquito extract and cell culture supernatants yielded cecropin peptides with molecular masses corresponding both to those calculated for amidated (3530.2 MH^+) and glycine-extended (3588.2 MH^+) forms as deduced from the cDNA sequence, and as measured by MALDI-TOF-MS analysis of the synthetic peptides. To analyse the role of amidation, we compared the antimicrobial activity of the *An. gambiae* glycine-extended cecropin and a mature form bearing a leucine amide as C terminus. Our results showed interesting patterns of activity for these two forms as compared to *Drosophila* amidated cecropin assayed in parallel (Table 1). Callaway *et al.* (1993) reported that the amidation of *Hyalophora* cecropin A significantly enhanced the activity of this peptide against both Gram-negative and Gram-positive bacteria. Our experiments revealed that amidation did not systematically increase the antimicrobial activity of *Anopheles* cecropin, but did so selectively, against some species. Thus, the two *Anopheles* forms did not differ significantly from each other or from *Drosophila* cecropin A in terms of their activity against most Gram-negative bacteria; the exceptions were *Alcaligenes faecalis* and *Salmonella typhimurium*, where the amidated form of AngCecA was two- to fivefold more effective than the non-amidated form, and tenfold more effective than *Drosophila* cecropin. Similarly, minimal differences were observed against filamentous fungi with the exception of *Aspergillus fumigatus*, against which the *Drosophila* cecropin was inactive while both mosquito forms were effective (the amidated form slightly more so). Comparable or lower activities against this fungus were reported for *Hyalophora* cecropins (DeLucca *et al.*, 1997, 1998).

By contrast, substantial differences in activity were apparent against Gram-positive bacteria: *Drosophila* cecropin was active against only three of the thirteen tested species, while AngCecA glycine-extended and amidated forms were active against six and ten species, respectively. For this microbial group the amidated mosquito cecropin was usually more active than the glycine-extended form. A similar pattern was observed with yeasts, against which *Drosophila* cecropin was ineffective while the *Anopheles* cecropins (particularly the amidated one) were active. Of special interest is the susceptibility to AngCecA of the medically important yeast species *Candida albicans*. This pathogen showed fiftyfold lower susceptibility to *Ae. aegypti* cecropin (Lowenberger *et al.*, 1999a), and sensitivity to other cecropins has not been reported (Moore *et al.*, 1996; DeLucca *et al.*, 1998). AngCecA, like the two other mosquito cecropins so far described (Sun *et al.*, 1998; Lowenberger *et al.*, 1999a), is devoid of any tryptophan residue. The presence of a tryptophan at the beginning of the N-terminal α-helix of the molecule, in position 1 or 2, was reported to be important for the activity of *Hyalophora* cecropin A (Andreu *et al.*, 1985) and is a common feature of almost all the insect cecropins (Hetru *et al.*, 1998). Despite the absence of this residue, the AngCecA is, in general, equal to or more active against Gram-negative bacteria than the *Drosophila* cecropin, which contains tryptophan at position 2. Our results further suggest that the presence of tryptophan may inhibit the activities of cecropins against yeasts as well as some Gram-positive bacteria.

Expression of cecropin messenger RNA

Experimental bacterial infection of *An. gambiae* induced approximately fourfold upregulation of cecropin RNA

within 2 h. Such an 'acute phase response' is characteristic of genes encoding insect antimicrobial factors (Hoffmann *et al*., 1996). A robust activation of cecropin gene expression was also observed in the haemocyte-like cell line 4a-2, in this case induced by exposure to killed bacteria. The same challenged cells also responded by activating defensin gene expression (data not shown). Differential display analysis of control vs. bacteria infected adult mosquitoes yielded a cDNA clone encoding the *An. gambiae* cecropin peptide, validating the usefulness of this approach for identification of additional immune-responsive genes in this insect. Upregulation of cecropin RNA expression was observed 24 h after ingestion of infective *P. berghei* parasites. This time period is associated with active invasion of midgut cells by motile ookinete stage parasites (Richman *et al*., 1997). Recent studies demonstrated that several immune-responsive genes of *An. gambiae*, including defensin, are activated during midgut invasion (Dimopoulos *et al*., 1998). In *Drosophila* different antimicrobial peptide-encoding genes are subject to regulation by distinct microbial elicitors and multiple signalling pathways (Lemaitre *et al*., 1997). It will therefore be of interest to compare rates and levels of induction of cecropin and defensin RNAs under different infection conditions in both mosquitoes and cultured cells.

During development the cecropin gene is expressed at negligible levels in eggs, embryos and larvae. Young pupae, however, display a striking elevation in cecropin RNA levels (Fig. 5A) which decrease in older pupae. Such a pupal stage-specific enhancement of expression was also observed for the *D. melanogaster* cecropin B and C genes (Tryselius *et al*., 1992) and for the defensin genes of *An. gambiae* (Richman *et al*., 1996) and *Ae. aegypti* (Lowenberger *et al*., 1999b). Interestingly, the *Ae. aegypti* cecropin A gene does not appear to be expressed during the pupal stage (Lowenberger *et al*., 1999a).

In uninfected mosquitoes the cecropin gene was differentially expressed in distinct body parts and tissues, in a pattern resembling that described for defensin RNA (Dimopoulos *et al*., 1997). Cecropin expression may therefore be inducible at multiple sites in response to *Plasmodium* parasite dissemination and tissue invasion during the course of infection. Midgut RNA expression of both defensin and cecropin was observed exclusively in the anterior midgut. As the midgut represents a vulnerable target for potential pathogens, localized synthesis and secretion of multiple immune peptides into the gut lumen may act as a barrier to microbial growth or invasion through the gut epithelium. In this regard it is interesting to note that midguts of African anopheline mosquitoes naturally harbour diverse species of bacteria (Straif *et al*., 1998). These include members of genera such as *Enterobacter*, *Klebsiella*, *Escherichia* and *Bacillus* that are sensitive to AngCecA activity *in vitro* (Table 1).

As in other dipteran insects, bacterial infection challenge of *An. gambiae* rapidly induces the expression of genes encoding defensin and cecropin antimicrobial peptides. In this malaria vector mosquito, these humoral immune factors are also expressed as a consequence of midgut invasion by *Plasmodium* parasites. Thus, cecropin represents not only a valuable additional molecular marker with which to monitor the mosquito response to parasite invasion, but also a potent humoral immune effector molecule. Exogenous cecropin has previously been reported to kill or inhibit growth of *Plasmodium* in mosquitoes (Gwadz *et al*., 1989). Future studies will determine the levels of activity of the *An. gambiae* cecropin against different stages of *Plasmodium*, as well as possible synergistic effects between defensin, cecropin and possibly other endogenous peptides.

Experimental procedures

Mosquito rearing, insect cell culture, and infection with bacteria and Plasmodium berghei

Anopheles gambiae mosquitoes were reared and infected, as previously described, with bacteria (Dimopoulos *et al*., 1997) and *P. berghei* (Richman *et al*., 1997). *Anopheles gambiae* cell lines were established from neonate larvae and maintained as described (Müller *et al*., 1999). Cells were bacteria-stimulated by a 7.5-h exposure to heat-killed *Micrococcus luteus* (approximately 1000 killed bacteria/cell) as described (Dimopoulos *et al*., 1997; Müller *et al*., 1999).

Complementary DNA cloning

Complementary DNA was prepared from bacteria-infected adults of the G3 strain using the Marathon cDNA Amplification kit (Clontech). The degenerate oligonucleotide 5′T(ATCG)AA-(AG)AA(AG)(TC)T(ATCG)GG(ATCG)AA(AG)AA(AG)AT(ATC)GA3′, corresponding to the amino acid sequence LKKLGKKIE, was designed based on the recently identified *Aedes aegypti* cecropin A as well as other known insect cecropins (Lowenberger *et al*., 1999a). This oligonucleotide was used in combination with the Marathon kit adaptor primer AP1 to amplify cecropin-related sequences from the cDNA template. An amplified product of 366 bp was obtained and cloned in the pGEM-T vector (Promega). PCR amplification was used to clone additional upstream sequences, using the primer Cec1 (5′-TTCCCCTCCCTGCCCAC-3′) in combination with an 'anchored' T3 primer (5′-ATTAACCCTCACTAAAG-3′). Template DNA was prepared from a directionally cloned cDNA library obtained from bacteria-infected mosquito larvae (Richman *et al*., 1996). The resulting product was cloned in the pCR 2.1 vector (TOPO-TA; Invitrogen). Six clones were sequenced using the Sequenase kit (USB) and were shown to encode the same cecropin peptide. A cDNA fragment representing 280 bp from the 5′ end of one clone (including the entire protein-coding region) was radioactively labelled using the Megaprime kit (Amersham) and used to probe a cDNA library (Zap Express; Stratagene) prepared from *An. gambiae* 4a-3B (Müller *et al*., 1999) and 4a-2 cells, immune-challenged with bacteria as described above. A cDNA clone containing a 469-bp insert was excised from

the pBK-CMV vector (Stratagene) according to the manufacturer's instructions. Sequence analysis revealed a cecropin cDNA beginning at nt 22 of the consensus sequence shown in Fig. 1.

Differential display polymerase chain reaction

Differential display PCR (DD-PCR) was performed as described (Dimopoulos *et al*., 1996) on cDNA prepared from naïve and bacteria-infected adult female mosquitoes 24 h after infection. PCR products were resolved on a 1.4% ethidium bromide-stained agarose gel. DNA was extracted from the gel using a QIAEX gel extraction kit (Qiagen), and cloned in the pCR2.5-TOPO TA vector (Invitrogen).

Northern blot analysis

Total RNA was prepared (RNAID; Bio 101) from 3- to 6-day-old female mosquitoes (strain L3–5), 24 h after feeding on BALB/c mice infected with a wild-type strain of *P. berghei* (Anka 2.34) or with a non-infective variant strain (Anka 2.33) as control. Infectivity was monitored as described (Richman *et al*., 1997). For each sample, 12 µg of RNA were electrophoresed on a 1% formaldehyde agarose gel, transferred to Hybond-N^+ membrane (Amersham), hybridized and washed as described (Sambrook *et al*., 1989). Successive hybridizations were performed using the 280 bp cecropin cDNA probe (described above) and a probe specific for the ribosomal protein S7 RNA as a constitutive control (Salazar *et al*., 1993). Probes were labelled using the Megaprime kit (Amersham). Northern analysis was also performed with 20 µg of total RNA isolated (RNeasy; Qiagen) from *An. gambiae* 4a-2 cells either uninduced or exposed to heat-killed bacteria as described above.

Reverse transcription polymerase chain reaction analysis and quantification

RT-PCR assays were performed as described (Richman *et al*., 1997). Cecropin-specific primers CecA (5′-CATCTTATCAACCCAGA-3′) and CecB (5′-GCCATATCATGTTAGCA-3′) were used, producing a PCR product of 287 bp. Primers CecA and CecB were designed to produce the largest possible product, avoiding polymorphic non-coding regions (see Results). Reaction products were electrophoresed on 2% agarose gels, stained with either ethidium bromide or Sybr Green (Molecular Probes), and quantified in a Fuji 2000 image quantifier. Induction levels of cecropin were calculated on the basis of the cecropin/S7 ratio and expressed as fold induction relative to control.

Isolation of a cecropin-containing bacterial artificial chromosome clone

A BAC library (average insert size 110 kb) prepared from genomic DNA of the *An. gambiae* PEST strain (kindly provided by Z. Ke and F. H. Collins) was screened using a PCR-based approach (details of the library and screening methodology will be published elsewhere) utilizing the primer combination Cec1 (see above) and Cec2 (5′-TGTCGTGCTGGCAGTGCT-3′). A BAC clone (designated 17J21) producing a cec1–2 PCR product of approximately 400 bp was identified.

In situ *hybridization to polytene chromosomes*

Hybridization of either the cecropin-containing BAC clone (see above) or a cDNA clone (in pBK-CMV) to polytene chromosome preparations from *An. gambiae* (strain Suakoko 2La) was performed as described (Kumar & Collins, 1994).

Peptide synthesis

Cecropin synthesis was performed by classic Fluoren-9-ylmethoxycarbonyl (Fmoc) methodology as previously described (Fehlbaum *et al*., 1996). Following purification by solid phase extraction and RP-HPLC, peptide purity and identity were established by partial Edman degradation, capillary zone electrophoresis, and mass spectrometry measurements.

Antimicrobial assays

Antimicrobial assays were performed against the same bacteria, filamentous fungi and yeast strains previously reported (Cociancich *et al*., 1994; Lowenberger *et al*., 1999a). In addition, the following microorganisms were used: *Enterococcus faecalis*, *Lactobacillus* sp., *Staphylococcus haemolyticus*, *S. epidermidis* (H. Monteil, Institute of Bacteriology, University of Strasbourg, France) and *Botrytis cinerea* (MUCL 30158; W. F. Broekaert, Janssens Laboratory of Genetics, Leuven, Belgium). Cecropin antibacterial and antifungal activities were assayed by liquid growth-inhibition as previously described (Hetru & Bulet, 1997). Antiyeast assays were performed in Sabouraud liquid medium (Bio-Mérieux) following the same procedure used for bacteria. In all assays the minimal inhibitory concentration (MIC) values were expressed according to Casteels *et al*. (1993).

Purification of native cecropin from bacteria-infected mosquitoes and cell line supernatant

Adult female mosquitoes (2 days old) were infected with bacteria and maintained under normal rearing conditions for 13 h. Proteins were extracted and prepurified as previously described (Hetru & Bulet, 1997). Peptides eluted with 40% acetonitrile (ACN) in acidified water (0.05% trifluoracetic acid) were separated by RP-HPLC on an Aquapore RP 300 C_8 column (250 × 4.6 mm; Brownlee) using a linear gradient of 2–62% ACN in acidified water, over 120 min, at a flow rate of 0.8 ml/min (Waters model 626 HPLC system). Supernatants of 15-day cultures of the *An. gambiae* cell lines 4a-3A and 4a-3B were acidified and prepurified as above. Step-wise elution was performed with 15% and 50% ACN in acidified water. The 50% ACN eluted fractions were subjected to size exclusion chromatography (Beckman Gold HPLC system) on two serially linked columns (Beckman Sec 3000 and Sec 2000; 300 × 7.5 mm) using 30% ACN in acidified water at a flow rate of 0.4 ml/min. Individual fractions were then analysed by RP-HPLC (as above), using a linear gradient of 2–60% ACN in acidified water, over 90 min, at 30 °C. Putative cecropin-containing fractions from mosquito extract and cell supernatants were identified by comparative RP-HPLC analysis using synthetic cecropin peptides as reference. Selected fractions were further analysed by RP-HPLC on a microbore Aquapore RP 300 C_8 column (0.1 × 10 cm; Brownlee) with a linear biphasic gradient of ACN in acidified water from 2% to 20% over 10 min, and from 20% to 30% over 50 min, at a flow rate of 0.08 ml/min, at 35 °C. Eluates were monitored by UV absorption at 225 nm for the first purifications and at 214 nm for the last step.

Microsequence analysis

Peptides were sequenced by Edman degradation on a pulse liquid automatic sequencer (*PE* Applied Biosystems, model 473A) and on a capillary liquid chromatography sequencer (*PE* Applied Biosystems, model 492 cLC).

Mass measurement by MALDI-TOF-MS

Peptide mass measurements were carried out with a Bruker (BIFLEX) mass spectrometer equipped with a delayed extraction ion source. Peptide samples were deposed on a thin layer of α-cyano-4-hydroxycinnamic acid (Sigma) as matrix and analysed as previously described (Uttenweiler-Joseph *et al.*, 1998).

GenBank accession number

The nucleotide sequence reported in this paper has been submitted to GenBank with the accession number AF200686.

Acknowledgements

The authors thank Douglas Seeley, Michael Spiegel and Belen Miñana for expert technical assistance, and Nükhet Cavusoglu for mass spectrometry analysis. Aspects of this research were funded by the John D. and Catherine T. MacArthur Foundation, the Human Frontier Science Program, a European Union Training and Mobility of Researchers (TMR) Network on Insect–Parasite Interactions, and an NIH Program Project Grant No. 1 P01 AI44220–01 on Innate Immunity in Vertebrates and Insects. This study was also supported by the Centre National de la Recherche Scientifique and the University Louis Pasteur of Strasbourg (France). J.V. was supported by a postdoctoral fellowship from the Fondazione Pasteur–Cenci Bolognetti. H.-M.M. was supported by a fellowship from the Deutsche Forschungsgemeinschaft.

References

Andreu, D., Merrifield, R.B., Steiner, H. and Boman, H.G. (1985) N-Terminal analogues of cecropin A: synthesis, antibacterial activity, and conformational properties. *Biochemistry* **24**: 1683–1688.

Beier, J.C. (1998) Malaria parasite development in mosquitoes. *Ann Rev Entomol* **43**: 519–543.

Boman, H.G. (1994) Cecropins: antibacterial peptides from insects and pigs. In: *Phylogenetic Perspectives in Immunity: the Insect Host Defense* (Hoffmann, J.A., Janeway, C.A.J. and Natori, S., eds), pp. 3–17. R. G. Landes Company, Austin, TX.

Bradbury, A.F. and Smyth, D.G. (1991) Peptide amidation. *Trends Biochem Sci* **16**: 112–115.

Callaway, J.E., Lai, J., Haselbeck, B., Baltaian, M., Bonnesen, S.P., Weickmann, J., Wilcox, G. and Lei, S.P. (1993) Modification of the C-terminus of cecropin is essential for broad-spectrum antimicrobial activity. *Antimicrob Agents Chemother* **37**: 1614–1619.

Casteels, P., Ampe, C., Jacobs, F. and Tempst, P. (1993) Functional and chemical characterization of hymenoptaecin, an antibacterial polypeptide that is infection-inducible in the honeybee (*Apis mellifera*). *J Biol Chem* **268**: 7044–7054.

Chalk, R., Townson, H., Natori, S., Desmond, H. and Ham, P.J. (1994) Purification of an insect defensin from the mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol* **24**: 403–410.

Chalk, R., Townson, H. and Ham, P.J. (1995) *Brugia pahangi*: the effects of cecropins on microfilariae *in vitro* and in *Aedes Aegypti*. *Exp Parasitol* **80**: 401–406.

Cho, W.L., Fu, Y.C., Chen, C.C. and Ho, C.M. (1996) Cloning and characterization of cDNAs encoding the antibacterial peptide, defensin A, from the mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol* **26**: 395–402.

Cociancich, S., Dupont, A., Hegy, G., Lanot, R., Holder, F., Hetru, C., Hoffmann, J.A. and Bulet, P. (1994) Novel inducible antibacterial peptides from a hemipteran insect, the sap-sucking bug *Pyrrhocoris apterus*. *Biochem J* **300**: 567–575.

DeLucca, A.J., Bland, J.M., Jacks, T.J., Grimm, C., Cleveland, T.E. and Walsh, T.J. (1997) Fungicidal activity of cecropin A. *Antimicrob Agents Chemother* **41**: 481–483.

DeLucca, A.J., Bland, J.M., Jacks, T.J., Grimm, C. and Walsh, T.J. (1998) Fungicidal and binding properties of the natural peptides cecropin B and dermaseptin. *Med Mycol* **36**: 291–298.

Dimopoulos, G., Richman, A., della Torre, A., Kafatos, F.C. and Louis, C. (1996) Identification and characterization of differentially expressed cDNAs of the vector mosquito, *Anopheles gambiae*. *Proc Natl Acad Sci USA* **93**: 13066–13071.

Dimopoulos, G., Richman, A., Müller, H.M. and Kafatos, F.C. (1997) Molecular immune responses of the mosquito *Anopheles gambiae* to bacteria and malaria parasites. *Proc Natl Acad Sci USA* **94**: 11508–11513.

Dimopoulos, G., Seeley, D., Wolf, A. and Kafatos, F.C. (1998) Malaria infection of the mosquito *Anopheles gambiae* activates immune-responsive genes during critical stages of the parasite life cycle. *EMBO J* **17**: 6115–6123.

Durvasula, R.V., Gumbs, A., Panackal, A., Kruglov, O., Aksoy, S., Merrifield, R.B., Richards, F.F. and Beard, C.B. (1997) Prevention of insect-borne disease: an approach using transgenic symbiotic bacteria. *Proc Natl Acad Sci USA* **94**: 3274–3278.

Fehlbaum, P., Bulet, P., Chernysh, S., Briand, J.P., Roussel, J.P., Letellier, L., Hetru, C. and Hoffmann, J.A. (1996) Structure-activity analysis of thanatin, a 21-residue inducible insect defense peptide with sequence homology to frog skin antimicrobial peptides. *Proc Natl Acad Sci USA* **93**: 1221–1225.

Gwadz, R.W., Kaslow, D., Lee, J.Y., Maloy, W.L., Zasloff, M. and Miller, L.H. (1989) Effects of magainins and cecropins on the sporogonic development of malaria parasites in mosquitoes. *Infect Immun* **57**: 2628–2633.

Hara, S., Taniai, K., Kato, Y. and Yamakawa, M. (1994) Isolation and α-amidation of the non-amidated form of cecropin D from larvae of *Bombyx mori*. *Comp Biochem Physiol* **108**: 303–308.

Hetru, C. and Bulet, P. (1997) Strategies for the isolation and characterization of antimicrobial peptides of invertebrates. In: *Methods in Molecular Biology Antimicrobial Peptide Protocols*, Vol. 78 (Shafer, W.M., ed.), pp. 35–49, Humana Press, Totowa, NJ.

Hetru, C., Hoffmann, D. and Bulet, P. (1998) Antimicrobial peptides from insects. In: *Molecular Mechanisms of Immune Responses in Insects* (Brey, P.T. and Hultmark, D., eds), pp. 40–66, Chapman & Hall, London.

Hoffmann, J.A., Reichhart, J.M. and Hetru, C. (1996) Innate immunity in higher insects. *Curr Opin Immunol* **8**: 8–13.

Holak, T.A., Engstrom, A., Kraulis, P.J., Lindeberg, G., Bennich, H., Jones, T.A., Gronenborn, A.M. and Clore, G.M. (1988) The solution conformation of the antibacterial peptide cecropin A: a nuclear magnetic resonance and dynamical simulated annealing study. *Biochemistry* **27**: 7620–7629.

Kneller, D.G., Cohen, F.E. and Langridge, R. (1990) Improvements in protein secondary structure prediction by an enhanced neural network. *J Mol Biol* **214**: 171–182.

Kumar, V. and Collins, F.H. (1994) A technique for nucleic acid *in situ* hybridization to polytene chromosomes of mosquitoes in the *Anopheles gambiae* complex. *Insect Mol Biol* **3**: 41–47.

Kylsten, P., Samakovlis, C. and Hultmark, D. (1990) The cecropin locus in *Drosophila*; a compact gene cluster involved in the response to infection. *EMBO J* **9**: 217–224.

Lemaitre, B., Reichhart, J.M. and Hoffmann, J.A. (1997) *Drosophila* host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms. *Proc Natl Acad Sci USA* **94**: 14614–14619.

Li, Z.Q., Merrifield, R.B., Boman, I.A. and Boman, H.G. (1988) Effects on electrophoretic mobility and antibacterial spectrum of removal of two residues from synthetic sarcotoxin IA and addition of the same residues to cecropin B. *FEBS Lett* **231**: 299–302.

Lowenberger, C., Bulet, P., Charlet, M., Hetru, C., Hodgeman, B., Christensen, B.M. and Hoffmann, J.A. (1995) Insect immunity: isolation of three novel inducible antibacterial defensins from the vector mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol* **25**: 867–873.

Lowenberger, C., Charlet, M., Vizioli, J., Kamal, S., Richman, A., Christensen, B.M. and Bulet, P. (1999a) Antimicrobial activity spectrum, cDNA cloning and mRNA expression of a newly isolated member of the cecropin family from the mosquito vector *Aedes aegypti*. *J Biol Chem* **274**: 20092–20097.

Lowenberger, C.A., Smartt, C.T., Bulet, P., Ferdig, M.T., Severson, D.W., Hoffmann, J.A. and Christensen, B.M. (1999b) Insect immunity: molecular cloning, expression, and characterization of cDNAs and genomic DNA encoding three isoforms of insect defensin in *Aedes aegypti*. *Insect Mol Biol* **8**: 107–118.

Moore, A.J., Beazley, W.D., Bibby, M.C. and Devine, D.A. (1996) Antimicrobial activity of cecropins. *J Antimicrob Chemother* **37**: 1077–1089.

Müller, H.M., Dimopoulos, G., Blass, C. and Kafatos, F.C. (1999) A hemocyte-like cell line established from the malaria vector *Anopheles gambiae* expresses six prophenoloxidase genes. *J Biol Chem* **274**: 11727–11735.

Nielsen, H., Engelbrecht, J., Brunak, S. and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10**: 1–6.

Richman, A.M., Bulet, P., Barillas-Mury, C., Hoffmann, J.A. and Kafatos, F.C. (1996) Inducible immune factors of the vector mosquito *Anopheles gambiae*: biochemical purification of a defensin antibacterial peptide and molecular cloning of preprodefensin cDNA. *Insect Mol Biol* **5**: 203–210.

Richman, A. and Kafatos, F.C. (1996) Immunity to eukaryotic parasites in vector insects. *Curr Opin Immunol* **8**: 14–19.

Richman, A.M., Dimopoulos, G., Seeley, D. and Kafatos, F.C. (1997) *Plasmodium* activates the innate immune response of *Anopheles gambiae* mosquitoes. *EMBO J* **16**: 6114–6119.

Salazar, C.E., Mills-Hamm, D., Kumar, V. and Collins, F.H. (1993) Sequence of a cDNA from the mosquito *Anopheles gambiae* encoding a homologue of human ribosomal protein S7. *Nucleic Acids Res* **21**: 4147.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Shahabuddin, M., Fields, I., Bulet, P., Hoffmann, J.A. and Miller, L.H. (1998) *Plasmodium gallinaceum*: differential killing of some mosquito stages of the parasite by insect defensin. *Exp Parasitol* **89**: 103–112.

Shai, Y. (1998) Mode of action of antibacterial peptides. In: *Molecular Mechanisms of Immune Responses in Insects* (Brey, P.T. and Hultmark, D., eds), pp. 111–134, Chapman & Hall, London.

Straif, S.C., Mbogo, C.N.M., Toure, A.M., Walker, E.D., Kaufman, M., Toure, Y.T. and Beier, J.C. (1998) Midgut bacteria in *Anopheles gambiae* and *An. funestus* (Diptera: Culicidae) from Kenya and Mali. *J Med Entomol* **35**: 222–226.

Sun, D., Eccleston, E.D. and Fallon, A.M. (1998) Peptide sequence of an antibiotic cecropin from the vector mosquito, *Aedes albopictus*. *Biochem Biophys Res Commun* **249**: 410–415.

Tryselius, Y., Samakovlis, C., Kimbrell, D.A. and Hultmark, D. (1992) CecC, a cecropin gene expressed during metamorphosis in *Drosophila* pupae. *Eur J Biochem* **204**: 395–399.

Uttenweiler-Joseph, S., Moniatte, M., Lagueux, M., Van Dorsselaer, A., Hoffmann, J.A. and Bulet, P. (1998) Differential display of peptides induced during the immune response of *Drosophila*: a matrix-assisted laser desorption ionization time-of-flight mass spectrometry study. *Proc Natl Acad Sci USA* **95**: 11342–11347.

Warburg, A. and Miller, L.H. (1991) Critical stages in the development of *Plasmodium* in mosquitoes. *Parasitol Today* **7**: 179–181.