Analysis of salivary gland proteins of the Southeast Asia malaria vector *Anopheles dirus* B

การศึกษาโปรตีนในต่อมน้ำลายยุงชนิด *Anopheles dirus* B ซึ่งเป็นพาหะของมาลาเรียในแถบเอเชียตะวันออกเฉียงใต้

[1]Dr. Narissara Jariyapan

[1]Assoc. Prof. Wej Choochote

[1]Ms. Thasaneeya Harnnoi

[1]Asst. Prof. Dr. Atchariya Jitpakdi

[2]Asst. Prof. Dr. Padet Siriyasatein

[3]Dr. Mark Wilkinson

[1]Department of Parasitology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[2]Department of Parasitology, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

[3]Department of Biochemistry, School of Biological Sciences, University of Liverpool, Liverpool, United Kingdom

Acknowledgements

This research was supported by Faculty of Medicine Endowment Fund for Medical Research, Chiang Mai University.

Contents

บทคัดย่อ

เชื้อมาลาเรียอาศัยอยู่ในต่อมน้ำลายของยุงพาหะก่อนการถ่ายทอดไปสู่โฮสต์ใหม่ การศึกษาเกี่ยวกับโปรตีนในต่อมน้ำลายยุงจะช่วยให้เข้าใจถึง ความสัมพันธ์ที่จำเพาะระหว่างสปอโรซอยท์ของมาลาเรียกับยุงพาหะได้ การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อศึกษาโปรตีนในต่อมน้ำลายยุง *Anopheles dirus* B ซึ่งเป็นพาหะหลักของมาลาเรียในแถบเอเชียตะวันออกเฉียงใต้ และเพื่อแยกและหาลำดับเบสที่เป็นรหัสสำหรับการสังเคราะห์โปรตีนเหล่านั้น จากการศึกษาโดยวิธี SDS-PAGE พบว่ามีโปรตีนหลักอย่างน้อย 7 ชนิด ที่จำเพาะสำหรับต่อมน้ำลายยุงตัวเมีย (63, 44, 43, 37, 33, 30 และ 18 kDa) และในแต่ละส่วนของต่อมน้ำลาย ประกอบไปด้วยโปรตีนหลักที่แตกต่างกัน เมื่อเปรียบเทียบรูปแบบของโปรตีนในต่อมน้ำลายของยุงที่กินเลือดกับไม่กินเลือดพบว่า มีความคล้ายคลึงกัน เราได้ทำการแยกลำดับของเปบไทด์ส่วน N-terminal ของโปรตีน 4 ชนิด และเซ็ทของลำดับเปบไทด์ภายในของโปรตีน 37 kDa ออกมาจาก two-dimensional polyacrylamide gel นอกจากนี้ได้สร้าง cDNA library ของต่อมน้ำลายยุงตัวเมีย *An. dirus* B และ แยกสาย cDNA 5 สาย จาก library นั้น ซึ่งพบว่าสาย cDNA ที่ได้ 2 สาย มีลำดับเบสที่ครบสมบูรณ์สำหรับการสังเคราะห์โปรตีน ส่วนอีก 3 สาย มีลำดับเบสสำหรับการสังเคราะห์โปรตีนเพียงบางส่วน จากการศึกษาลำดับเบสที่ครบสมบูรณ์ทั้ง 2 สาย พบว่า คล้ายกับโปรตีนที่มีอยู่ในแฟมิลี่ ต่อมน้ำลาย 1 (salivary gland 1 protein family, SG1) คือ SG1B-like หนึ่งสาย และอีกสายมีความคล้ายกับโปรตีนต่อมน้ำลาย GE-rich สำหรับสายที่ไม่ครบสมบูรณ์ พบว่ามีสองสายคล้ายกับโปรตีนที่มีอยู่ในแฟมิลี่ SG1 คือ SG1-like และ SG1D-like และอีกสายมีความคล้ายกับโปรตีนที่อยู่ในแฟมิลี่ antigen 5 คือ antigen 5-related 2

Abstract

Malarial parasites reside in the salivary glands of vectors prior to transmission. Analysis of mosquito salivary gland proteins will improve understanding of the specific interaction between malarial sporozoites and their mosquito vectors. The objectives of this study were to analyze salivary gland proteins of the Southeast Asia malaria vector, *Anopheles dirus* B and, isolate and sequence complementary DNAs (cDNAs) encoding the salivary gland proteins. SDS-PAGE analysis revealed that at least 7 major female-specific salivary gland protein bands (63, 44, 43, 37, 33, 30 and 18 kDa) were identified, each morphological region of the salivary glands containing different major protein bands. Similar electrophoretic protein profiles were detected comparing unfed and blood-fed mosquitoes. Four N-terminal peptide sequences of the major proteins were obtained and, a set of internal peptide sequences of the 37 kDa was extracted from two-dimensional polyacrylamide gels. Also, a female *An. dirus* B salivary gland cDNA library was constructed. Five unique cDNA fragments encoding 2 mature-protein and 3 partial-protein sequences were isolated from the cDNA library. Sequence analysis revealed that both mature-protein sequences were predicted to be a novel member of the salivary gland 1 (SG1) protein family, SG1B-like salivary protein and a GE-rich salivary gland protein. The partial proteins were related to two members of SG1 protein family, SG1-like and SG1D-like salivary proteins; and a member of the antigen 5 family, antigen 5-related 2 salivary protein.

Figures

Table

Introduction

In the last ten years, a new strategy of applying molecular genetic techniques to control the mosquito vectors has been proposed and received substantial attention (Curtis and Graves, 1988; Meredith and James, 1990; Crampton, 1994; Collins and James, 1996). Several studies have been focusing on the development of tools for the genetic alteration of mosquito vectors, with the final goal to block a parasite life cycle within mosquitoes, making them incapable of transmitting the disease. An important outcome of these studies was the success in achieving stable transformation of the yellow fever mosquito *Aedes aegypti* (Coates *et al.*, 1998; Jasinskiene *et al.*, 1998) and the malaria vector *Anopheles stephensi* (Catteruccia *et al.*, 2000). This has raised hopes for the production of mosquito strains that are unable to transmit various parasites (Beerntsen *et al.*, 2000). The development of transgenic mosquitoes refractory to malaria transmission requires not only the development of appropriate germ line transformation but also the identification of genes/effector molecules involved in parasite-vector interaction and the isolation of endogenous promoters able to drive the tissue-specific expression of a chosen gene (Beerntsen *et al.*, 2000; Aultman *et al.*, 2001).

We have focused our initial effort on the salivary glands of mosquitoes because they are the final site where malaria sporozoites reside before being passed to the vertebrate host (Ghosh *et al.*, 2000). They also express genes whose products are involved in the ability of mosquitoes to feed efficiently on blood. Several secreted proteins and gene expressed in the mosquito salivary glands have been identified and characterized (reviewed by Ribeiro and Francischetti, 2003). However, only six salivary gland-specific genes, four of which are expressed specifically in the female glands, have been isolated and characterized from the mosquito *Ae. aegypti* (James *et al.*, 1999). At least six cDNAs have been isolated from the *An. gambiae* and identified as salivary gland-specific genes (Arca *et al.*, 1999). Suwan *et al.* (2002) has reported two female salivary gland-specific cDNAs in *An. stephensi*, AnsD7 and AnsD7r1. None of them has been shown to affect the parasites in their invasion and development in the mosquito salivary glands. Recently, the salivary gland transcriptomes and proteomes of the mosquitoes, *Ae. aegypti* (Valenzuela *et al.*, 2002b), *An. gambiae* (Francischetti *et al.*, 2002) and *An. stephensi* (Valenzuela *et al.*, 2003) have been excellently described.

An. dirus B, an important vector in Southeast Asia, has been studied in many areas including molecular studies of insect glutathione S-transferases (Prapanthadara *et al.*, 1996; Prapanthadara *et al.*, 1998; Prapanthadara *et al.*, 2000; Oakley *et al.*, 2001a; Oakley *et al.*, 2001b), genetic studies of *Anopheles* species complexes (Baimai *et al.*, 1984; Green *et al.*, 1992; Walton *et al.*, 1999; Walton *et al.*, 2000a; Walton *et al.*, 2000b; Walton *et al.*, 2001), and malaria transmission (Somboon and Morakote, 1990, Klein *et al.*, 1991; Frances *et al.*, 1996; Singhasivanon *et al.*, 1999; Coleman *et al.*, 2001). However, very little is known about the salivary gland proteins of *An. dirus* B. In this study we therefore analyzed its salivary gland proteins and isolated cDNAs encoding specific-salivary gland proteins. The information obtained from this study would help to predict and understand the role of the salivary proteins in the mosquito and also be an initial step for further identification and characterization of antiparasite-effector molecules and promoters that may be useful in the development of genetically-transformed *Plasmodium*-refractory mosquitoes.

Objectives

1. To analyze the salivary gland proteins of *An. dirus* B using SDS-PAGE and two-dimensional gel electrophoresis.
2. To obtain N-terminal sequences of major salivary gland proteins of *An. dirus* B.
3. To construct a female *An. dirus* B salivary gland cDNA library.
4. To isolate and sequence cDNAs obtained from the library.

Materials and Methods

Mosquito

Anopheles dirus B mosquitoes (originally from the Armed Forces Research Institute of Medical Sciences (AFRIMS) laboratory, Bangkok, Thailand) were used in this study. This strain has been maintained in the insectary of the Department of Parasitology, Faculty of Medicine, Chiang Mai University, since 1982. This mosquito strain was proven to be highly susceptible to *Plasmodium vivax* and *P. falciparum* (Somboon and Morakote, 1990). Method for rearing of mosquitoes was followed standard techniques described by Choochote *et al.* (1983).

Salivary gland dissection

Salivary gland dissection was performed as the method described by Suwan *et al.* (2002). Adult mosquitoes aged between 3-10 days were cold anaesthetized on ice before salivary gland dissection. Salivary glands of the adult mosquitoes were dissected in RNase-free PBS, transferred to a microcentrifuge tube with a small volume of PBS, and then kept at -80°C until use. Dissection of the various regions of the female salivary glands was performed with 25 gauge needles under a dissecting microscope at 4x magnification. The medial lobes were cut at the junction of the medial lobes and the lateral lobes (Fig. 1). The distal-lateral and proximal-lateral lobes were cut at the intermediate region separating the two lobes. The gland parts were immediately removed to separate tubes to avoid possible protein contamination between the different sections of the glands. The gland parts were placed in a small volume of PBS and stored at -80°C until use. Salivary glands of blood-fed mosquitoes were dissected within 1 hour after taking a blood meal.

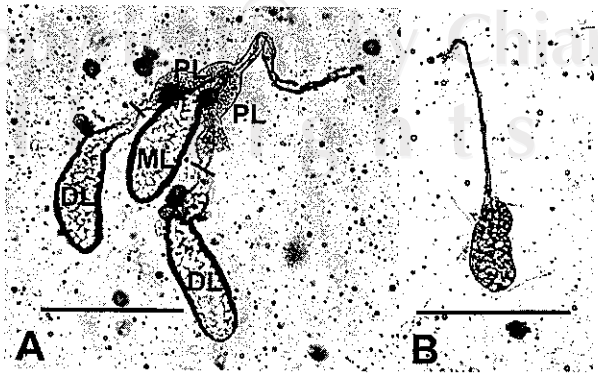

Fig. 1. Representative adult salivary glands of the mosquito, *Anopheles dirus* B. A. A female salivary gland. B. A male salivary gland. PL: proximal region of lateral lobe; DL: Distal region of lateral lobe; ML: median lobe. Bar represents 1 mm.

Isolation of messenger RNA and construction of complementary DNA library

Approximately 2-3 μg of poly(A) RNA were isolated from 150 pairs of *An. dirus* B female salivary glands using a Micro FastTrack^{TM}2.0 kit (Invitrogen, USA) and used as a template for double-stranded (ds) cDNA synthesis using cDNA Synthesis Kit (Phamarcia Biotech Inc., The Netherlands). A Zero Background^{TM}/Kan Cloning Kit (Invitrogen, USA) was used to construct the female salivary gland cDNA library. *Eco*RI/*Not*I adaptors were added into the blunt-ended ds cDNA before ligating into pZEro^{TM}-2 vector (Invitrogen, USA) and then 2 μl of ligation mixture (from a total volume of 10 μl) were transformed into TOP 10 Competent cells. Kanamycin was used for colony selection. Transformants per 1μg of mRNA was calculated. A pool of the bacterial colonies was mixed with glycerol and this cDNA library was stored at -20°C.

Analysis of salivary gland proteins in adult An. dirus B

Analysis of the salivary gland proteins of *An. dirus* B was performed by investigating their electropheretic profiles using SDS-PAGE and two-dimentional (2D) gel electrophoresis. Of interest were differences of protein expression in salivary glands of female and male mosquitoes, between unfed and blood-fed mosquitoes, as well as among individual female regions.

SDS-PAGE. Salivary gland samples were thawed on ice and mixed in 1:2 (v/v) 1XSDS gel loading buffer [50mM Tris-HCl (pH 6.8), 100mM DTT, 2% (w/v) SDS, 0.1% (w/v) Bromphenol blue, 10% (v/v) glycerol]. Then, the samples were heated for 5 minutes in a boiling water bath and loaded on 12% SDS polyacrylamide gels. Molecular weight markers (Bio-rad, USA) were applied in each gel.

Two-dimentional gel electrophoresis. Two-dimentional gel electrophoresis was performed using 2D system of Amersham Biosciences, Sweden. Five pairs of female salivary glands were solubilized in 125 μl sample solubilization solution [8M urea, 50 mM DTT, 4% CHAPS, 0.2% 3/10 Bio-lyte Ampholyte, 0.0002% Bromopheno Blue] and then loaded on an IPG strip (pI 3-10, 7 cm, Amersham Biosciences, Sweden) to re-hydrate for IPGphor (Amersham Biosciences, Sweden) to perform the first dimension isoelectric focusing (IEF) separation. The strip was incubated in equilibration buffer [6M urea, 2% SDS, 0.05M Tris pH

8.8, 20% glycerol] for 15 minutes. SDS-PAGE slab gels (12%) were used to separate proteins in the second dimension.

Following the electrophoresis, gels were Coomassie Brilliant Blue (CBB) stained. First, the gels were fixed in 50% methanol and 10% acetic acid for 30 min, then stained with 1% CBB in 10% methanol and 5% acetic acid for 2 hours, and finally de-stained in 10% methanol and 5% acetic acid until dark protein bands or spots were visible. Digital images of both SDS-PAGE and 2D CBB-stained gels were captured by scanning at 300 dpi using a color scanner. The images were stored and manipulated in PDF and TIFF formats using Photoshop™ 6.0 graphic software (Adobe Systems Inc., CA, USA)

N-terminal sequencing and internal sequencing

Proteins that express abundantly only in the female salivary glands of the mosquitoes were chosen for N-terminal sequencing. Briefly, fifty female salivary glands (10 glands per lane) were separated on a 12% SDS-polyacrylamide gel. Following transfer to PVDF membrane, N-terminal sequences were determined by Edman degradation on a Model 471A Protein Sequenator (Applied Biosystems, Cheshire, UK) at the School of Biological Sciences, University of Liverpool, UK. Internal sequencing was performed after N-terminal sequencing of some major proteins was not succeeded. After separating 5-10 salivary gland pairs on 2D gel, the gel was CBB stained and de-stained with 1% acetic acid (v/v) and 30% methanol (v/v) for 1 hour with 2-3 changes. The relevant piece of gel was excised and digested with trypsin *in situ* to general peptides. These were eluted from the gel slice, separated from each other by HPLC, individually collected and sequenced.

Isolation of the cDNAs encoding for major An. dirus B salivary gland proteins

Fifty to one hundred colonies were picked randomly from the female *An. dirus* B salivary gland cDNA library. Plasmid DNA of each clone was purified using the alkaline lysis method (Sambrook *et al.*, 1989) and then 1 µl of each plasmid DNA was electrophoresed through a 1% agarose gel, and visualized with ethidium bromide staining to determine the size of inserted plasmid DNA. Recombinant plasmids with insert size larger than 600 bp were purified using the QIAGEN miniprep (QIAGEN, Germany) before sequencing using a automated sequencing system at the BSU Bioservice Unit, National Science and Technology Development Agency (NSTDA) Building, Bangkok, Thailand.

Sequence analysis

Sequence editing and translation were carried out using DNAStar program. Analysis of sequence data by comparison to the Genbank sequence databases was performed by using BLAST program at National Centre for Biotechnology Information (NCBI) server (http://www.ncbi.nlm.nih.gov/). Signal peptides were predicted by submission of the sequences to Signal P server (http://www.cbs.dtu.dk/services/SignalP/), allowing the identification of putative secretory peptides of each sequences (Nielsen *et al.*, 1997). Sequence alignments were performed using the CLUSTALW program (Higgins *et al.*, 1996).

Results

Distribution of male and female salivary gland proteins

Total proteins in whole male and female salivary glands of *An. dirus* B, as well as the various micro-dissected morphological regions of female salivary glands were examined in Coomassie blue stained SDS-polyacrylamide gels (Fig. 2). At least 7 major and several minor protein bands were detected in the female salivary glands (Fig. 2, lane F), some of which are labeled (P1 through P7). The molecular masses of these major protein bands were estimated to be 63, 44, 43, 37, 33, 30 and 18 kDa, respectively. The male gland protein profile differed from the female profile and the protein content was lower (compare lane M, fifty male glands, with lane F, five female glands). The different morphological regions of the female salivary glands also displayed distinct protein electrophoretic profiles. Salivary gland protein bands P1, P4, P5, P6 and P7 appeared predominantly in the distal region (Fig.2, lane DL), while the female specific protein bands P1, P2 and P3 were predominant in the median lobe (Fig. 2, lane ML). The protein profile of the proximal-lateral region (Fig. 2, lane PL) appeared similar to the profile of male salivary glands (Fig. 2, lane M).

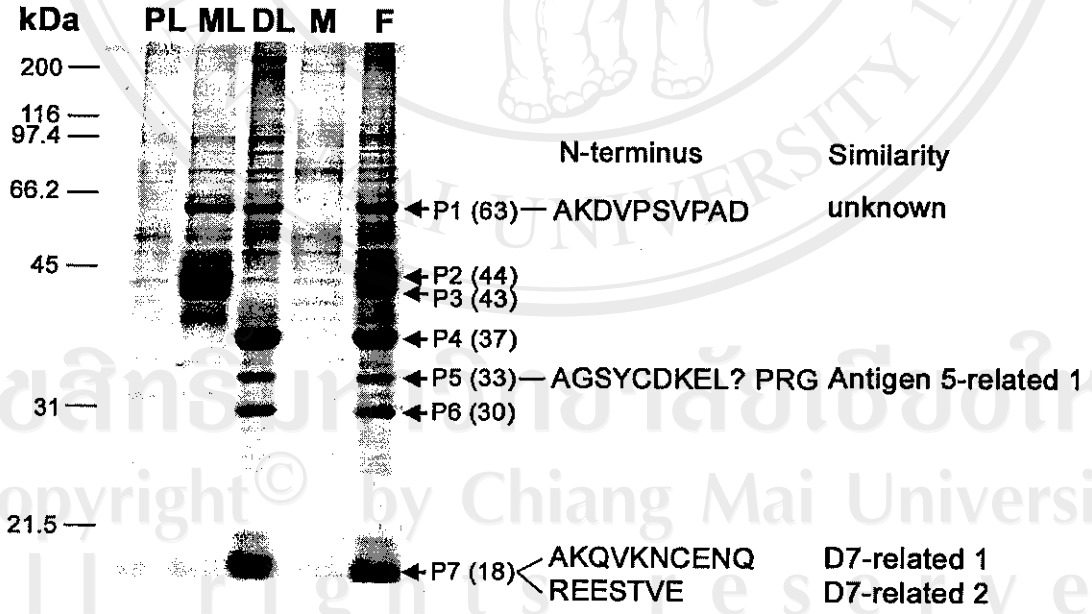

Fig. 2. Female and male salivary gland proteins of *An. dirus* B mosquitoes. Salivary proteins were separated on 12% SDS-polyacrylamide gels and Comassie blue stained. Lane PL, ten proximal-lateral lobes; lane ML, ten median lobes; lane DL, ten distal-lateral lobes; lane M, fifty whole male salivary glands; lane F, ten whole female salivary glands. Molecular mass markers are indicated on the left in kDa. Labels on the right indicate protein bands found specifically in the female glands (P1 – P7) and their estimated molecular mass (in blankets). The amino acid sequences obtained by Edman degradation and their similarity to known sequences in GenBank are shown on the right.

Figure 3 shows the salivary gland electrophoretic profiles of the blood-fed and the sugar-fed mosquitoes. The protein profiles are basically similar, although there are minor differences in both profiles after several repeats.

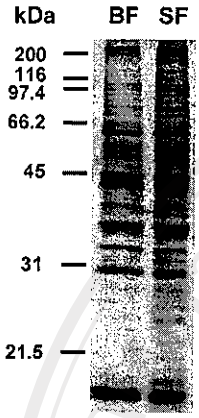

Fig. 3. Comparison of female salivary gland protein profiles between blood-fed and sugar-fed *An. dirus* B. The salivary gland proteins were separated on 12% SDS-polyacrylamide gels and Coomassie blue stained. Lane BF, ten whole glands of blood fed mosquitoes 1 hour previously; lane SF, ten whole glands of sugar fed mosquitoes. Molecular mass markers are indicated on the left in kDa.

Protein sequencing of salivary gland proteins

To identify these major proteins, they were transferred to PVDF membranes and the protein bands were cut from the membrane and submitted to Edman degradation. After several attempts to sequence N-terminal peptides of the seven major protein bands, only four peptides were successfully sequenced from the three protein bands (P1, P5 and P7) (Fig. 2). The P2, P3 and P6 were not pure, therefore, no sequence was obtained. The N-terminal sequences were compared with protein sequences in the GenBank databases. The N-terminal sequence of P5 (AGSYCDKELXPRG) shared homology with the predicted N-terminal sequences of antigen 5-related 1 of *An. stephensi* (Valenzuela *et al.*, 2003) (7 out of 13 residues identical) and of *An. gambiae* (Francischetti *et al.*, 2002b) (5 out of 13 residues identical). Protein band P7 contained a mixture of two N-terminal sequences, AKQVKNCENQ and REESTVE. The first one was homologued with the N-terminal sequence of *An. gambiae* D7-related 1 (Francischetti *et al.*, 2002b) (5 out of 9 residues identical). The last one matched the N-terminal sequence of D7-related 2 of *An. gambiae* (Francischetti *et al.*, 2002b) (6 out of 7 residues identical) and *An. stephensi* (Valenzuela *et al.*, 2003) (6 out of 7 residues identical). However, no match was found for P1 N-terminal sequence. As Edman degradation for the P4 band was unsuccessfully, internal sequencing of the 37 kDa spot from 2D gels was performed (Fig. 4). The tryptic-peptide sequences of the spot were QVHDOL, DGYLK and SFVVAR. These internal-peptide sequences were identical to the deduced amino acids sequence of a cDNA clone, Andi054 (see below).

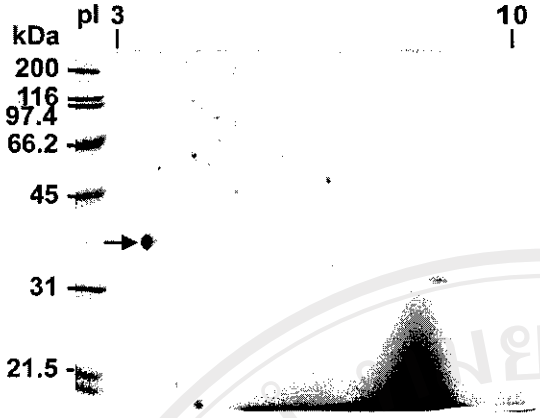

Fig. 4. Coomassie blue-stained 2D gel of proteins from 3 female salivary glands of *An. dirus* B. Arrow indicates 37 kDa spot, the most abundant protein expressed in the salivary glands. Molecular mass markers are indicated on the left in kDa. Isoelectric points (pI) are indicated at the top.

DNA sequences of salivary gland cDNA fragments and alignment of translated cDNA fragments to known proteins

A female *An. dirus* B salivary gland cDNA library was constructed, and screened by randomly picking clones and sequencing plasmids with large inserted. Five cDNA fragments ranging in size from 658 to 1,300 nucleotides (nt) were obtained from the cDNA library. A polyadenylation signal sequence, AATAAA, was found at the 3'-terminus of all sequences indicating that the carboxy-terminal region of the 6 deduced polypeptides was complete. The cDNA sequence data were translated in 6 frames for deduced amino acid sequences. The individual cDNA sequences and their deduced amino acid sequences were subjected to a DNA data bank search using the BlastP program at NCBI. Signal peptides were predicted by submission of the sequences to the Signal P server. Sequence analysis showed that the 5 cDNA fragments were unique (Table 1 and Appendix).

Table 1. Properties of the *Anopheles dirus* B salivary gland cDNAs isolated in this study.

clone	Type[1]/ Size[2]	GenBank[3]	Best match to NR protein database[4]	MW1[5]	SP[6]	MW2[7]	PI[8]
Andi006	P/ 1269 (402)	AY296729	SG1D *An. stephensi*	NA[9]	20-21	44085.95	9.43
Andi027	P/ 1233 (388)	AY299325	SG1-like3 *An. gambiae*	NA	17-18	41990.62	5.98
Andi053	F/ 1300 (391)	AY299326	SG1B *An. stephensi*	44858.03	20-21	42777.37	6.65
Andi054	F/ 863 (257)	AY299327	GE rich *An. stephensi*	27454.11	20-21	25370.40	4.15
Andi099	P/ 799 (230)	AY299329	Antigen 5-related 2 *An. gambiae*	NA	NA	NA	NA

Type[1], clone type (P = partial or F = full-length). Size[2], length of the cDNA fragments expressed in base pair (deduced amino acids). GenBank[3], NR database accession number. Similarity[4], amino acid similarities to known sequences deposited in GenBank or European Molecular Biology Laboratory databases. MW1[5], molecular mass before signal peptide removal. SP[6], the most likely cleavage site for signal peptide. MW2[7], molecular mass after signal peptide removal. PI[8], Isoelectric point. NA[9], not avialable.

Andi006 cDNA fragment encoded a protein of 402 amino acids (aa), 1269 nucleotides (nt). No 5' UTR was found. After submission the amino acid sequence to the Signal P server, the most likely cleavage site for a putative signal peptide was found between position 21 and 22 (CRG-KF). Andi006 had one N-linked glycosylation site at Asn[75]. Figure 5 shows that the Andi006 was closely related to *An. gambiae* SG1D salivary precursor with 59% similarity (38% identity) and to *An. gambiae* hypothetical protein with 55% similarity (36% identity).

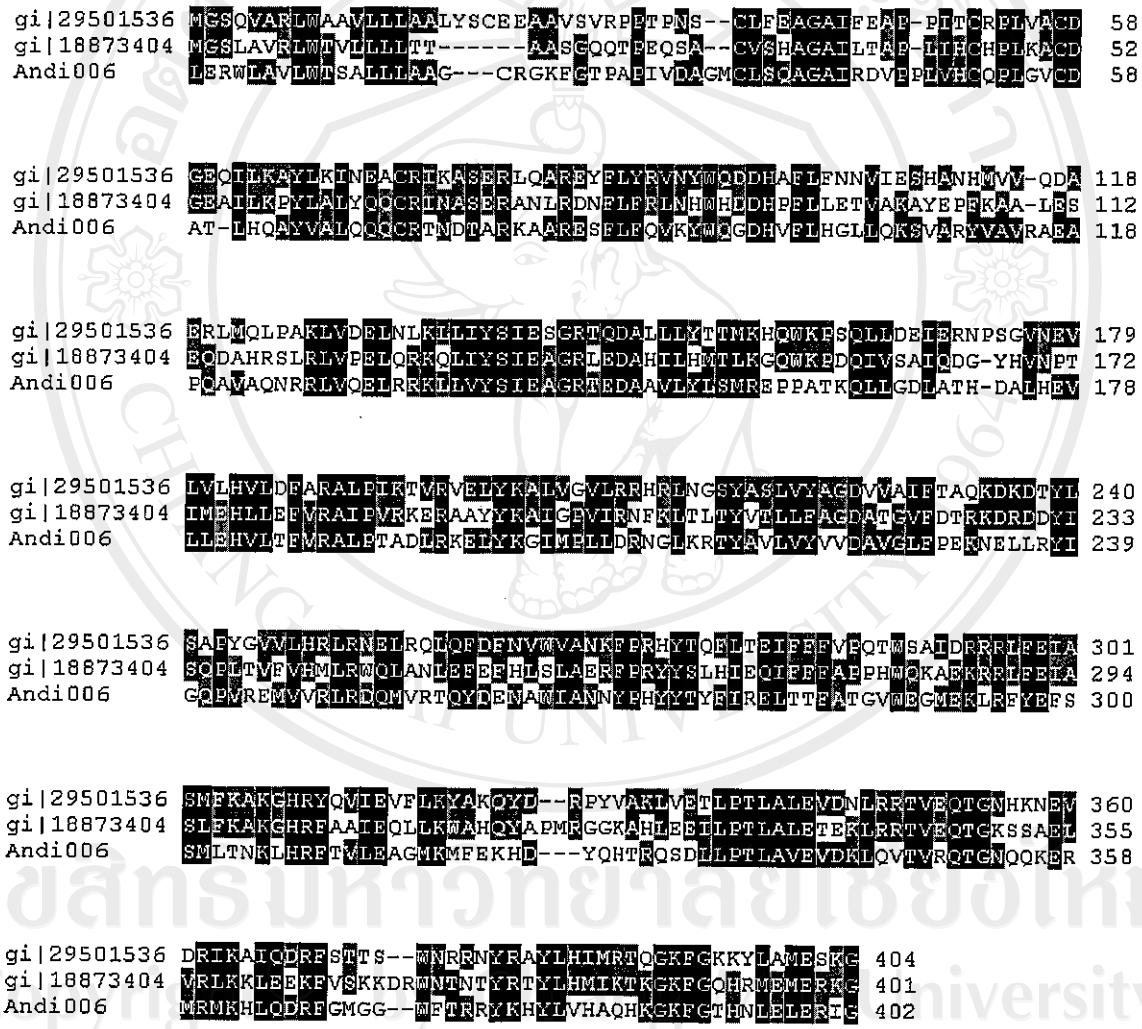

Fig. 5. CLUSTAL alignment of *An. dirus* B SG1D-like (Andi006; GenBank accession number AY296729), *An. stephensi* SG1D salivary protein precursor (gi|29501536) and *An. gambiae* hypothetical protein (gi|18873404). Similar amino acid residues are marked with a gray background, identical amino acids with a black background.

Andi027 cDNA fragment consisted of 1233 nt. The deduced protein sequences of Andi027 contained 388 aa. The amino acid sequence between position 17 and 18 (ADG-LP) was the most likely cleavage site for a putative signal peptide. However, no 5'UTR was observed. The protein contained three potential N-linked glycosylation sites at Asn^{98}, Asn^{162} and Asn^{258}. Andi027 showed 71% similarity (59% identity) with *An. gambiae* SG1-like 3 and 69% similarity (54% identity) with *An. gambiae* ENSANGP00000019238. There were several regions that showed good conservation of sequence among the three proteins (Fig. 6).

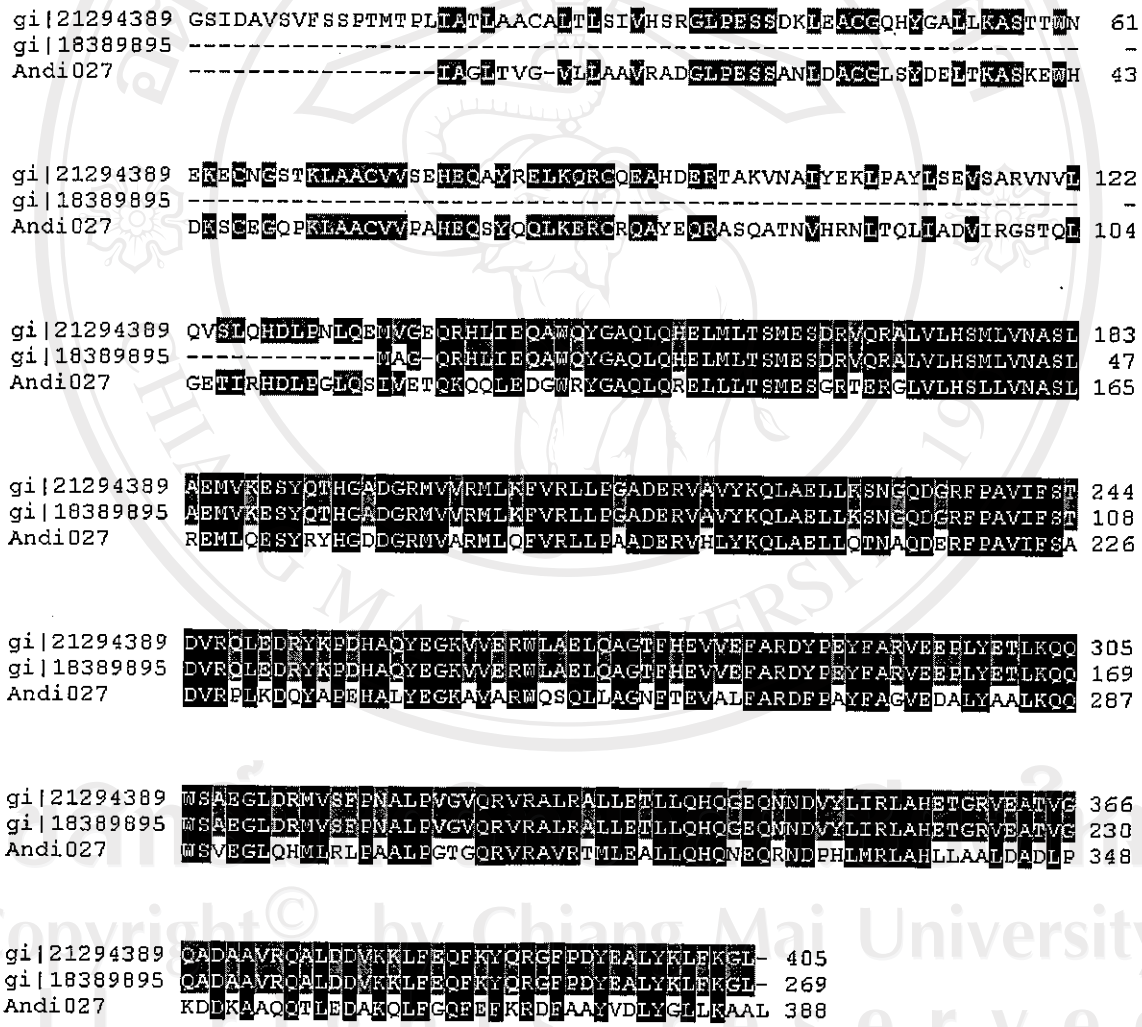

Fig. 6. CLUSTAL alignment of *An. dirus* B SG1-like (Andi027; GenBank accession number AY299325), *An. gambiae* ENSANGP00000019238 (gi|21294389) and *An. gambiae* salivary gland 1-like 3 protein (gi|18389895). Similar amino acid residues are marked with a gray background, identical amino acids with a black background.

Andi053 was one of the two full-length cDNAs obtained in this study. The cDNA specified a protein of 391 aa, 391 nt. The protein sequence showed high similarity (65%) to putative salivary protein SG1B of *An. stephensi*. A signal peptide indicative of secretion was found between position 20 and 21 (AGA-RP), producing a predicted mature molecular mass of 42.78 kDa (pI 6.65). One potential N-linked glycosylation site was found at Asn^{74}. However, no amino terminal sequence was detected from a SDS-PAGE gel band. The CLUSTAL alignment of SG1B-like protein of *An. dirus* B and SG1B protein of *An. stephensi* and *An. gambiae* ENSANGP00000019156 is shown in Fig. 7.

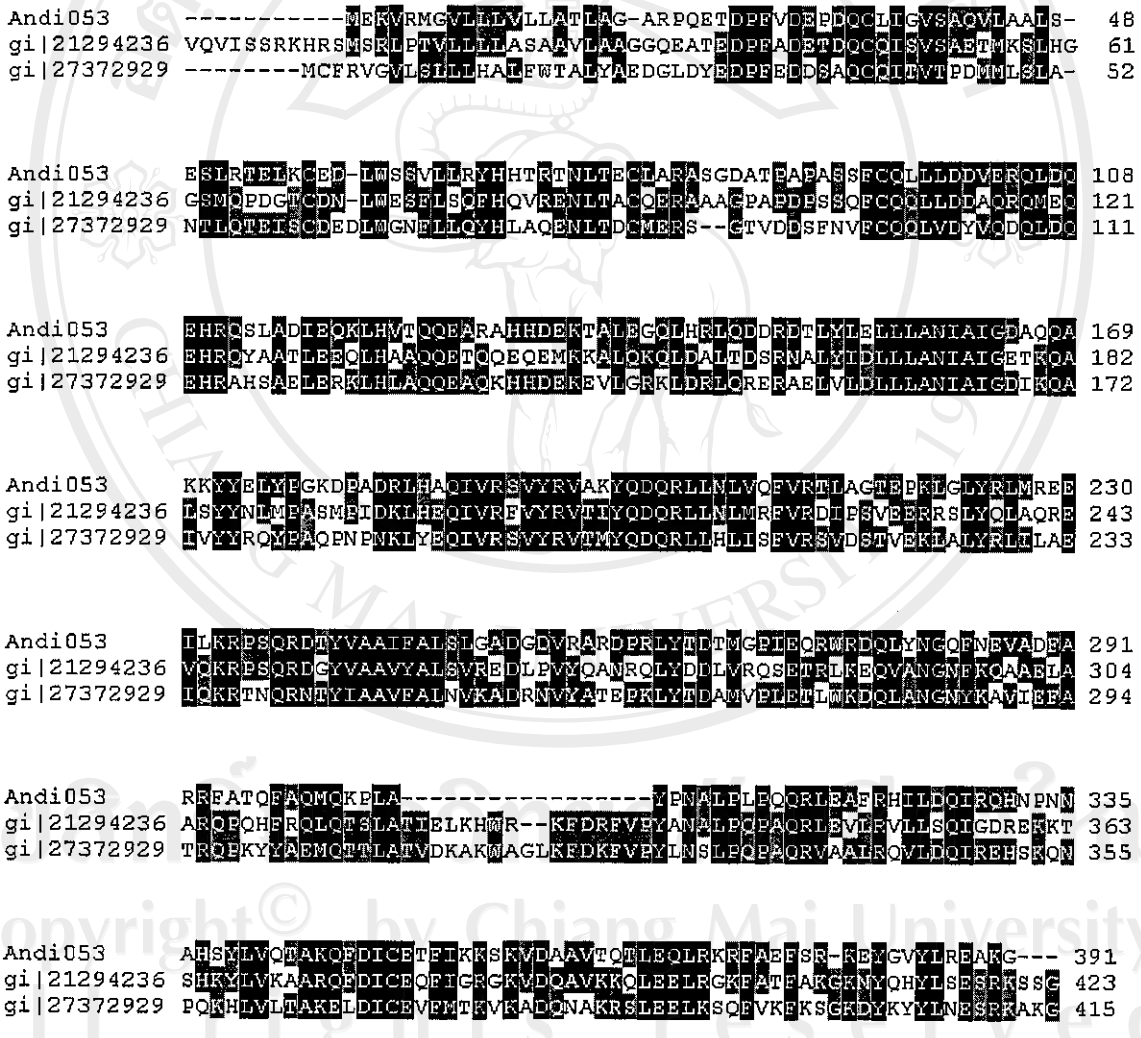

Fig. 7. CLUSTAL alignment of *An. dirus* B SG1B-like (Andi053; GenBank accession number AY299326), *An. gambiae* ENSANGP00000019156 (gi|21294236) and *An. stephensi* putative salivary protein SG1B (gi|27372929). Similar amino acid residues are marked with a gray background, identical amino acids with a black background.

Another full-length cDNA, Andi054, encoded a protein of 257 aa, 863 nt, with 89% similarity (79% identity) to *An. gambiae* ENSANGP00000022344 and 87% similarity (74% identity) to *An. stephensi* GE-rich salivary gland protein precursor. In the Andi054 protein sequence, 2 N-linked glycosylation sites were found at Asn^{78} and Asn^{224}. A putative signal peptide was found between position 20 and 21 (VTA-RP). A predicted minimum size of the mature protein is 25.37 kDa with pI 4.15. Although no N-terminal sequence data was matched with the Andi054 amino terminal sequence, the internal peptide sequences of the 37 kDa spot (Fig. 4) were identical to the sequence. Figure 8 shows the CLUSTAL alignment of GE-rich salivary gland protein of *An. dirus* B, *An. gambiae* ENSANGP00000022344 and *An. stephensi* GE-rich salivary gland protein precursor.

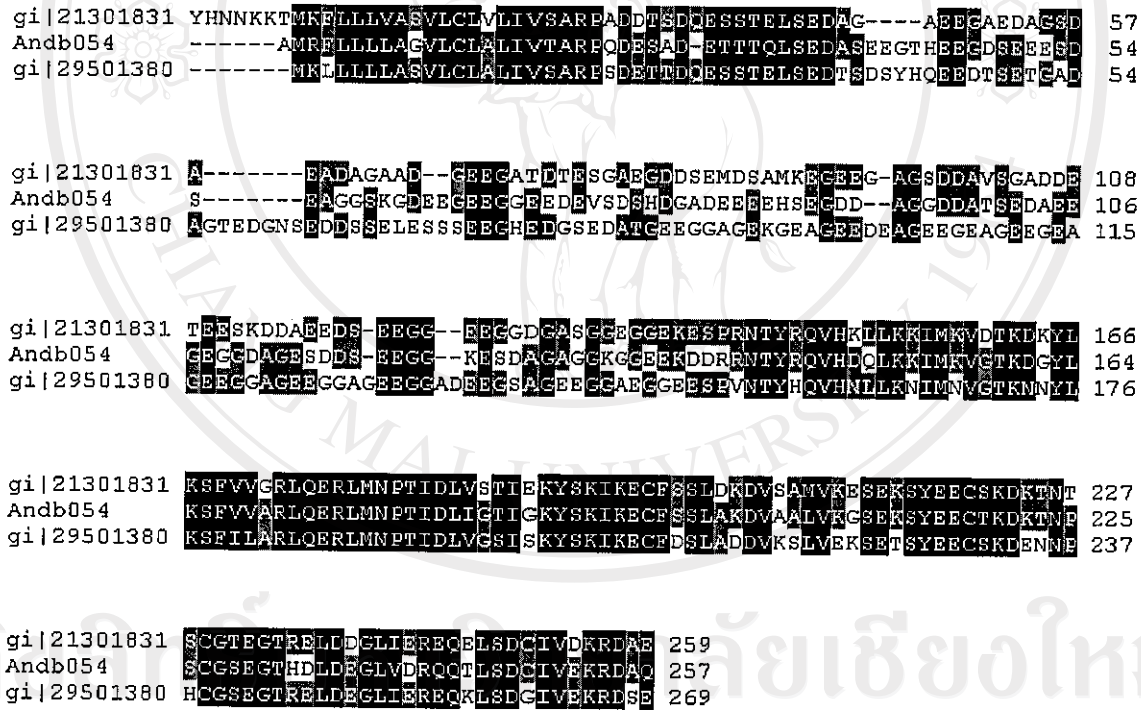

Fig. 8. CLUSTAL alignment of *An. dirus* B GE-rich (Andi054; GenBank accession number AY299327), *An. gambiae* ENSANGP00000022344 (gi|21301831) and *An. stephensi* GE-rich salivary gland protein precursor (gi|29501380). Similar amino acid residues are marked with a gray background, identical amino acids with a black background.

Andi099 cDNA fragment was similar to *An. gambiae* ENSANGP00000021046 with 77% similarity (64% identity) and *An. gambiae* antigen-5 related 2 with 76% similarity (64% identity). No putative signal peptide was found. The deduced Andi099 amino acid sequence had 2 consensus glycosylation sites, Asn^{48} and Asn^{126}. The CLUSTAL alignment of antigen 5-related protein of *An. dirus* B, *An. gambiae* ENSANGP00000021046, *An. gambiae* antigen 5-related 2 and is shown in Fig. 9.

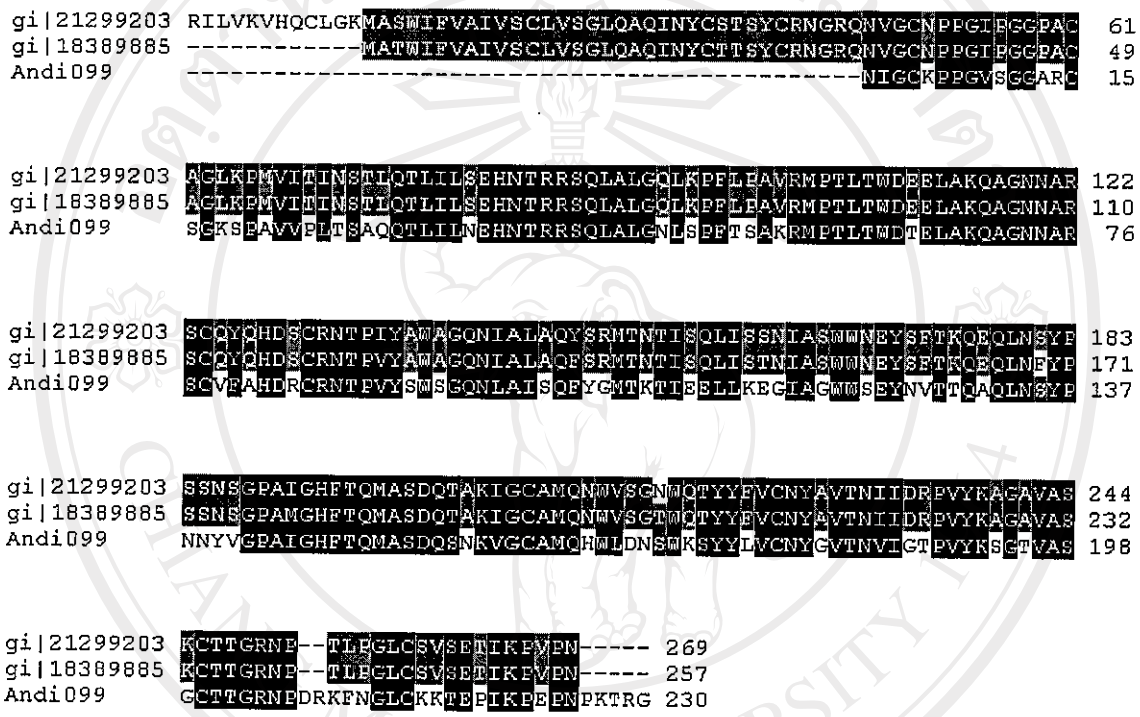

Fig. 9. CLUSTAL alignment of *An. dirus* B antigen 5-related 2 (Andi099; GenBank accession number AY299329), *An. gambiae* ENSANGP00000021046 (gi|21299203) and *An. gambiae* antigen 5-related 2 protein (gi|18389885). Similar amino acid residues are marked with a gray background, identical amino acids with a black background.

Comparison of the N-terminal sequence data shown in Fig. 2 to the amino acid sequences obtained from the salivary gland cDNA library was performed. No match was found for all of these sequences.

Discussions

In this study the overall profiles of female and male salivary gland proteins of *An. dirus* B was analyzed. The protein profiles present in male and female glands are distinctly different. At least 7 major proteins visualized after SDS-PAGE are female specific. One possibility is that the polypeptides not found in males are synthesized by female-specific cells and are involved in blood feeding. The predominant protein bands found in the distal-lateral region and the median lobes of female glands were not present in the glands from non-blood sucking males, which lack these regions. Specific proteins produced in different parts of the salivary glands of female *An. dirus* B are consistent with the previous studies on salivary gland profiles of *An. stephensi* (Suwan *et al.*, 2002) and *Ae. togoi* (Jariyapan *et al.*, 2002).

The protein profile of the salivary glands of sugar-fed female mosquitoes was compared with that of blood-fed ones. The major protein bands in the glands of sugar-fed mosquitoes showed similar profiles with the blood-fed ones. Although, the total salivary gland protein content of blood-fed anopheline mosquitoes (*An. stephensi*, *An. albimanus*, *An. gambiae*, and *An. freeborni*) are at least 10% less than that of unfed control (Golenda *et al.*, 1995), small differences in the amount of proteins are difficult to visualize in Coomassie stained gels and/or silver stained gels. Soliman *et al.* (1999) reported that after *Cx. pipiens* had blood-fed, the total saliva was depleted by 64% within 24 hours, but the protein level returned to the unfed value by the next 24-48 hours. Furthermore, Orr *et al.* (1961) observed change in salivary gland cells 24 hours after *Ae. aegypti* females had taken blood meals; the nucleoli of the median and lateral acini became greatly enlarged and there was a concomitant increase in RNA around the nuclei. The authors concluded that blood feeding may deplete the glands and this depletion leads to resynthesis of secretory products.

Recently, salivary gland transcriptome and proteome of blood-sucking arthropods, *Ae. aegypti* (Valenzuela *et al.*, 2002b), *An. gambiae* (Francischetti *et al.*, 2002b), *An. stephensi* (Valenzuela *et al.*, 2003), and tick (Valenzuela *et al.*, 2002a) were explored. The important aims of these studies were to generate hypotheses on evolution of blood feeding in general and in discovery of novel anti-hemostatic substances and to understand the roles of salivary proteins in host-vector interactions. Presently, N-terminal and internal amino acid sequences of some major proteins of female *An. dirus* B salivary glands were determined. Three N-terminal peptide sequences matched D7-related 1, D7-related 2 and antigen 5 proteins that

were commonly found in salivary glands of several mosquito species. The N-terminal data can be used to design primers to the 5'–terminus of each protein to isolate full-length clones from the *An. dirus* B salivary gland cDNA library. For the internal peptide sequences of 37 kDa protein, Andi054 protein sequence matched the internal peptide sequences (17 out of 17 aa).

From this study, a female salivary gland cDNA library of *An. dirus* B was constructed. Three clones (Andi006, Andi027 and Andi053) were identified having sequence homology to members of salivary gland 1 (SG1) protein family of *An. gambiae*. The proteins called salivary gland (SG) proteins, SG1-8, were first reported in *An. gambiae* by Arca *et al.* (1999). SG1 or gSG1 family of anopheline salivary proteins (Arca *et al.*, 1999; Lanfrancotti *et al.*, 2002) does not yield significant similarity (by BlastP) to other proteins in the NCBI database except among its own members. Recently, Holt *et al.* (2002) constructed two cDNA libraries from adult female *An. gambiae* of same age that were not blood fed and from mosquitoes that blood fed 24 h previously. After analysis of expressed sequence tags sequenced from each library, transcripts for two proteins found in *An. gambiae* salivary glands have increased expression after the blood meal (Ribeiro, 2003). These two transcripts code for proteins of the SG1 family of salivary proteins (Francischetti *et al.*, 2002; Lanfrancotti *et al.*, 2002). Both these transcripts derive from genes closedly located in the X chromosome of *An. gambiae* and are probably under the same transcriptional control. In addition Valenzuela *et al.* (2003) reported 9 salivary proteins of *An. stephensi* as new members of the SG1 family. The function of this protein family is still unknown. So far, the SG1 family has been reported only in anopheline mosquitoes. The new members of SG1 family from *An. dirus* B found in this study consistent with these studies and might help to confirm that the proteins are *Anopheles* specific.

Andi054 amino acid sequence was similar to *An. gambiae* ENSANGP00000022344 sequence and *An. stephensi* GE-rich salivary gland protein precursor. The GE-rich salivary gland protein was first reported in *An. stephensi* (Valenzuela *et al.*, 2003); the sequence did not produce a clear match to any of the translation products of the NCBI database, but closely matched the salivary 30-kDa protein of *An. gambiae*. Presently, the full-length Andi054 cDNA showed more than 80% similarity to that of *An. stephensi*. The results from this study are evidence that Andi054 is homologous with *An. stephensi* GE-rich salivary gland protein precursor. Further analysis may provide insights into the biological function of this molecule.

Another clone called Andi099 showed similarity to *An. gambiae* antigen-5 related 2. Antigen 5 belongs to the larger CAP family of proteins. Closely related proteins from this family have been reported in the salivary glands of Hymenoptera sand flies (Charlab *et al.*, 1999), tsetse (Li *et al.*, 2001), and mosquitoes (Francischetti *et al.*, 2002b; Valenzuela *et al.*, 2002b). They belong to a ubiquitous family of extra-cellular proteins with mostly unknown function (Schreiber *et al.*, 1997). From this study, Andi099 might be another member of this protein family. The full-length sequence of Andi099 gene should be isolated and characterized to provide more information.

The information obtained from this study would be an initial step for further identification and characterization of the salivary proteins in the mosquito species. Expression of these proteins in large amounts and screening for their possible role in multiple bioassays will facilitate understanding how these organisms have adapted to disarm host hemostasis and inflammation, as was recently done for the D7 protein hamadarin (Isawa *et al.*, 2002); Ixolaris, the tissue factor pathway inhibitor of tick, *Ixodes scapularis* (Francischetti *et al.*, 2000a); the tick histamine-binding proteins (Paesen *et al.*, 2000); and *Rhodnius* biogenic amine-binding protein (Andersen *et al.*, 2003). In particular, 3 members of the SG1 family that could be good markers of anopheline exposure, such has been accomplished with ticks (Schwartz *et al.*, 1990) and sand flies (Barral *et al.*, 2000).

References

Arca B, Lombardo F, Capurro M.L, et al. Trapping cDNA encoding secreted proteins from the salivary glands of the malaria vector *Anopheles gambiae*. PNAS 1999;96:1516-21.

Andersen JF, Francischetti IM, Valenzuela JG, Schuck P, Ribeiro JM. Inhibition of hemostasis by a high-affinity biogenic amine-binding protein from the saliva of a blood-feeding insect. J Biol Chem 2003;278:4611-7.

Aultman KS, Beaty BJ, Walker ED. Genetically manipulated vectors of human disease: a practical overview. Trends Parasitol 2001;17:507-9.

Baimai V, Green CA, Andre RG, Harrison BA, Peyton EL. Cytogenetic studies of some species complexes of *Anopheles* in Thailand and Southeast Asia. Southeast Asian J Trop Med Public Health 1984;15:536-46.

Barral A, Honda E, Caldas A, et al. Human immune response to sand fly salivary gland antigens: a useful epidemiological marker? Am J Trop Med Hyg 2000;62:740-5.

Beerntsen BT, James AA, Christensen BM. Genetics of mosquito vector competence. Microbiol Mol Biol Rev 2000;64:115-37.

Catteruccia F, Nolan T, Loukeris TG, et al. Stable germline transformation of the malaria mosquito *Anopheles stephensi*. Nature 2000;405:959-62.

Charlab R, Valenzuela JG, Rowton ED, Ribeiro JM. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagus sand fly *Lutzomyia longipalpis*. PNAS 1999;96:15155-60.

Choochote W, Sucharit S, Abeywickreme W. A note on adaptation of *Anopheles annularis* van Der Wulp, Kanchanaburi, Thailand, to free mating in a 30x30x30 cm cage. Southeast Asian J Trop Med Public Health 1983;14:559-60.

Coates CJ, Jasinskiene N, Miyashiro L, James AA. Mariner transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. PNAS 1998;95:3748-51.

Coleman RE, Polsa N, Eikarat N, Kollars TM Jr, Sattabongkot J. Prevention of sporogony of *Plasmodium vivax* in *Anopheles dirus* mosquitoes by transmission-blocking antimalarials. Am J Trop Med Hyg 2001;65:214-8.

Collins FH, James AA. Genetic modification of mosquitoes. Science Med 1996;3:52-61.

Crampton JM. Molecular studies of insect vectors of malaria. Adv Parasitol 1994;34:1-31.

Curtis CF, Graves PM. Methods for replacement of malaria vector populations. J Trop med Hyg 1988;91:43-8.

Frances SP, Klein TA, Wirtz RA, Eamsila C, Pilakasiri C, Linthicum KJ. *Plasmodium falciparum* and *P. vivax* circumsporozoite proteins in *Anopheles* (Diptera:Culicidae) collected in eastern Thailand. J Med Entomol 1996;33:990-1.

Francischetti IM, Valenzuela JG, Anderson JF, Mather TN, Ribeiro JM. Ixolaris, a novel recombinant tissue factor pathway inhibitor (TFPI) from the salivary gland of the tick, *Ixodes scapularis*: identification of factor X and factor Xa as scaffolds for the inhibition of factor VIIa/tissue factor complex. Blood 2002a;99:3602-12.

Francischetti IM, Valenzuela JG, Pham VM, Garfield MK, Ribeiro JM. Toward a catalog for the transcripts and proteins (sialome) from the salivary gland of the malaria vector *Anopheles gambiae*. J Exp Biol 2002b;205:2429-51.

Ghosh A, Edwards MJ, Jacobs-Lorena M. The journey of the malaria parasite in the mosquito: hope for the new century. Parasitol Today 2000;16:196-201.

Golenda CF, Starkweather WH, Wirtz RA. The distribution of circumsporozoite protein (CS) in *Anopheles* stephensi mosquitoes infected with *Plasmodium falciparum* malaria. J Histochem Cytochem 1995;38:75-81.

Green CA, Munstermann LE, Tan SG, Panyim S, Baimai V. Population genetic evidence for species A, B, C, and D of the *Anopheles dirus* complex in Thailand and enzyme electropmorphs for their identification. Med Vet Entomol 1992;6:29-36.

Higgins DG, Thompson JD, Gibson TJ. Using CLUSTAL for multiple sequence alignments. Methods Enzymol 1996;266:383-402.

Holt RA, Subramanian GM, Halpern A, et al. The genome sequence of the malaria mosquito *Anopheles gambiae*. Science 2002;298:129-49.

Isawa H, Yuda M, Orito Y, Chinzei Y. A mosquito salivary protein inhibits activation of the plasma contact system by binding to factor XII and high molecular weight kininogen. J Biol Chem 2002;277:27651-8.

Jariyapan N, Harnnoi T. preliminary study of salivary gland proteins of the mosquito *Aedes togoi* (Theobald) Chiang Mai Med Bull 2002;41:21-8.

Klein TA, Harrison BA, Dixon SV, Burge JR. Comparative susceptibility of Southeast Asian *Anopheles* mosquitoes to the simian malaria parasite *Plasmodium cynomolgi*. J Am Mosq Control Assoc 1991;7:481-7.

Lanfrancotti A, Lombardo F, Santolamazza F, et al. Novel cDNAs encoding salivary gland proteins from the malaria vector *Anopheles gambiae*. FEBS Lett 2002;517:67-71.

Li S, Kwon J. Aksoy S. Characterization of genes expressed in the salivary glands of the tsetse fly, *Glossina morsitans morsitans*. Insect Mol Biol 2001;10:69-76.

Meredith SEO, James AA. Biotechnology as applied to vectors and vector control. Ann Parasitol Hum Comp 1990;65:113-8.

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

Oakley AJ, Ketterman A, Wilce MC. Structural biology and its applications to the health sciences. Croat Med J 2001a;42:375-8.

Oakley AJ, Harnnoi T, Udomsinprasert R, Jirajaroenrat K, Ketterman A, Wilce MC. The crystal structures of glutathione S-transferases isozymes 1-3 and 1-4 from *Anopheles dirus* species B. Protein Sci 2001b;10:2176-85.

Orr CWM, Hudson A, West AS. The salivary glands of *Aedes* aegypti: histological-histochemical studies. Can J Zool 1961;39:265-72.

Paesen GC, Adams PL, Nuttall PA, Stuart DL. Tick histamine-binding proteins: lipocalins with a second binding cavity. Biochim Biophys Acta 2000;14882:92-101.

Prapanthadara L, Koottathep S, Promtet N, Hemingway J, Ketterman AJ. Purification and characterization of a major glutathione S-transferase from the mosquito *Anopheles dirus* species B. Insect Biochem Mol Biol 1996;26:277-85.

Prapanthadara L, Promtet N, Koottathep S, Somboon P, Ketterman AJ. Isoenzymes of glutathione S-transferase from the mosquito *Anopheles dirus* species B: the purification, partial characterization and interaction with various insecticides. Insect Biochem Mol Biol 2000;30:395-403.

Prapanthadara L, Ranson H, Somboon P, Hemingway J. Cloning, expression and characterization of an insect class I glutathione S-transferase from *Anopheles dirus* species B. Insect Biochem Mol Biol 1998;28:321-9.

Ribeiro JMC. A catalogue of *Anopheles gambiae* transcripts singnificantly more or less expressed following a blood meal. Insect Biochem Mol Biol 2003;33:865-82.

Ribeiro JMC, Francischetti MB. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. Annu Rev Entomol 2003;48:73-88.

Schreiber MC, Karlo JC, Kovalick GE. A novel cDNA from *Drosophila* encoding a protein with similarity to mammalian cysteine-rich secretory proteins, wasp venom antigen 5, and plant group 1 pathogenesis-related proteins. Gene 1997;191,135-141.

Schwartz BS, Ribeiro JMC, Goldstein MD. Anti-tick antibodies: an epidemiological tool in Lyme disease research. Am J Epidemiol 1990;132:58-66.

Singhasivanon P, Thimasarn K, Yimsamran S, et al. Malaria in tree crop plantations in south-eastern and western provinces of Thailand. Southeast Asian J Trop Med Public Health 1999;30:399-404.

Soliman MA, Abdel-Hamid ME, Mansour MM, Seif AI, Kamel KI, el Hamshary EM. Total salivary gland proteins of female *Culex pipiens* and *Aedes caspius* (Diptera: Culicidae) and their fractionation during adult development and after blood sucking. J Egypt Soc Parasitol 1999;29:619-34.

Somboon P, Morakote N. Infectivity of gametocytes of *Plasmodium falciparum* and *Plasmodium vivax* after storage *in vitro*. Ann Trop Med Parasitol 1990;84:89-91

Suwan N, Wilkinson MC, Crampton JM, Bates PA. Expression of D7 and D7-related proteins in the salivary glands of the human malaria mosquito, *Anopheles stephensi*. Insect Mol Biol. 2002;11:223-32.

Valenzuela JG, Francischetti IMB, Pham VM, Garfield MK, Mather TN, Ribeiro JMC. Exploring the sialome of the tick, *Ixodes scapularis*. J Exp Biol 2002a;205:2843-64.

Valenzuela JG, Pham VM, Garfield MK, Francischetti IM, Ribeiro JMC. Toward a description of the sialome of the adult female mosquito *Aedes aegypti*. Insect biochem mol Biol 2002b;32:1101-22.

Valenzuela JG, Francischetti IMB, Pham VM, Garfield MK, Ribeiro JMC. Exploring the salivary gland transcriptome and proteome of the *Anopheles stephensi* mosquito. Insect Biochem Mol Biol 2003;33:717-32.

Walton C, Chang MS, Handley JM, et al. The isolation and characterization of microsatellites from *Anopheles dirus* mosquitoes. Mol Ecol 2000a;9:1665-7.

Walton C, Handley JM, Collins FH, et al. Genetic population structure and introgression in *Anopheles dirus* mosquitoes in Southeast Asia. Mol Ecol 2001;10:569-80.

Walton C, Handley JM, Kuvangkadilok C, et al. Identification of five species of the *Anopheles dirus* complex from Thailand, using allele-specific polymerase chain reaction. Med Vet Entomol 1999;13:24-32.

Walton C, Handley JM, Tun-Lin W, et al. Population structure and population history of *Anopheles dirus* mosquitoes in Southeast Asia. Mol Biol Evol 2000b;17:962-74.

Appendix

Five cDNA sequences isolated from an *Anopheles dirus* B salivary gland cDNA library.

cDNA sequences : The first translated amino acid is listed in bold and underlined. The termination codon (TGA or TAG or TAA) is listed in bold. The consensus polyadenylation signal sequence (AATAAA) is underlined.

Clone Andi006 : 1269 base pairs

CTCTCGAACGGTGGCTTGCCGTGCTTTGGACGAGCGCACTCCTGCTCGCGGCAGGCTGCCGCGGCAAGTTCGGCACGCCCGCA
CCGATCGTCGATGCGGGCATGTGCCTGTCGCAGGCGGGAGCCATCCGGGACGTCCCGCCACTGGTGCACTGTCAGCCGCTCGG
CGTCTGCGACGCGACGTTGCACCAGGCGTACGTCGCGCTACAGCAACAGTGTCGCACGAACGACACCGCCCGGAAGGCGGCGC
GCGAATCGTTCCTGTTCCAGGTGAAGTACTGGCAGGGCGATCACGTCTTCCTGCACGGGCTGCTGCAGAAGTCGGTCGCCCGG
TACGTAGCGGTGCGGGCGGAGGCGCCGCAGGCGGTCGCGCAGAACCGCCGGCTCGTGCAGGAGCTGCGCCGAAAGCTGCTCGT
GTACAGCATCGAGGCGGGCCGGACGGAGGACGCGGCGGTGCTGTACCTGTCGATGCGGGAGCCACCCGCCACCAAGCAGCTGC
TGGGCGACCTCGCCACGCACGACGCCCTGCACGAGGTGCTGCTCGAGCACGTGCTGACGTTCGTGCGCGCCCTGCCGACGGCC
GACCTGCGGAAGGAGCTGTACAAGGGCATCATGCCGCTGCTGGACCGCAACGGGCTAAAGCGGACGTACGCGGTGCTCGTGTA
CGTCGTCGACGCCGTCGGTCTCTTCCCGGAGAAGAACGAGCTGCTCCGCTACATCGGGCAGCCGGTCCGCGAGATGGTCGTGC
GGCTGCGCGACCAGATGGTCCGCACGCAGTACGACGAGAACGCGTGGATCGCGAACAACTATCCGCACTACTACACGTACTTC
ATCAGGGAGCTCACCACGTTCGCCACCGGCGTGTGGGAGGGCATGGAGAAGCTGCGGTTCTACGAGTTCTCGAGCATGCTGAC
GAACAAGCTGCACCGGTTCACGGTGCTGGAGGCGGGGATGAAGATGTTCGAGAAGCACGACTACCAGCACACGCGCCAGTCGG
ACCTGCTGCCGACGCTGGCGGTCGAGGTGGACAAGCTGCAGGTGACGGTGCGGCAGACGGGCAACCAGCAGAAGGAGCGCATG
CGGATGAAGCACCTGCAGGACCGGTTCGGGATGGGCGGGTGGTTCACGCGACGCTACAAACACTACCTGGTGCACGCTCAGCA
CAAGGGCAAGTTCGGCACGCACAATCTGGAACTGGAGCGCATCGGATAGGGGTACGGGGGTTGTCAAGCAATAAAGAGAACGG
CCTATGTCGGCACGTAAAAAAAAA

Clone Andi006 : 402 amino acids

LERWLAVLWTSALLLAAGCRGKFGTPAPIVDAGMCLSQAGAIRDVPPLVHCQPLGVCDATLHQAYVALQQQCRTNDTARKAAR
ESFLFQVKYWQGDHVFLHGLLQKSVARYVAVRAEAPQAVAQNRRLVQELRRKLLVYSIEAGRTEDAAVLYLSMREPPATKQLL
GDLATHDALHEVLLEHVLTFVRALPTADLRKELYKGIMPLLDRNGLKRTYAVLVYVVDAVGLFPEKNELLRYIGQPVREMVVR
LRDQMVRTQYDENAWIANNYPHYYTYFIRELTTFATGVWEGMEKLRFYEFSSMLTNKLHRFTVLEAGMKMFEKHDYQHTRQSD
LLPTLAVEVDKLQVTVRQTGNQQKERMRMKHLQDRFGMGGWFTRRYKHYLVHAQHKGKFGTHNLELERIG

Sequences producing significant alignments:		Score (bits)	E Value
gi|29501536|gb|AAO74845.1|	SG1D salivary protein precursor ...	247	2e-64
gi|18873404|emb|CAA76824.2|	hypothetical protein [Anopheles...	230	3e-59
gi|30177827|gb|EAA06620.2|	ENSANGP00000023463 [Anopheles ga...	229	5e-59
gi|21294389|gb|EAA06534.1|	ENSANGP00000019238 [Anopheles ga...	58	3e-07
gi|21294374|gb|EAA06519.1|	ENSANGP00000007337 [Anopheles ga...	54	4e-06
gi|4210615|emb|CAA10258.1|	SG1 protein [Anopheles gambiae]	52	2e-05
gi|18389895|gb|AAL68784.1|AF457554_1	salivary gland 1-like ...	51	3e-05
gi|27372941|gb|AAO06844.1|	putative salivary protein SG1C [...	45	0.002
gi|27372929|gb|AAO06838.1|	putative salivary protein SG1B [...	44	0.003

Clone Andi027 : 1233 base pairs

GATCGCCGGACTTACTGTGGGGGTGCTGCTGGCCGCTGTTCGGGCCGATGGGCTCCCGGAGTCATCGGCCAACCTGGACGCGT
GCGGGCTAAGCTACGACGAGCTGACGAAAGCGTCCAAGGAGTGGCACGACAAGTCGTGCGAGGGGCAGCCGAAGCTCGCCGCG
TGCGTGGTGCCAGCGCACGAGCAGTCCTACCAGCAGCTAAAGGAGCGCTGCCGGCAGGCGTACGAGCAGCGCGCGTCGCAGGC
TACCAATGTCCACCGCAACCTGACGCAGCTCATTGCCGACGTGATCCGGGGCTCCACCCAGCTCGGCGAGACGATCCGGCACG
ACCTGCCCGGGCTGCAGAGCATCGTGGAGACGCAGAAGCAGCAGCTGGAGGACGGCTGGCGTTATGGAGCGCAGCTGCAGCGG
GAGCTGCTCCTCACCAGCATGGAGTCGGGCCGTACGGAGCGGGGGCTCGTGCTCCACTCGCTGCTGGTGAACGCATCGCTGCG
GGAGATGCTGCAGGAGAGCTACCGGTACCACGGAGACGACGGTCGTATGGTCGCGCGGATGTTGCAGTTCGTCGTGCGGTTGCTGC
CTGCGGCGGACGAGCGCGTTCACCTCTACAAGCAGCTGGCGGAGCTGCTCCAAACCAACGCGCAGGACGAGCGCTTCCCGGCC
GTCATCTTCTCGGCGGACGTGCGGCCTCTGAAGGACCAGTACGCACCGGAGCACGCACTGTACGAGGGTAAGGCGGTCGCCCG
CTGGCAGTCGCAGCTGCTCGCCGGCAACTTCACCGAGGTCGCCCTGTTCGCCCGCGACTTCCCGGCGTACTTCGCCGGCGTCG
AGGACGCGCTGTACGCGGCGCTCAAGCAGCAGTGGTCGGTCGAGGGCCTGCAGCACATGCTCCGGCTGCCCGCCGCCCTCCCC
GGGACCGGGCAGCGCGTCCGTGCCGTTCGCACCATGCTGGAGGCACTGCTGCAGCATCAGAACGAGCAGCGCAACGACCCGCA
CCTGATGCGGCTCGCACACCTGCTGGCCGCACTGGACGCTGATCTCCCCAAGGACGACAAAGCCGCCCAGCAGACGCTGGAAG
ACGCCAAGCAGCTGTTCGGGCAGTTCGAGTTCAAGCGCGACTTCGCCGCGTACGTCGACCTGTACGGGCTGCTCAAGGCAGCG
CTCTAACCGCACAGCACAATAAACCACCGCTTTTCGGGGATCGATTTCACCGCTACACCGCAAAAAAAAAA

Clone Andi027 : 388 amino acids

IAGLTVGVLLAAVRADGLPESSANLDACGLSYDELTKASKEWHDKSCEGQPKLAACVVPAHEQSYQQLKERCRQAYEQRASQA
TNVHRNLTQLIADVIRGSTQLGETIRHDLPGLQSIVETQKQQLEDGWRYGAQLQRELLLTSMESGRTERGLVLHSLLVNASLR
EMLQESYRYHGDDGRMVARMLQFVRLLPAADERVHLYKQLAELLQTNAQDERFPAVIFSADVRPLKDQYAPEHALYEGKAVAR
WQSQLLAGNFTEVALFARDFPAYFAGVEDALYAALKQQWSVEGLQHMLRLPAALPGTGQRVRAVRTMLEALLQHQNEQRNDPH
LMRLAHLLAALDADLPKDDKAAQQTLEDAKQLFGQFEFKRDFAAYVDLYGLLKAAL

Sequences producing significant alignments:	Score (bits)	E Value
gi|21294389|gb|EAA06534.1| ENSANGP00000019238 [Anopheles ga...	388	e-107
gi|18389895|gb|AAL68784.1|AF457554_1 salivary gland 1-like ...	291	1e-77
gi|30267888|gb|AAP21784.1| secretion protein gp65 [Anophele...	155	1e-36
gi|4127309|emb|CAA76820.1| hypothetical protein [Anopheles ...	97	3e-19
gi|27372929|gb|AAO06838.1| putative salivary protein SG1B [...	72	1e-11
gi|21294374|gb|EAA06519.1| ENSANGP00000007337 [Anopheles ga...	66	1e-09
gi|4210615|emb|CAA10258.1| SG1 protein [Anopheles gambiae]	64	3e-09
gi|29501536|gb|AAO74845.1| SG1D salivary protein precursor ...	58	3e-07

Clone Andi053 : 1300 base pairs

CGATGGAGAAAGTTAGGATGGGCGTACTGCTGCTGGTGCTGCTCGCCACGTTGGCCGGGGCACGACCGCAAGAGACCGATCCG
TTCGTGGACGAACCGGACCAGTGTCTGATCGGCGTCAGTGCGCAGGTCCTCGCCGCGCTGTCCGAGTCGCTGCGTACCGAGCT
CAAGTGCGAGGACCTCTGGAGCAGCGTGCTGCTGCGGTACCACCACACGCGCACCAACCTCACCGAGTGTCTGGCGCGGGCGA
GCGGCGACGCGACGCCCGCCCCGGCCAGCAGCTTCTGCCAGCTGCTGCTCGACGACGTCGAGCGCCAGCTCGACCAGGAGCAC
CGCCAGTCGCTGGCCGACATCGAGCAGAAGCTGCACGTCACCCAGCAGGAGGCGCGCGCCCACCACGACGAGAAGACCGCGCT
GGAGGGCCAGCTGCACCGCCTGCAGGACGACCGGGACACGCTGTACCTGGAGCTGCTGCTCGCGAACATCGCGATCGGTGACG
CGCAGCAGGCGAAGAAGTACTACGAGCTGTACCCCGGCAAAGACCCGGCCGACAGGCTGCACGCGCAGATCGTGCGGTCCGTG
TACCGCGTGGCCAAGTACCAGGACCAGCGGCTGCTCAACCTCGTCCAGTTCGTGCGCACGCTCGCCGGCACCGAGCCGAAGCT
CGGGCTGTACCGGCTGATGCGGGAGGAGATCCTGAAGCGGCCGAGCCAGCGCGACACCTACGTCGCGGCCATCTTCGCGCTCA
GCCTGGGCGCGGACGGGGACGTGCGGGCCCGCGACCCCCGGCTCTACACCGACACGATGGGACCGATCGAGCAGCGCTGGAGG
GACCAGCTGTACAACGGGCAGTTCAACGAGGTGGCCGACTTTGCCCGCCGCTTTGCGACGCAGTTCGCGCAGATGCAGAAGCC
GCTCGCCTACCCGAACGCGCTCCCGCTGCCGCAGCAGCGGCTGGAAGCGTTCCGGCACATACTGGACCAGATCCGGCAGCACA
ACCCGAACAACGCGCACAGCTATCTGGTGCAGACGGCGAAGCAGTTCGACATCTGCGAGACGTTCATCAAGAAGAGCAAGGTG
GACGCCGCCGTCACGCAGACGCTCGAGCAGCTGCGCAAGAGGTTTGCGGAGTTCTCGCGCAAGGAGTACGGCGTGTACCTCCG
GGAGGCGAAGGGATGACCCGTATGGCCCGTGACCTTGACCGCTTTCTTCCTTCCGTAGTCTAGCATCTAAGTGGCACCATGAA
ATAAAACTTGTCCTCGAGGCAGCTCGAGCAGATAGCAGAAAAAAAAAAAAAAAAA

Clone Andi053 : 391 amino acids

MEKVRMGVLLLVLLATLAGARPQETDPFVDEPDQCLIGVSAQVLAALSESLRTELKCEDLWSSVLLRYHHTRTNLTECLARAS
GDATPAPASSFCQLLLLDDVERQLDQEHRQSLADIEQKLHVTQQEARAHHDEKTALEGQLHRLQDDRDTLYLELLLANIAIGDA
QQAKKYYELYPGKDPADRLHAQIVRSVYRVAKYQDQRLLNLVQFVRTLAGTEPKLGLYRLMREEILKRPSQRDTYVAAIFALS
LGADGDVRARDPRLYTDTMGPIEQRWRDQLYNGQFNEVADFARRFATQFAQMQKPLAYPNALPLPQQRLEAFRHILDQIRQHN
PNNAHSYLVQTAKQFDICETFIKKSKVDAAVTQTLEQLRKRFAEFSRKEYGVYLREAKG

Sequences producing significant alignments:	Score (bits)	E Value
gi|27372929|gb|AAO06838.1| putative salivary protein SG1B [...	327	2e-88
gi|21294236|gb|EAA06381.1| ENSANGP00000019156 [Anopheles ga...	301	1e-80
gi|4210615|emb|CAA10258.1| SG1 protein [Anopheles gambiae]	176	5e-43
gi|21294374|gb|EAA06519.1| ENSANGP00000007337 [Anopheles ga...	176	8e-43
gi|21294237|gb|EAA06382.1| ENSANGP00000019154 [Anopheles ga...	132	1e-29
gi|21294389|gb|EAA06534.1| ENSANGP00000019238 [Anopheles ga...	83	7e-15
gi|18389895|gb|AAL68784.1|AF457554_1 salivary gland 1-like ...	81	3e-14
gi|4127301|emb|CAA76813.1| gSG1 protein [Anopheles gambiae]	67	5e-10
gi|3378533|emb|CAA03873.1| D3 protein [Anopheles gambiae]	62	2e-08
gi|27372939|gb|AAO06843.1| putative salivary protein SG1A [...	52	2e-05
gi|18389897|gb|AAL68785.1|AF457555_1 salivary gland 1-like ...	50	5e-05
gi|27372941|gb|AAO06844.1| putative salivary protein SG1C [...	39	0.10
gi|13537664|emb|CAC35521.1| gSG1b protein [Anopheles gambiae]	39	0.17
gi|21294133|gb|EAA06278.1| ENSANGP00000017327 [Anopheles ga...	38	0.23
gi|16418019|gb|AAL18964.1|AF432352_1 putative alpha-1,3-glu...	38	0.24
gi|22988357|ref|ZP_00033423.1| hypothetical protein [Burkho...	36	1.0
gi|30267888|gb|AAP21784.1| secretion protein gp65 [Anophele...	35	2.6
gi|345511|pir||PC1232 copia polyprotein - fruit fly (Drosop...	33	6.6

Clone Andi054 : 863 base pairs

AGCAATGAGGTTCCTACTACTGCTGGCAGGCGTGCTTTGCCTGGCATTAATCGTCACTGCACGGCCACAGGATGAAAGCGCAG
ACGAAACAACGACGCAGCTAAGTGAAGATGCGTCAGAGGAAGGCACACATGAAGAAGGAGATTCGGAGGAAGAATCAGATTCC
GAAGCCGGTGGTAGCAAAGGTGACGAAGAAGGTGAGGAAGGTGGAGAGGAGGACGAAGTAAGTGATTCACATGATGGAGCGGA
TGAGGAAGAAGAGCATTCTGAAGGAGATGATGCGGGCGGCGATGATGCGACAAGTGAAGATGCAGAAGAAGGTGAAGGAGGCG
ATGCTGGGGAAAGTGATGATTCAGAAGAAGGTGGTAAAGAAAGTGATGCTGGAGCAGGTGGAAAGGGCGGTGAAGAAAAGGAT
GATCGCAGAAATACGTACCGTCAGGTGCACGACCAGCTGAAAAAGATCATGAAGGTCGGAACGAAAGACGGGTACCTGAAGTC
GTTCGTTGTGGCCCGCCTCCAGGAGCGTCTGATGAATCCGACAATTGATCTGATAGGCACGATCGGAAAGTACTCGAAGATCA
AGGAATGCTTCAGCTCGCTGGCGAAAGATGTGGCCGCCCTGGTGAAGGGGTCGGAGAAATCTTACGAAGAATGTACCAAGGAT
AAGACCAACCCTAGCTGTGGCAGTGAAGGCACCCATGATCTCGACGAAGGACTCGTCGACCGTCAACAGACTCTGTCCGATTG
CATCGTTGAGAAACGTGATGCACAATAGACGGTCGCAAACTAAAACGCGACTTTGTACGATGCAGCTGAAGAGTTGGCGACAA
ATAAATGCATTATAAAACGATAAAAAAAAAAAA

Clone Andi054 : 257 amino acids

AMRFLLLLAGVLCLALIVTARPQDESADETTTQLSEDASEEGTHEEGDSEEESDSEAGGSKGDEEGEEGGEEDEVSDSHDGAD
EEEEHSEGDDAGGDDATSEDAEEGEGGDAGESDDSEEGGKESDAGAGGKGGEEKDDRRNTYRQVHDQLKKIMKVGTKDGYLKS
FVVARLQERLMNPTIDLIGTIGKYSKIKECFSSLAKDVAALVKGSEKSYEECTKDKTNPSCGSEGTHDLDEGLVDRQQTLSDC
IVEKRDAQ

Sequences producing significant alignments:	Score (bits)	E Value
gi|21301831|gb|EAA13976.1| ENSANGP00000022344 [Anopheles ga...	165	6e-40
gi|29501380|gb|AAO74840.1| GE rich salivary gland protein p...	158	6e-38
gi|18568322|gb|AAL76031.1|AF466608_1 putative 30 kDa allerg...	59	1e-07
gi|14423642|sp|O01949|ALL3_AEDAE 30 kDa salivary gland alle...	55	1e-06
gi|18389879|gb|AAL68776.1|AF457546_1 30 kDa protein [Anophe...	35	0.98

Clone Andi099 : 799 base pairs

AAACATCGGTTGCAAGCCTCCCGGTGTTTCGGGTGGTGCCCGTTGCAGTGGTAAGAGTCCCGCCGTTGTACCGCTCACCTCCG
CACAGCAAACGCTTATCCTGAACGAGCACAACACTCGTCGGTCCCAGCTCGCCCTGGGCAATCTGAGTCCGTTCACGTCGGCA
AAGCGCATGCCCACGCTCACCTGGGACACGGAGCTGGCAAAGCAGGCGGGTAACAATGCGCGAAGCTGTGTCTTCGCGCACGA
CAGATGTCGCAACACGCCCGTGTACAGCTGGTCCGGACAGAATCTCGCTATATCACAGTTCTACGGTATGACCAAGACGATCG
AGGAGTTGCTGAAAGAAGGCATTGCAGGTTGGTGGAGCGAGTACAACGTTACCACCCAAGCTCAACTGAACAGCTATCCCAAC
AATTATGTGGGCCCGGCAATCGGACACTTCACTCAGATGGCCAGCGATCAGTCGAACAAAGTGGGATGCGCCATGCAGCACTG
GTTGGATAATTCGTGGAAATCCTACTACTTGGTCTGCAACTACGGTGTCACGAATGTTATCGGCACTCCGGTCTATAAGAGTG
GTACTGTTGCATCCGGCTGTACCACTGGGCGTAACCCAGATCGAAAGTTCAACGGCCTGTGCAAGAAAACCGAACCCATCAAA
CCTGAACCAAATCCGAAAACTCGGGGCTGAATTCTTCAATAACGGGCAGTTTACAGTAACACGTTTGAGCGTCGTTAAAAGTC
ACCAAAACCAATAAATAAAATGAGCTTAAATCTAAAAAAAAAAAAAAAAAAA

Clone Andi099 : 230 amino acids

NIGCKPPGVSGGARCSGKSPAVVPLTSAQQTLILNEHNTRRSQLALGNLSPFTSAKRMPTLTWDTELAKQAGNNARSCVFAHD
RCRNTPVYSWSGQNLAISQFYGMTKTIEELLKEGIAGWWSEYNVTTQAQLNSYPNNYVGPAIGHFTQMASDQSNKVGCAMQHW
LDNSWKSYYLVCNYGVTNVIGTPVYKSGTVASGCTTGRNPDRKFNGLCKKTEPIKPEPNPKTRG

Sequences producing significant alignments:	Score (bits)	E Value
gi|21299203|gb|EAA11348.1| ENSANGP00000021046 [Anopheles ga...	304	7e-82
gi|18389885|gb|AAL68779.1|AF457549_1 antigen 5-related 2 pr...	301	7e-81
gi|21299230|gb|EAA11375.1| ENSANGP00000004285 [Anopheles ga...	295	4e-79
gi|21298576|gb|EAA10721.1| ENSANGP00000020404 [Anopheles ga...	284	8e-76
gi|21299047|gb|EAA11192.1| ENSANGP00000020471 [Anopheles ga...	284	8e-76
gi|30176055|gb|EAA11396.2| ENSANGP00000021028 [Anopheles ga...	220	1e-56
gi|27372895|gb|AAO06821.1| salivary antigen-5 related prote...	207	8e-53
gi|21299125|gb|EAA11270.1| ENSANGP00000021178 [Anopheles ga...	197	2e-49
gi|18568308|gb|AAL76024.1|AF466601_1 putative secreted prot...	192	4e-48
gi|18568284|gb|AAL76012.1|AF466589_1 putative secreted prot...	186	2e-46
gi|21294230|gb|EAA06375.1| ENSANGP00000019483 [Anopheles ga...	179	2e-44
gi|30176054|gb|EAA44225.1| ENSANGP00000025115 [Anopheles ga...	159	2e-38