Published in final edited form as:

Mol Cell Biochem. 2009 June ; 326(1-2): 67–77. doi:10.1007/s11010-008-0007-z.

Characterization of the yellow fever mosquito sterol carrier protein-2 like 3 gene and ligand-bound protein structure

David H. Dyer^{1,2}, Irina Vyazunova², Jeffery M. Lorch², Katrina T. Forest¹, and Que Lan²

¹ Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706, USA

² Department of Entomology, University of Wisconsin-Madison, 1630 Linden Drive, Madison, WI 53706, USA e-mail: qlan@entomology.wisc.edu

Abstract

The sterol carrier protein-2 like 3 gene (*AeSCP-2L3*), a new member of the SCP-2 protein family, is identified from the yellow fever mosquito, *Aedes aegypti*. The predicted molecular weight of *AeSCP-2L3* is 13.4 kDa with a calculated *pI* of 4.98. *AeSCP-2L3* transcription occurs in the larval feeding stages and the mRNA levels decrease in pupae and adults. The highest levels of *AeSCP-2L3* gene expression are found in the body wall, and possibly originated in the fat body. This is the first report of a mosquito SCP-2-like protein with prominent expression in tissue other than the midgut. The X-ray protein crystal structure of *AeSCP-2L3* reveals a bound C16 fatty acid whose acyl tail penetrates deeply into a hydrophobic cavity. Interestingly, the ligand-binding cavity is slightly larger than previously described for *AeSCP-2* (Dyer et al. *J Biol Chem* 278:39085–39091, 2003) and *AeSCP-2L2* (Dyer et al. *J Lipid Res* M700460–JLR200, 2007). There are also an additional 10 amino acids in *SCP-2L3* that are not present in other characterized mosquito SCP-2s forming an extended loop between $\beta 3$ and $\beta 4$. Otherwise, the protein backbone is exceedingly similar to other SCP-2 and SCP-2-like proteins. In contrast to this observed high structural homology of members in the mosquito SCP2 family, the amino acid sequence identity between the members is less than 30%. The results from structural analysis imply that there have been evolutionary constraints that favor the SCP-2 C α backbone fold while the specificity of ligand binding can be altered.

Keywords

Sterol carrier; Mosquito; Cholesterol; Fatty acid

Introduction

Sterol carrier protein-2 (SCP-2) is known as a nonspecific intracellular lipid carrier. The majority of the members of the SCP-2 protein family are multidomain proteins [1]. For example, within the human genome only two out of seven genes in the SCP-2 family, SCP-x/SCP-2 (ENSG00000116171) and the SCP-2-like (ENSG00000132631) are present which produce single SCP-2 domain proteins. However, four genes encoding single SCP-2 domain proteins have been identified in the yellow fever mosquito *Aedes aegypti* [2–5]. It is evident that single SCP-2 domain genes are expanded in mosquitoes. Whether those mosquito SCP-2 and SCP-2-like proteins are functionally divergent or redundant is unknown. The mosquito SCP-2 (*AeSCP-2*) and SCP-2-like proteins have very similar temporal and spatial expression

Correspondence to: Que Lan.

Electronic supplementary material The online version of this article (doi: 10.1007/s11010-008-0007-z) contains supplementary material, which is available to authorized users.

profiles [2,5], but vary in their affinity to different lipids [2,4]. The in vivo physiological roles of these mosquito SCP-2 and SCP-2-like proteins remain unclear. However, it is evident that mosquito SCP-2 and SCP-2-like proteins differentially aid lipid absorption in cultured cells and in vivo. While AeSCP-2 contributes to the uptake of cholesterol from a blood meal in female mosquitoes; on the other hand, SCP-2-like 2 (AeSCP-2L2) only affects the uptake of free fatty acid [6]. A comparison between three-dimensional protein structures of mosquito AeSCP-2 and AeSCP-2L2 also reveals another interesting difference. AeSCP-2 binds to fatty acids as a monomer [7], whereas AeSCP-2L2 binds to fatty acids as a homodimer at a ratio 2:3 of protein:ligand [6].

We found an annotated gene (AAEL012704) in the *Ae. aegypti* genome that codes for a predicted protein with homology to the described AeSCP-2 and AeSCP-2L2 proteins, and speculated that it represented another member of the mosquito SCP-2 family. To investigate this, we isolated the complete cDNA by 5' and 3' rapid amplification of cDNA ends (RACE) using primers matching to AAEL012704 sequence. Our results indicated that AAEL012704 codes for the mosquito SCP-2-like 3 protein (AeSCP-2L3), a new member in the mosquito SCP-2 protein family. AeSCP-2L3 is a single SCP-2 domain protein that has only an overall identity of 22 and 29% to AeSCP-2L2 and AeSCP-2, respectively. We also analyzed the expression profiles of AeSCP-2L3 in the yellow fever mosquito and solved the three-dimensional protein structure of AeSCP-2L3 by means of X-ray crystallography (to a resolution of 1.4 Å with synchrotron radiation under cryo-cooled conditions and at room temperature to 2.1 Å resolution). The results of our analysis shed light on the complex divergent expression regulation of the mosquito SCP-2-like genes and the different modes of ligand/protein interactions among SCP-2 protein structures.

Materials and methods

Chemicals

Unless otherwise mentioned in the text, all chemicals and reagents were purchased from Sigma (Sigma, St. Louis, MO), Fisher Scientific (Pittsburgh, PA), or ICN (Costa Mesa, CA).

Mosquitoes

The mosquitoes used in these experiments, *Ae. aegypti*, were taken from an inbred laboratory strain (Rockefeller). The colony was maintained as described in Krebs and Lan [2]. Larvae and pupae were staged as described in Krebs and Lan [2].

RNA extraction

Ten staged animals were washed with ddH₂O, rinsed once with DEPC-H₂O, and the excess water blotted off using a clean Kimwipe. Cleaned animals were put into a 1.5 ml Eppendorf test tube and homogenized with a micropestle in 1 ml of Trizol reagent according to the Manufacturer's instructions (Invitrogen Corporation, Carlsbad, CA).

Tissues were dissected in insect saline solution [8] under the dissecting microscope and immediately transferred into an Eppendorf tube containing 0.5 ml of Trizol reagent. Tissues from more than 20 animals per time point were collected and pooled. The samples were then homogenized with a micropestle and another 0.5 ml of Trizol reagent added. The resulting total RNA samples were stored at -80°C until use.

Isolation of AeSCP-2 cDNA

We found that AAEL012704 codes for a predicted protein displaying high homology to AeSCP-2 (51% similarity) in the *Ae. aegypti* genome. We speculated the AAEL012704 represents a new mosquito SCP-2-like protein gene. Two primers were designed to match the

N- and C-terminal sequences based on the genomic DNA information: 5'-ATGGCTCTAAAAACGGATCAAA-3' (AeSCP-2L3 forward primer) and 5'-TTATTTCTGAACTTTTCGTCTAGAGTT-3' (AeSCP-2L3 reverse primer). The Smart RACE cDNA amplification Kit (ClonTech, Palo Alto, CA) was used for 5' and 3' RACE with cDNAs made from 24-h-old fourth instar larvae. PCR was carried out using Phusion Master Mix (NBL). PCR conditions were as following: one cycle at 98°C for 1', 35 cycles of 98°C 10"/58°C 20"/72°C 1', and followed by one cycle at 72°C for 1'. PCR products were cloned into pCR-Blunt II TOPO (Invitrogen Corporation, Carlsbad, CA), transformed into *E. coli* TOP10 strain (Invitrogen), and plated onto LB plates containing kanamycin. Plasmid minipreps of 4–6 clones containing inserts were performed using the QiaSpin columns (QIAGEN, Valencia, CA) and sequenced in an automatic sequencer (ABI 377XL) with BigDye labeling (Amersham Pharmacia Biotech AB, Uppsala, Sweden).

Dot blot analysis

For dot blot analysis, 5 µg of total RNA from three parallel samples (10 larvae per sample) of each developmental time point were blotted onto a positively charged nylon membrane (Millipore) as described [9]. The membrane was rinsed in 2× SSC buffer at room temperature for 5 min and air-dried. RNA was permanently attached to the membrane via a UV cross-linker (Stratagene Cloning System, La Jolla, CA). The membranes were baked at 80°C under a vacuum for 1 h.

Primers synthesized for AeSCP-2L3 5' and 3' RACE were used to produce gene-specific probes. Radioactive probes were made using [α -³²P] dCTP (3,000 Ci/mM) in a PCR reaction with DNA polymerase *Tfl* (Epicentre, Madison, WI). The PCR reaction was performed under one cycle of 94°C, 3 min to denature the template, then three cycles of 94°C for 2 min, 50°C for 3 min, and 72°C for 5 min. The probes were cleaned using a QiaSpin columns (QIAGEN).

Membranes prehybridization, hybridization, and washing were performed as described in Krebs and Lan [2]. The membranes were then exposed to a Phosphor Image Screen for at least 1 h and scanned with a Storm820 (Molecular Dynamics, Piscataway, NJ). The data were analyzed using the Storm820 software.

Semi-quantitative RT-PCR analysis

Total RNA samples (1.5 µg each sample) were treated with DNase I twice using a kit (Ambion, Austin, TX) to eliminate genomic DNA contamination. Cleaned total RNA samples were reverse transcribed using a reverse transcription kit (Applied Biosystems, Foster City, CA) to generate the first strand cDNAs. For semi-quantitative RT-PCR, 1.5 µl first strand cDNA representing 0.5 µg of total RNA from each sample were used in the PCR reaction. Primers for AeSCP-2L3 transcript were 5'-atggctctaaaaacggatcaaa-3' (forward) and 5'-ttatttctgaacttttcgtctagagtt-3' (reverse). Mosquito muscle actin-2 and ribosome protein L8 were used as internal controls for the PCR reaction. Primers for AeAct-2 were 5'-cg aactcctccagccactac-3' (forward) and 5'-gcagtttcttagcgggtgtc-3' (reverse). Primers for Ae-rpL8 were 5'-tacctgaaggg aaccgtcaagcaa-3' (forward) and 5'-acaatggtacctcgggcacaga-3' (reverse). The PCR conditions for AeSCP-2L3 were one cycle at 98°C for 1', 40 cycles at 98°C for 20", 58°C for 30", and 72°C for 30"; one cycle at 72°C for 1'. For AeAct-2 and Ae-rpL8 the number of amplification cycles was 30.

Production of AeSCP-2L3 recombinant proteins

The coding region of the AeSCP-2L3 was cloned into the pGEX-4T-2 GST tag vector (Amersham Pharmacia, Biotech AB, Uppsala, Sweden). PCR primers were: 5'-gtgaattcgaatggctctaaaaacggatcaaa-3' and 5'-tactcgagttatttctgaacttttcgtctagagtt-3' (bold letters represent coding sequence and the stop codon; *Eco*RI and *Xho*I sites were incorporated

for cloning). Sequence analysis was performed to confirm that the fusion protein was in frame with GST. Purification of recombinant AeSCP-2L3 was carried out as described in Dyer et al. [7]. The predicted molecular weight of recombinant AeSCP-2L3 is 14,134 Da, and the molecular weight of purified recombinant AeSCP-2L3 determined via MS was 14,165 Da. The MS was performed at the Mass Spectrometer/Proteomics facility, Biotechnology Center, University of Wisconsin-Madison. Purified recombinant AeSCP-2L3 was concentrated to 8 mg/ml in phosphate saline buffer (PBS) pH 7.4 then subsequently stored at -80°C .

Production of AeSCP-2L3 antibodies

Purified rAeSCP-2L3 (8 mg/ml) was sent to Immunology Services (Covance Research Products Inc., Denver, PA) for the production of rabbit polyclonal anti-AeSCP-2L3 antibodies. Anti-AeSCP-2L3 serum was tested against rAeSCP-2L3 by ELISA to ensure that a high titer of AeSCP-2L3 antibodies was obtained. After two boosts using rAeSCP-2L3, anti-AeSCP-2L3 serum was partially purified by passing it through a GST column to remove anti-GST antibodies. Partially purified anti-AeSCP-2L3 antibodies did not cross-react to the bacterial GST in Western blotting analysis (data not shown).

Western blot analysis

Dissected tissues from intact mosquitoes were homogenized in lysis buffer [0.25 M Tris-HCl, pH 8.0/0.2% Triton X-100/1 mM dithioerythritol/5 mM EDTA/10 mM β -mercaptoethanol/1 mM phenylmethylsulfonyl fluoride/protease inhibitor cocktail (Sigma)], and centrifuged at $12,000\times g$ at 4°C for 15 min. Supernatant containing soluble protein was stored at -80°C . Protein concentrations were determined using a BCA kit (Pierce, Rockford, IL).

Western blotting analysis was performed as described in McPherson [10] using either SDS 15% PAGE or SDS 4–20% gradient PAGE (ISC BioExpress, Kaysville, UT). Affinity purified anti-AeSCP-2 primary antibodies [4] were used at a 1:2,000 dilution; partially purified anti-AeSCP-2L3 antibodies were used at a 1:500 dilution; goat anti-rabbit horse radish peroxidase conjugated (for anti-AeSCP-2) or alkaline phosphatase conjugated (for anti-AeSCP-2L3) secondary antibodies (Jackson ImmunoResearch Laboratories, Inc., West Grove, PA) were used at a 1:2,000 dilution. Developing solution was used to visualize the bound antibodies, which developed within 10 min at room temperature as described in Ausubel et al. [9]. For Western blotting of two different proteins on the same blot, the blot was completely dried after staining for the first protein and then blotted for the second protein with different conjugated secondary antibody.

Crystallization and data collection

Crystals of AeSCP-2L3 were grown at room temperature using the hanging drop technique [10] and first identified from a sparse matrix screen based on the University of Toronto structural genomics minimal 24 screen [11]. Two datasets for AeSCP-2L3 were collected. One of them was at The Advanced Photon Source LSCAT ID-21 to $1.4\text{ \AA}$ under a cryo-cooled nitrogen stream (-180°C), and the other was done in-house to $2.1\text{ \AA}$ on a Proteum CCD detector (Bruker AXS) at room temperature (23°C) by means of capillary mounting. The space group was identified as P3221 for both crystals although they were grown under different conditions.

To grow the crystal used for the cryo-cooled structure, $2\text{ }\mu\text{l}$ AeSCP-2L3 at 10 mg/ml was combined with $2\text{ }\mu\text{l}$ of the well solution containing 25% polyethylene glycol 8000, 200 mM ammonium sulfate, and 200 mM ammonium fluoride with a final pH = 7.5. Useful crystals could be grown over a 3- to 4-day period, which yielded unit cell constants of $a = 62.99\text{ \AA}$, $b = 62.99\text{ \AA}$, $c = 67.98\text{ \AA}$, and $\gamma = 120.0^{\circ}$. Optimization of freezing conditions included the addition of 500 mM NaCl to a synthetic mother liquor of 25% polyethylene glycol 8000, 200 mM ammonium sulfate, 15% glycerol (pH = 7.5), which resulted in a significantly lower

mosaic spread than when native crystals were flash cooled directly from mother liquor. Data were processed and scaled for the cryo-temperature structure using HKL-2000 software [12].

To grow the crystal used for the room temperature structure, 2 μ l AeSCP-2L3 at 10 mg/ml was combined with 2 μ l of the well solution consisting of 2 M ammonium sulfate and 100 mM Bis-Tris pH = 5.5. This crystal was identified 4 months after initial screening with the JCSG commercial sparse matrix screen (Qiagen, Valencia, CA, USA) and required a minimum of 2 months to grow to dimensions of $0.5 \times 0.4 \times 0.4$ mm. For the room temperature structure the unit cell (constants of $a = 63.33$ Å, $b = 63.33$ Å, $c = 68.41$ Å, and $\gamma = 120.0^\circ$). Data were processed and scaled for the room temperature structure using Proteum 2 software (Bruker AXS).

Both the room temperature and cryo-cooled structures were solved with the molecular replacement program Phaser [13,14] utilizing a poly-alanine model of AeSCP-2 (pdb entry 1PZ4). The initial phases were significantly improved with ARP/wARP [15] by switching on the option “improvement of maps by atoms update and refinement.” Automatic model building was performed with the new phases and ARP/wARP had the ability to build over 90% of each structure. The models were further improved with iterative cycles of refinement with Refmac5 [16] and manual building with Xfit to produce the final structures.

Residues not observed in the electron density include Met¹ and the expression tag residues (Gly⁻⁶, Pro⁻⁵, Ser⁻⁴, Gly⁻³, Ile⁻², and Arg⁻¹) which remained after thrombin cleavage of the recombinant protein. All other residues are observed continuously from Ala2 to the C-terminal residue Lys120.

The final R-factor for the cryo-temperature structure is 18.5% (R-free with 5% of data omitted is 21.8%). For the room temperature structure R and R-free are 16.6 and 18.9%, respectively. Omit maps reduced model biases and clearly showed the positions of the one fatty acid identified as palmitic acid based upon the clearly defined density. Structure and coordinate analysis with Procheck [17] revealed 92.9% (91.2% in the room temperature dataset) of the nonglycine, nonproline amino acids lie within the most favored regions of the Ramachandran plot, with all remaining residues in the additionally allowed region.

Results

AeSCP-2L3 gene

We cloned the cDNA of AeSCP-2L3 based on the sequence information of the predicted gene represented by AEEL012704. Complete cDNA sequence was obtained via 5' and 3' RACE. Four out of seven independent cDNA sequences (EU357904) were the same in the 5'UTR from the EST sequences, three out of seven independent cDNAs had different 5'UTR, which suggest IntronI may have an alternative splicing site, which gives rise to a 33 bp longer 5'UTR in the transcript (EU357905). In the coding region, there was one amino acid substitution (A60) from that predicted in the genomic sequence (V60). The predicted molecular weight of AeSCP-2L3 is 13.4 kDa with a calculated pI of 4.98. The C-terminus of AeSCP-2L3 does not have the peroxisome target sequence (AKL or SKL), which is the same characteristic of all *Aedes* SCP-2 and SCP-2-like proteins (Fig. 1). AeSCP-2L3 amino acid sequence identity to AeSCP-2, AeSCP-2L1, AeSCP-2L2, and the C-terminal SCP-2 domain in AeSCP-x is 29, 29, 21, and 38%, respectively (Fig. 1), although the amino acid sequence similarities are about 47–51% between AeSCP-2L3 and other *Aedes* single SCP-2 domain proteins. The AeSCP-2L3 gene has four exons which is unique among single SCP-2 domain protein genes in the *Ae. aegypti* SCP-2 family (Fig. 2). Interestingly, there is no predicted AeSCP-2L3 ortholog in the *Anopheles gambiae* genome, indicating the AeSCP-2L3 gene might be derived after the split of *Anopheles* and *Aedes* mosquitoes about 145–200 million years ago [18].

Expression profiles of AeSCP-2L3 gene

Total RNA samples from staged fourth instars were used to determine the temporal transcription profile of AeSCP-2L3. The levels of AeSCP-2L3 mRNA throughout the fourth stadium were stable with only a small, but significant increase at the 24 h time point and a slight decrease in pharate pupae at the 62 h time point (Fig. 3, red colored line). The temporal expression pattern of AeSCP-2L3 in the fourth instars contrasted sharply to other SCP-2-like genes that show an approximate 2 to 3-fold increase in transcription (Fig. 3, purple and green lines), suggesting that expression regulation of the AeSCP-2L3 gene is different from previously described mosquito SCP-2 and SCP-2-like genes [2,5]. In the adult stage, AeSCP-2L3 transcripts were at very low levels in both males and females through the first 4 days (Fig. 4a). Within 10 h post blood meal (PBM), the AeSCP-2L3 mRNA levels did not change (Fig. 4a), contrasting to the observations of increased expression of AeSCP-2 and AeSCP-2L2 within 10 h PBM [2,19]. The results imply that AeSCP-2L3 may not be functionally important in the initial phase of blood meal digestion/absorption during the vitellogenic cycle.

Tissues prominently transcribing the AeSCP-2L3 gene are in the body wall of larvae, pupae, and adults (Fig. 4b, lanes A/T or H/T or A). The overall levels of AeSCP-2L3 mRNA were higher in larvae than adults (Fig. 4b, fourth instar versus Day 1 adults). AeSCP-2L3 is larger (predicted molecular mass of 13.4 kDa) than AeSCP-2 (12.3 kDa), which was visible as the slightly slower migration of the AeSCP-2L3 band when blotted twice with antibodies against each protein (Fig. 4c, purple stained band versus brown stained bands in right panel). Anti-AeSCP-2L3 antibodies detected a protein of about 14 kDa in the body wall in larvae and adults (Fig. 4c, lanes 2, 3, and 6 in the right panel) but not in the ovaries either before or 5 h PBM (Fig. 4c, lanes 5 and 7 in the right panel). In contrast, Western blotting with anti-AeSCP-2 antibodies on the same blot showed that AeSCP-2 was more prominent in the larval midgut (Fig. 4c, lane 3 blotted in the left panel) as described previously [2]. The *Ae. aegypti* Aag-2 cell also showed a detectable level of AeSCP-2L3 expression (Fig. 4c, lane 8 blotted in the right panel), but AeSCP-2 levels were undetectable in Aag-2 cells (Fig. 4c, lane 8 in the left panel). As the body wall includes fat body, epidermis, CNS, and attached muscle cells, it is possible that the fat body is the likely origin of AeSCP-2L3 expression in the body wall due to the fact that 7 out of 8 AeSCP-2L3 ESTs are reportedly isolated from the female fat body (DV419323, DV361177, DV259641, DV259640, DV432324, DV432325, DV419324, and DV423800). AeSCP-2L3 is the first member of the mosquito single SCP-2 domain gene family that does not have higher expression levels in the gut tissue than in the fat body; all the other genes in the SCP-2 family are highly expressed in the larval gut [2,5]. The spatial expression patterns of AeSCP-2L3 suggest that the primary function of AeSCP-2L3 is as a lipid carrier in the fat body, which is divergent from other members of the mosquito SCP-2 family [2,4, 5].

Protein structure of SCP-2L3

The three-dimensional protein crystal structure of *Aedes aegypti* sterol carrier protein-2 like-3 (AeSCP-2L3) was solved by molecular replacement with a polyalanine model of AeSCP-2 and refined to 1.4 Å resolution (Table 1). It is the third SCP-2 family member solved from the yellow fever mosquito. The protein is coincidentally ligated with a single palmitic acid coordinated to the protein, which appears to have been advantageously acquired by the protein during the purification process. Other members of the SCP-2 structural family that have been solved with a natural ligand include AeSCP-2 (pdb entry 1PZ4) and the dimeric AeSCP-2 Like-2 which possesses three bound fats (pdb entry 2QZT). The overall backbone structure of AeSCP-2L3 is indeed similar to all the other members of the mosquito SCP-2 family, displaying an $\alpha + \beta$ tertiary classification (Fig. 5) with a five-stranded antiparallel β -sheet in the order 3-2-1-4-5 as defined by structural classification of proteins [20].

A large internal cavity is observed in AeSCP-2L3 with one C16 fatty acid (suppl. Table 1) ligated by coordination through the carboxylate groups of the fat near outer openings to the surface. The ligand was identified unequivocally as palmitic acid, based on omit map electron density contoured at 4σ (Fig. 6a). The cavity occupies space between β -sheets $\beta 1$, $\beta 2$, $\beta 3$, and α -helices $\alpha 1$, $\alpha 2$, and $\alpha 4$. The amino acid residues within 3–4 Å of the C16 fatty acid include: Lys15, Leu16, Val19, Arg24, Ser25, Phe26, Phe48, Phe117, and Lys120. It should be noted that Lys120 on $\alpha 4$ is the C-terminal residue that provides a direct hydrogen bond from the terminal oxygen to the $N\epsilon$ of Lys15 providing a salt bridge over the cavity (Fig. 6b). The residues lining the deep interior of the cavity that are within 5 Å of the fat include Leu12, Leu46, Leu51, Ile53, Phe82, Val85, Ala86, Phe92, and Leu113. It is interesting to note that the cavity appears as if it could accommodate a much larger substrate. Calculation of the cavity volume was done in the presence of computer generated hydrogen atoms with the exclusion of fat and waters. The AeSCP-2L3 cavity was determined to have a volume of 695 Å³, as calculated with Sybyl (www.tripos.com) using the Fast Connolly Channel method utilizing a probe radius of 1.4 Å [21]. For the purpose of comparison the internal cavity of AeSCP-2 was calculated in a likewise manner to be 606 Å³.

Binding of the carboxylate head group of the C16 fatty acid is coordinated by the loop between $\alpha 1$ and $\beta 1$. Carboxylate atom O1 forms shared hydrogen bonds to NE and NH2 of Arg24, the amide backbone of Ser25, and a well-defined water molecule (Fig. 6b). This water is coordinated through hydrogen bonds by not only O1 of the fat but also by the NE of Arg24 as well as the main chain oxygen's of Asp20 and Lys23. O2 of the fat carboxylate group is coordinated via hydrogen bonds to the amide backbone as well as the OG of Ser25. There is also an additional side chain coordination of O2 via a water mediated hydrogen bonding pattern from the NZ of Lys116 which resides on helix $\alpha 4$.

Another noticeable structural difference between AeSCP-2L3 and the two previously described mosquito SCP-2 proteins is the extra 10 or 11 amino acids between $\beta 3$ and $\beta 4$ (Figs. 1, 7), which in AeSCP-2L3 form a long extended loop on the other side the layer of stranded β -sheets (Fig. 7).

AeSCP-2L3 cryo-cooled and room temperature structures

When comparing two very different conditions for crystallographic data collection, such as cryo-cooled at -180°C in a nylon loop versus room temperature in a capillary, it has been observed that differences arise when constructing the three-dimensional models to fit the data [22]. For this reason we have added to the protein data bank, in addition to the 1.4 Å resolution cryo-cooled dataset of AeSCP-2L3 (pdb entry 3BKR), a room temperature dataset with both structure factors and model of AeSCP-2L3 refined to 2.1 Å (pdb entry 3BKS). Differences observed due to contraction of the bulk solvent within the lattice during the plunge of the crystal into liquid nitrogen include changes in the number of waters refined (168 for the cryo-cooled model versus 41 observed for the room temperature model) along with changes in the volume of the unit cell volume (232,991 Å³ for the cryo-cooled structure versus 237,651 Å³ for the room temperature structure). The decrease in unit cell volume is roughly 5,000 Å³ for the cryo-cooled structure. This represents a change of only 2%, which is consistent with the previously identified differences between cryo-cooled versus room temperature AeSCP-2L2 [6]. Likewise the waters observed in the room temperature structure appear to have nearly identical counterparts (> 95%) in the cryo-cooled model even though the crystallization conditions were significantly different.

Discussion

AeSCP-2L3 gene and gene expression

Isolation of AeSCP-2L3 cDNAs proves that AAEL012704 is an active SCP-2-like 3 protein gene, encoding the fifth member of a single SCP-2 domain protein in the mosquito SCP-2 family. Other members of the SCP-2 family such as AeSCP-x, AeSCP-2, AeSCP-2L1, and AeSCP-2L2 all have orthologs with over 75% amino acid sequence identities in the *Anopheles gambiae* genome [5]. However, a search in the *Anopheles gambiae* genome did not find any predicted protein with more than 26% amino acid sequence identity to AeSCP-2L3. Extensive search in the EST databank also failed to identify *Anopheles* transcript ortholog to AeSCP-2L3 gene. The lack of orthologous AeSCP-2L3 in the *Anopheles gambiae* genome suggests that AeSCP-2L3 gene evolved after the split of *Anopheles* and *Aedes* mosquitoes, which was about 145–200 million year ago [18]. The organization of the AeSCP-2L3 gene is quite different from that of other SCP-2-like genes (Fig. 2), indicating that the AeSCP-2L3 gene might not have been derived via a simple gene duplication event.

Another interesting difference between AeSCP-2L3 and other AeSCP-2-like genes is their spatial expression profiles. AeSCP-2L3 expression is higher in the body wall (Fig. 4b, c), presumably in the fat body, whereas other SCP-2-like genes are expressed highly in the gut tissue [2, 4, 5]. It is speculated that SCP-2 and SCP-2-like proteins aid the uptake of lipids from ingested food in the midgut, and the spatial expression patterns of AeSCP-x, AeSCP-2, AeSCP-2L1, and AeSCP-2L2 coincide with this proposed function [2, 4, 5]. Evidently, the primary function of AeSCP-2L3 may not be to enhance lipid uptake in the midgut. It has been shown that lipophorin endocytosis is not a relevant pathway to transfer or store lipids into the fat body [23, 24]. Therefore, it is speculated that intracellular lipid carriers such as sterol carrier protein and fatty acid binding protein in the fat body aid the transfer of lipids to lipid droplets. Whether AeSCP-2L3 transfers lipids for storage in the fat body needs further investigation. AeSCP-2L3 may not be involved in the absorption of lipids from the blood meal. In adults, AeSCP-2L3 expression is not induced by a blood meal, which differs from blood meal up-regulated expression of AeSCP-2 and AeSCP-2L2 [5, 19]. Both AeSCP-2 and AeSCP-2L2 are involved in lipid uptake from the blood meal [6]. However, AeSCP-2L3 may participate in the mobilization of lipids in the fat body since the higher level of AeSCP-2L3 expression is observed in the body wall, likely the fat body, throughout the life stages (Figs. 3, 4).

Structural comparison between mosquito SCP-2 proteins

A key point of interest between the members of the mosquito SCP-2 family is their extremely well-conserved polypeptide backbone fold, in spite of the fact that the sequence divergence between family members can be dramatic (Figs. 1, 5). To highlight this point, a comparison of the monomeric AeSCP-2L3 fold to the previously determined structure of AeSCP-2L2 (chain A) displays a root mean square difference (rmsd) of 1.6 Å over 79 amino acids, even though the identical residues shared between AeSCP-2L3 and AeSCP-2L2 number only 13 (16.5%) while the sequence homology is 27.4% (CCP4 calculation [25]). A comparison of AeSCP-2L3 with AeSCP-2 shows a rmsd of 1.4 Å over 99 amino acids with identical residues numbering only 24 (24.2%) while displaying a sequence homology of 88% (CCP4 calculation [26]). The highly homologous three-dimensional structures of the SCP-2 family members indicate that there is a structural constraint for the organization of the ligand cavities in the SCP-2 protein family. It is interesting that AeSCP-2L3 has as many as five phenylalanines lining the ligand cavity (Fig. 6b), which differs from the high numbers of methionines and leucines in AeSCP-2 and AeSCP-2L2, respectively (Fig. 1). The amino acid sequence divergence could give rise to varied ligand affinities; thus, the ligand specificity in SCP-2 proteins may be determined by the residues lining the hydrophobic cavities. Conservative substitutions of amino acids within

the ligand-binding site in plant SCP-2 proteins have shown that ligand specificity might be defined by the residues in the cavity but not the overall structural change of the protein [27].

In addition to the divergent nature of the AeSCP-2L3 sequence, notable differences between the mosquito AeSCP-2L3, AeSCP-2, and AeSCP-2L2 protein structures include the mode of fatty acid coordination (Fig. 6), the size and shape of the cavity (Fig. 5), and the extended loop between $\beta 3$ and $\beta 4$ in AeSCP-2L3 (Fig. 7). Similar to AeSCP-2, the carboxylate group of the fatty acid in AeSCP-2L3 is sequestered primarily by the loop that connects helix $\alpha 1$ to sheet $\beta 1$, which appears to be a conserved secondary structural element for ligand binding among mosquito SCP-2s (Figs. 1, 6). Intriguingly, each protein possesses an alternative mechanism of water mediated or direct interactions with the fatty acid carboxylate moiety (Fig. 6b) [6,7]. These results suggest that the divergent modes of ligand interaction could be accommodated by the highly conserved structure. One common thread between the protein structures is the amino acid Arg24 in AeSCP-2L3 that is conserved at an equivalent position in other *Ae. aegypti* SCP-2 proteins (Fig. 1) providing a direct or indirect hydrogen bonding partner to the O1 of the bound fatty acid albeit it is positioned as a different conformer within each protein. It is worth mentioning that Arginine has been observed as a critical residue for fatty acid binding in other fatty acid binding proteins even though the secondary structures are vastly different [22,27–29].

AeSCP-2L3 is the third protein structure of a single domain SCP-2 family member solved from the yellow fever mosquito *Ae. aegypti*. This undertaking is part of a larger effort to understand the structure and function of proteins that shuttle hydrophobic molecules within insect cells. Based on our observations, prediction of the binding specificity of a long chain lipid from the amino acid sequence would prove to be difficult, especially considering the utilization of water molecules indirectly acting to coordinate the ligand dissimilarly in each mosquito SCP-2 structure.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by the University of Wisconsin-Madison College of Agriculture and Life Sciences' USDA-CSREES Hatch project WIS04963, by grant W9113 M-05-1-0006 from the Deployed War Fighter Protection Research Program (DWFP) administered by the US Armed Forces Pest Management Board (AFPMB), by the National Institute of Health research grant #5R01AI067422 to Q.L. and by the Alex and Lillian Feir Graduate Fellowship to I.V.

References

1. Gallegos AM, et al. Gene structure, intracellular localization, and functional roles of sterol carrier protein-2. *Prog Lipid Res* 2001;40:498–563.10.1016/S0163-7827(01)00015-7 [PubMed: 11591437]
2. Krebs KC, Lan Q. Isolation and expression of a sterol carrier protein-2 gene from the yellow fever mosquito, *Aedes aegypti*. *Insect Mol Biol* 2003;12:51–60.10.1046/j.1365-2583.2003.00386.x [PubMed: 12542635]
3. Lan Q, Massey RJ. Subcellular localization of the mosquito sterol carrier protein-2 and sterol carrier protein-x. *J Lipid Res* 2004;45:1468–1474.10.1194/jlr.M400003-JLR200 [PubMed: 15145982]
4. Lan Q, Wessely V. Expression of a sterol carrier protein-x gene in the yellow fever mosquito, *Aedes aegypti*. *Insect Mol Biol* 2004;13:519–529.10.1111/j.0962-1075.2004.00510.x [PubMed: 15373808]
5. Vyazunova I, et al. Identification of two sterol carrier protein-2 like genes in the yellow fever mosquito, *Aedes aegypti*. *Insect Mol Biol* 2007;16:305–314.10.1111/j.1365-2583.2007.00729.x [PubMed: 17433070]

6. Dyer DH, et al. 3-D structure/function analysis of sterol carrier protein-2-like 2 reveals differences among SCP-2 family members. *J Lipid Res* 2007;M700460–JLR200.
7. Dyer DH, et al. The structural determination of an insect sterol carrier protein-2 with a ligand-bound C16 fatty acid at 1.35-Å resolution. *J Biol Chem* 2003;278:39085–39091.10.1074/jbc.M306214200 [PubMed: 12855689]
8. Riddiford LM, Curtis AT, Kiguch K. Culture of the epidermis of the tobacco hornworm, *Manduca sexta*. *Tissue Cult Assoc Man* 1979;5:975–985.10.1007/BF00919715
9. Ausubel, FM., et al. Current protocols in molecular biology. Greene Publishing Associates, Inc. and Wiley; New York: 1999.
10. McPherson, A. Preparation and analysis of protein crystals. Krieger; Malabar: 1982.
11. Page R, Stevens RC. Crystallization data mining in structural genomics: using positive and negative results to optimize protein crystallization screens. *Methods Macromol Cryst* 2004;34:373–389.
12. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 1997;276:307–326.
13. Storoni LC, McCoy AJ, Read RJ. Likelihood-enhanced fast rotation functions. *Acta Crystallogr* 2004;60:432–438.
14. McCoy AJ, et al. Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* 2005;61:458–464.10.1107/S0907444905001617 [PubMed: 15805601]
15. Zwart PH, Langer GG, Lamzin VS. Modelling bound ligands in protein crystal structures. *Acta Crystallogr D Biol Crystallogr* 2004;60:2230–2239.10.1107/S0907444904012995 [PubMed: 15572776]
16. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr* 1997;53:240–255.
17. Leszczynski M. The lattice constant of a nonperfect crystal measured by X-ray diffraction. *J Appl Cryst* 1993;26:280–283.10.1107/S0021889892011282
18. Krzywinski J, Grushko OG, Besansky NJ. Analysis of the complete mitochondrial DNA from *Anopheles funestus*: an improved dipteran mitochondrial genome annotation and a temporal dimension of mosquito evolution. *Mol Phylogenet Evol* 2006;39:417–423.10.1016/j.ympev.2006.01.006 [PubMed: 16473530]
19. Blitzer EJ, Vyazunova I, Lan Q. Functional analysis of AeSCP-2 using gene expression knockdown in the yellow fever mosquito, *Aedes aegypti*. *Insect Mol Biol* 2005;14:301–307.10.1111/j.1365–2583.2005.00560.x [PubMed: 15926899]
20. Murzin AG, et al. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* 1995;247:536–540. [PubMed: 7723011]
21. Connolly ML. Solvent-accessible surfaces of proteins and nucleic acids. *Science* 1983;221:709–713.10.1126/science.6879170 [PubMed: 6879170]
22. Juers DH, Matthews BW. Cryo-cooling in macromolecular crystallography: advantages, disadvantages and optimization. *Q Rev Biophys* 2005;37:105–119.10.1017/S0033583504004007 [PubMed: 15999418]
23. Canavoso LE, et al. Lipid transfer particle mediates the delivery of diacylglycerol from lipophorin to fat body in larval *Manduca sexta*. *J Lipid Res* 2004;45:456–465.10.1194/jlr.M300242-JLR200 [PubMed: 14679163]
24. Yun HK, Jouni ZE, Wells MA. Characterization of cholesterol transport from midgut to fat body in *Manduca sexta* larvae. *Insect Biochem Mol Biol* 2002;32:1151–1158.10.1016/S0965–1748(02)00051–6 [PubMed: 12213250]
25. Krissinel EB, et al. The new CCP4 coordinate library as a toolkit for the design of coordinate-related applications in protein crystallography. *Acta Crystallogr D Biol Crystallogr* 2004;60:2250–2255.10.1107/S0907444904027167 [PubMed: 15572778]
26. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr* 2004;D60:2256–2268.
27. Viitanen L, et al. Characterization of SCP-2 from *Euphorbia lagascae* reveals that a single Leu/Met exchange enhances sterol transfer activity. *FEBS J* 2006;273:5641–5655.10.1111/j.1742–4658.2006.05553.x [PubMed: 17212780]

28. Thompson JR, Banaszak LJ. Lipid–protein Interactions in lipovitellin. *Biochemistry* 2002;41:9398–9409.10.1021/bi025674w [PubMed: 12135361]
29. Haapalainen AM, et al. Crystal structure of the liganded SCP-2-like domain of human peroxisomal multifunctional enzyme type 2 at 1.75 Å resolution. *J Mol Biol* 2001;313:1127–1138.10.1006/jmbi.2001.5084 [PubMed: 11700068]
30. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–4680.10.1093/nar/22.22.4673 [PubMed: 7984417]

Fig. 1.

Amino acid sequence alignments of SCP-2 and SCP-2-like proteins from the yellow fever mosquito. This list includes: *Aedes aegypti* SCP-2L3, EU357904; SCP-2, AAO34708; SCP-2L1, ABF06628; SCP-2L2, ABF38969; SCP-x, AAQ24505. The alignment was prepared with the program CLUSTALW v1.8 [30]. The secondary structural elements correspond to the structure of yellow fever mosquito SCP-2 proteins where the residues in contact with the hydrophobic acyl chain are highlighted in *yellow* (fatty acid in vertical position) and *green* (fatty acid in horizontal position in AeSCP-2L2); those that coordinate the carboxyl moiety are depicted in *red* (fatty acid in vertical position) and *purple* (fatty acid in horizontal position in AeSCP-2L2); and those forming the dimer interface in AeSCP-2L2 are colored in *blue*.

Fig. 2. Transcript structures of *Aedes* SCP-2 gene family. Exons are shown as *boxes*; coding sequences are shown in *gray*. *Arrow* indicates the transcription direction. * indicates position of the start codon; x indicates the position of stop codon

Fig. 3.

Dot blot analysis of relative AeSCP-2L3 mRNA levels in fourth instars. *Red diamonds* AeSCP-2L3 gene, *blue triangles* AeSCP-2, *yellow circles* AeSCP-2L1, *green rectangle boxes* AeSCP-2L2. Bar = standard deviations ($N = 3$; each N represents the pool of total RNAs from 10 animals). Relative mRNA level = levels of each gene's mRNAs normalized to 0-h-old fourth instars that is arbitrarily set at 1

Fig. 4.

Expression profiles of AeSCP-2L3 gene in different life stages. **a** Semi-quantitative RT-PCR analysis of temporal patterns of transcription in adults. The muscle actin-2 gene (AY289764) is used as an internal control. **b** Semi-quantitative RT-PCR analysis of spatial profiles of transcription in different tissues. *H* head, *T* thorax, *A/T* abdomen and thorax, *G* gut. The ribosome protein L8 gene is used as an internal control. **c** Western blot analysis of AeSCP-2L3 in larvae, adults and cultured cells. *Lane 1* protein molecular weight markers, *Lane 2* the body wall from 12-h-old fourth instars, *Lane 3* the gut tissues from 12-h-old fourth instars, *Lane 4* whole body from 7-day-old male adults, *Lane 5* the ovaries from 7-day-old females, *Lane 6* the body wall from 4-day-old females 5 h post a blood meal, *Lane 7* the ovaries from 4-day-old females 5 h post a blood meal, *Lane 8* *Ae. aegypti* cell line Aag-2. Twenty micrograms of total soluble protein was loaded per lane

Fig. 5. Ribbon diagrams (Pymol) and surface/cavity representations of the mosquito SCP-2 family members. Refined structures include *Aedes aegypti* SCP-2L3 (pdb entry 3BKR) at 1.4 resolution, AeSCP-2 (pdb entry 1PZ4) at 1.35 resolution, and the dimeric AeSCP-2L2 (pdb entry 2QZT) at 1.7 resolution. The C16 fatty acids are highlighted as van der Waals space filled models and the cavities are calculated with the addition of hydrogens and without waters or fats present with the program Sybyl (www.tripos.com)

Fig. 6.

Stereo views (Pymol) of the C16 fatty acid along with selected residues. **a.** Coordination of C16 fatty acid emphasized with electron density from F_o to F_c omit map contoured at 4σ . The ribbon backbone and residues are colored according to B-factors which are highlighted in a gradient from yellow (highest B-factors) to blue (lowest B-factors). **b.** Coordination of C16 fatty acid and selected residues contributing to the formation of cavity. The phenylalanine residues are highlighted to show their abundance

Fig. 7. Backbone comparison of SCP-2 family monomers from *Aedes aegypti*. AeSCP-2L3 is shown as a cyan ball and stick $\text{C}\alpha$ backbone to differentiate it from AeSCP-2 (red) and the chain A monomer of AeSCP-2L2 (blue). Figure and overlays generated with Sybyl

Table 1

Data collection and refinement statistics

PDB identifier	3BKR	3BKS
Data collection temp. (C)	−180	+23
Resolution (Å)	25.0–1.4	25.0–2.1
High resolution shell (Å)	1.44–1.4	2.15–2.1
Total reflections	224,972	17,937
Unique reflections	30,445	9,607
R-merge (%) ^a	0.062 (0.300)	0.018 (0.063)
Completeness (%)	98.3 (96.9)	99.9 (99.0)
Average I/σ	23.3 (6.2)	38.3 (5.9)
Number of mol./asym. unit	1	
Number of fatty acid ligands	1	
Number of solvent atoms	168	41
R-factor (%) ^b	18.5 (21.4)	16.6 (16.0)
R-free (%)	21.8 (25.1)	18.9 (25.9)
Mean B value main chain (Å ²)	14.1	11.7
Mean B value fat (Å ²)	28.1	33.9
Mean B value water (Å ²)	28.4	30.4
ESU (maximum likelihood) Å	0.037	0.088
RMS deviation bond length (Å)	0.009	0.014
RMS deviation bond angles	1.18°	1.35°

^aR-factor = $(|F_O - F_C|/SF_O) \times 100$, where F_O is the observed structure-factor amplitude, and F_C is the calculated structure-factor amplitude

^bValues for the highest resolution shells are depicted in parentheses