Current Research in Structural Biology 3 (2021) 95–105

Contents lists available at ScienceDirect

Current Research in Structural Biology

journal homepage: www.journals.elsevier.com/current-research-in-structural-biology

The structures of two salivary proteins from the West Nile vector *Culex quinquefasciatus* reveal a beta-trefoil fold with putative sugar binding properties

Olivia Kern [a], Paola Carolina Valenzuela Leon [a], Apostolos G. Gittis [b], Brian Bonilla [a], Phillip Cruz [c], Andrezza Campos Chagas [a,2], Sundar Ganesan [d], Jose M.C. Ribeiro [a], David N. Garboczi [b], Ines Martin-Martin [a,**,1], Eric Calvo [a,*,1]

[a] *Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD, 20852, USA*
[b] *Structural Biology Section, Research Technologies Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, 20814, USA*
[c] *Bioinformatics and Computational Biosciences Branch. Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, 20892, USA*
[d] *Biological Imaging Section, Research Technologies Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, 20814, USA*

ARTICLE INFO

Keywords:
Arthropods
Mosquito
Salivary protein
Blood feeding
Vector borne diseases

ABSTRACT

Female mosquitoes require blood meals for egg development. The saliva of blood feeding arthropods contains biochemically active molecules, whose anti-hemostatic and anti-inflammatory properties facilitate blood feeding on vertebrate hosts. While transcriptomics has presented new opportunities to investigate the diversity of salivary proteins from hematophagous arthropods, many of these proteins remain functionally undescribed. Previous transcriptomic analysis of female salivary glands from *Culex quinquefasciatus*, an important vector of parasitic and viral infections, uncovered a 12-member family of putatively secreted proteins of unknown function, named the Cysteine and Tryptophan-Rich (CWRC) proteins. Here, we present advances in the characterization of two *C. quinquefasciatus* CWRC family members, CqDVP-2 and CqDVP-4, including their enrichment in female salivary glands, their specific localization within salivary gland tissues, evidence that these proteins are secreted into the saliva, and their native crystal structures, at 2.3 Å and 1.87 Å, respectively. The β-trefoil fold common to CqDVP-2 and CqDVP-4 is similar to carbohydrate-binding proteins, including the B subunit of the AB toxin, ricin, from the castor bean *Ricinus communis*. Further, we used a glycan array approach, which identifies carbohydrate ligands associated with inflammatory processes and signal transduction. Glycan array 300 testing identified 100 carbohydrate moieties with positive binding to CqDVP-2, and 77 glycans with positive binding to CqDVP-4. The glycan with the highest relative fluorescence intensities, which exhibited binding to both CqDVP-2 and CqDVP-4, was used for molecular docking experiments. We hypothesize that these proteins bind to carbohydrates on the surface of cells important to host immunology. Given that saliva is deposited into the skin during a mosquito bite, and acts as the vehicle for arbovirus inoculation, understanding the role of these proteins in pathogen transmission is of critical importance. This work presents the first solved crystal structures of *C. quinquefasciatus* salivary proteins with unknown function. These two molecules are the second and third structures reported from salivary proteins from *C. quinquefasciatus*, an important, yet understudied disease vector.

Abbreviations: SGs, Salivary glands; SGH, Salivary gland homogenate; CWRC, Cysteine and Tryptophan-Rich proteins from *Culex quinquefasciatus*; C, *Culex*.

* Corresponding author.
** Corresponding author.
E-mail addresses: ines.martin-martin@nih.gov (I. Martin-Martin), ecalvo@niaid.nih.gov (E. Calvo).

[1] These authors contributed equally to this work.

[2] Current address: Vector & Parasite Biology Department, Entomology Branch. Walter Reed Army Institute of Research, Silver Spring, Maryland, 20910, USA.

https://doi.org/10.1016/j.crstbi.2021.03.001
Received 21 December 2020; Received in revised form 1 March 2021; Accepted 4 March 2021

1. Introduction

Culex quinquefasciatus is a globally distributed mosquito vector of parasitic and viral pathogens, including Bancroftian filariasis and West Nile Virus (Farajollahi et al., 2011). These and other mosquito-borne pathogens are transmitted through the bite of an infected vector and injected into a vertebrate host's skin along with saliva. Mosquito saliva and its constituent proteins facilitate blood feeding through vasodilatory, anti-platelet, and anti-coagulant activities (Ribeiro, 1987). Further, mosquito saliva has immunomodulatory activity, and has been shown to promote pathogen infection and dissemination *in vivo* (Styer et al., 2011; Vogt et al., 2018; Pingen et al., 2016). In experimental settings, salivary gland homogenate (SGH) has been found to increase leukocyte recruitment to the dermis and to exacerbate arbovirus pathogenesis (Styer et al., 2011; Schneider and Higgs, 2008; Schmid et al., 2016). Given that mosquito saliva establishes a milieu within the dermis that promotes pathogen infection, designing vaccines to target mosquito salivary proteins offers a promising strategy for arbovirus control. Identifying optimal candidates for salivary-based vaccine design, however, will require a thorough understanding of the biochemical functions of major components of vector saliva.

While transcriptomic and proteomic analyses of salivary glands (SGs) have accelerated the discovery of new salivary protein families, many of these remain functionally undescribed (Ribeiro et al., 2010). It is estimated that among the 100–200 proteins contained within mosquito saliva, 30–40% belong to previously uncharacterized protein families and have unknown functions (Arca and Ribeiro, 2018). One such family, the cysteine and tryptophan-rich (CWRC) family, was first identified through a transcriptomic analysis of *Culex quinquefasciatus* female SGs (Ribeiro et al., 2004). The genome annotation of *C. quinquefasciatus* revealed additional members of this family, totaling 28 genes, 13 of which have EST representation (Arensburguer et al., 2010). A subsequent transcriptomic study of *Culex tarsalis* SGs uncovered orthologs of the CWRC proteins and found these proteins to be overexpressed in female salivary gland tissues compared to both female carcasses and male SGs (Ribeiro et al., 2018). While CWRC orthologs have since been identified in *Psorophora* mosquitoes and frog biting midges of the *Corethrella* genus (Chagas et al., 2013; Ribeiro et al., 2014), no orthologs have been found in either *Aedes* or *Anopheles* mosquitoes. Proteins belonging to the CWRC family are among the most abundant transcripts identified in *Culex* transcriptomes (Ribeiro et al., 2010; Calvo et al., 2010), yet no characterization of any CWRC protein has been published up until now.

Here, we describe two CWRC proteins from *C. quinquefasciatus*, CqDVP-2 and CqDVP-4. We show CqDVP-2 and CqDVP-4 transcripts are enriched in adult female salivary glands relative to other tissues, that both proteins are specifically localized within a region of the salivary glands consistent with proteins involved in blood feeding, and further, that these proteins are secreted into the saliva. In addition, we present the native crystal structures of both proteins, at 2.3 Å and 1.87 Å, respectively, and evidence of these proteins' carbohydrate binding activities.

2. Materials and methods

2.1. Ethics statement

All animal work was conducted in accordance with the Public Health Service Animal Welfare Assurance #A4149-01 according to the National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH) Animal Care and Use Committee (ACUC) guidelines, under the animal study protocol (ASP LMVR3).

2.2. Mosquito colony maintenance and salivary gland dissections

Culex quinquefasciatus mosquitoes were reared at 27 °C, 80% humidity, on a 12 h light/dark cycle at the Laboratory of Malaria and Vector Research insectary, NIAID, NIH. The colony was first established in 2015 from egg rafts collected in Hilo, Hawaii, US. Salivary glands from sugar-fed 5- to 8-day-old adult female mosquitoes were dissected in PBS pH 7.4 using a stereomicroscope (Zeiss, Thornwood, NY, USA). Salivary gland homogenate (SGH) was prepared by sonicating dissected salivary glands (Branson Sonifier 450) in PBS pH 7.4. Disrupted tissues were then centrifuged at 12,000×*g* for 5 min, and supernatants were recovered. The concentration of protein content was measured by spectrophotometry at A_{280} (DS-11, DeNovix, Wilmington, DE, USA), then stored at −80 °C until use.

2.3. Saliva collection

Saliva from starved female *Culex quinquefasciatus* mosquitoes was collected as described previously (Martin-Martin et al., 2020). Briefly, eight-day-old female *C. quinquefasciatus* mosquitoes were given water-soaked cotton and starved for 12 h. Mosquitoes were transferred to petri dishes using a mechanical aspirator and were sedated on ice for periods of 5–10 min. Mosquitoes were then transferred to platforms covered in double sided tape, with their dorsa and wings secured to the tape. Mosquitoes were injected with 200 nL of 10.8 mg/mL pilocarpine hydrochloride (Sigma P6503, St. Louis, MO, USA) intrathoracically and incubated at 27 °C for 1 h prior to saliva collection. Mosquito proboscises were then inserted into P10 XL pipette tips (Neptune 89,140–166, VWR International, Radnor, PA, USA) filled with sterile mineral oil. The hypopharynx was carefully separated from the outer sheath of the proboscis to encourage a higher yield of saliva. Mouthparts were left partially inserted, to avoid submersion of the maxillary palps. Mosquitoes were left to salivate for 1 h, and the resulting emulsion of saliva and mineral oil was collected into a Protein Lo-Bind tube (Eppendorf 0030108094, Enfield, CT, USA) containing 10 μL of sterile PBS. Saliva recovered from 61 mosquitoes was combined and centrifuged at 20,000×*g* for 5 min. The aqueous phase was recovered, and the concentration of the resulting protein solution was measured by spectrophotometry (DS-11, DeNovix). Saliva samples with protein concentrations of 1.93 mg/mL and 2.36 mg/mL were submitted to liquid chromatography coupled with mass spectrometry at the Research Technologies Branch (NIAID, NIH), as described previously (Martin-Martin et al., 2020). Briefly, samples of *C. quinquefasciatus* saliva were reduced using dithiothreitol (DTT), digested with trypsin, desalted, and dissolved in the injection solvent (0.1% formic acid, 3% acetonitrile). LC–MS analysis of digested peptides was performed on the Orbitrap Fusion mass spectrometer (ThermoFisher Scientific) connected with EASY nLC 1000 liquid chromatography system. Results were searched against the NCBInr proteome using PEAKS v10 (Bioinformatics Solutions Inc., Ontario, Canada). Peptides were filtered with a 0.5% false discovery rate using a decoy database approach and a two spectral matches/peptide requirement.

2.4. CqDVP-2 and CqDVP-4 gene expression analysis

Culex quinquefasciatus larvae, pupae and adult males were collected in Trizol (Life Technologies, Carlsbad, CA, USA) and stored at −80 °C. Additionally, female adults were dissected to separate the head and thorax from abdomens, and each tissue was stored alone in Trizol. Tissues were homogenized using pestles and total RNA was extracted with Trizol, according to manufacturer's instructions (Life Technologies). cDNA was prepared using the QuantiTect Reverse Transcriptase Kit (Qiagen), from 1 μg of extracted RNA, from two biological replicates of each tissue type. All nucleic acid concentrations and $OD_{260/280}$ ratios were measured by a DS-11 spectrophotometer (DeNovix). For qPCR, specific primers to target CqDVP-2 and CqDVP-4 genes were designed (CqDVP-2 –F: 5′-CTGCACTACCGGAAGCTCTA-3′, CqDVP-2-R: 5′-ACTCCGTATAAATTCTGGGCCT-3'; CqDVP-4-F: 5′-GGCAATCCAGGGTACTTTCA-3′, CqDVP-4-R: 5′- CACGACTCATAAGCCGCAAC-3′). Briefly, in a final volume of 20 μL, the reaction mix was prepared with 2X SsoAdvanced Universal SYBR Green Supermix (Bio-Rad), 300 nM of each primer, forward and reverse, and 100 ng of cDNA template. Reactions

were run on a CFX96 thermocycler (Bio-Rad) using the following amplification protocol: 95 °C, 5 min, 40 cycles of 95 °C, 10 min; 55 °C, 30 s and 72 °C, 20 s and a final step of 95 °C for 10 min. A melting curve was included to evaluate primer specificity. Two biological replicates of each tissue were analyzed, representing groups of 10 specimens each. All samples were analyzed in technical duplicates and non-template controls were included as negative controls. qPCR data were manually examined and analyzed using the ΔΔCt method. ΔCt values were obtained by normalizing the data against the *C. quinquefasciatus* 40 S ribosomal protein S7 transcript (GenBank AF272670; CxS7-F: 5′-GTGATCAAGTCCGGCGGTGC-3′ and CxS7-R: 5′-GCTTCAGGTCCGAGTTCATCTC-3′). Larvae samples were used as ΔΔCt controls. The relative abundance of genes of interest, or fold change, with respect to the larvae control, was calculated as $2^{-\Delta\Delta Ct}$. Graphs were prepared using GraphPad Prism software version 8.02.

2.5. Recombinant protein expression and purification

CqDVP-2 and CqDVP-4 cDNA sequences (AY388544, AY388554) were codon optimized for bacterial expression, cloned into a pET-17 b vector and transformed into *Escherichia coli* One Shot BL21 (DE3) pLysS competent cells, following the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). The bacterial cultures were grown in LB broth at 37 °C and induced with 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) for 3 h at 250 rpm. Inclusion bodies were recovered by centrifugation, washed and solubilized in 6 M guanidine hydrochloride. Soluble CqDVP-2 and CqDVP-4 were refolded dropwise in 4 L of 500 mM arginine, 50 mM Tris, 40 mM sucrose, 20 mM NaCl, 2 mM dithiothreitol (DTT), 2 mM ethylenediaminetetraacetic acid (EDTA), pH 8.6. Refolding solutions were then incubated at 4 °C with stirring overnight. Recombinant proteins were purified by size exclusion chromatography on a HiPrep16/60 Sephacryl S-100 H R column in PBS pH 7.4, followed by ion exchange chromatography, using a HiPrep SP FF 16/10 column in 25 mM MES, pH 6.0. A final round of analytical size exclusion chromatography was performed on a Superdex200 column in PBS pH 7.4. All HPLC was performed on a GE Äkta purifier 10, and columns were supplied by GE Healthcare Life Science (Piscataway, NJ, USA). Following purification, proteins were separated by electrophoresis on a 4–20% NuPAGE Tris-glycine polyacrylamide gel (Invitrogen) and stained with Coomassie blue using an eStain™ L1 Protein Staining System (GenScript Biotech, Piscataway, NJ, USA). To confirm protein identity, the resulting bands were analyzed by Edman degradation at the Research Technologies Branch, NIAID, NIH.

2.6. Production of polyclonal antibodies against recombinant CqDVP-2 and CqDVP-4

Six-week-old female BALB/c mice (5 per group) were inoculated with 50 μg of recombinant CqDVP-2 or CqDVP-4, in Magic Mouse adjuvant (Creative Diagnostics, Shirley, NY, USA). Animals received a total of three immunizations, one every 21 days, and terminal bleeds were performed four days following the final immunization. All injections and terminal bleeds were conducted by NIAID/NIH Animal Facility staff (NIAID/NIH ACUC). Levels of resulting IgG antibodies were quantified by ELISA, as described previously (Martin-Martin et al., 2018) and serum from the five animals immunized against each protein was pooled and stored at −30 °C. Previously prepared IgG from mice immunized against only Magic Mouse adjuvant was used as a control in salivary gland immunofluorescence assays.

2.7. Immunofluorescence assays

Dissected *Culex quinquefasciatus* salivary glands were transferred to shallow, concave wells in a 9-well plate (PYREX® Spot Test Plates, 85 × 100 mm; well dimensions: 22 mm outer diameter x 7 mm deep, Thomas Scientific, Swedesboro, NJ, USA) and fixed with 4% paraformaldehyde (Sigma) for 30 min at room temperature. Tissues were then washed 3 times for 10 min with PBS and blocked for 30 min with 2% bovine serum albumin (BSA, Sigma), 0.5% Triton X-100 (Sigma), PBS, pH 7.4. Glands were washed 3 times with PBS and incubated overnight at 4 °C with 200 μL of antibodies against either CqDVP-2 or CqDVP-4 diluted 1:500 in 2% BSA, PBS. Glands incubated in a 1:500 dilution of serum from mice immunized with Magic Mouse adjuvant only (in 2% BSA PBS) served as a negative control. Primary antibodies were removed by 3 washes of 10 min each with 2% BSA, PBS, and glands were incubated with 2 μg/mL anti-mouse Alexa Fluor 594 (Thermo Fisher, Waltham, MA, USA) for 2 h in the dark at 4 °C. Excess of conjugate antibody was removed by 3 additional washes with PBS. DNA was stained with 1 μg/mL DAPI (Sigma D9542) and actin by 4 U/mL Phalloidin Alexa 488 (Invitrogen) for 20 min. Glands were washed 3 times with PBS, then transferred to a glass slide, and mounted using a coverslip coated with 25 μL Prolong Gold mounting medium (Invitrogen). Slides were covered with foil and left to dry at room temperature overnight, then preserved at 4 °C until imaging. Fluorescent images were captured in a Leica SP8 Microscope using a 63x objective with 10X digital magnification. Images were processed with Imaris software version 9.2.1.

2.8. Western blots

Culex quinquefasciatus salivary gland homogenate (SGH) equivalent to 10 pairs of female salivary glands, and 150 ng of purified recombinant CqDVP-2 and CqDVP-4 were separated by NuPAGE. Proteins were transferred to a nitrocellulose membrane (iBlot, Invitrogen) then blocked for 1 h at 37 °C with blocking buffer (5% (w/v) powdered nonfat milk TBS). Polyclonal antibodies against CqDVP-2 and CqDVP-4 were diluted in blocking buffer (1:250) and incubated for 60 min. HRP-Goat anti-Mouse IgG (KPL) diluted in blocking buffer (1:3000) was used as a secondary antibody. Immunogenic bands were developed with Femto SuperSignal™ West Femto Maximum Sensitivity Substrate cat. 34,094 (Thermo Scientific, Rockford IL). Blots were imaged using an Azure 300 imaging system (Azure Biosystems, Dublin, CA, USA).

2.9. Crystallography, diffraction data collection and analysis

Purified recombinant CqDVP-2 and CqDVP-4 were crystallized using the hanging-drop diffusion method. Protein crystals grew after 14 days when CqDVP-2 (11 mg/mL) was incubated with 0.2 M Sodium acetate trihydrate, 0.1 M TRIS hydrochloride pH 8.5, 30% w/v Polyethylene glycol 4000 (Crystal Screen, Condition 22, Hampton Research, Aliso Viejo, CA, USA), in a 2.5:2 protein to buffer ratio (v/v), and after 4 days when CqDVP-4 (32.8 mg/mL) was incubated with 4 M Sodium formate (Crystal Screen, Condition 33, Hampton Research) in a 3:2 protein to buffer ratio. For data collection, the crystals were rapidly soaked in the mother liquor solution (the crystallization buffer described above) supplemented with 25% glycerol, then mounted on a loop and flash frozen in a nitrogen gas stream at 95 K. Data were collected on a Rigaku MicroMax-007 HF high-flux microfocus rotating anode X-ray generator equipped with a Rigaku SATURN A200 CCD detector using Cu radiation. For CqDVP-2, a crystal that diffracted to 2.3 Å resolution with cell dimensions (in Å) $a = 72.94$, $b = 73.05$ and $c = 106.96$ and belonging to the orthorhombic space group $P2_12_12_1$ was used to collect a data set (Table 1). For CqDVP-4, a crystal that diffracted to 2.4 Å resolution with cell dimensions (in Å) $a = 35.09$, $b = 59.59$ and $c = 85.310$ and belonging to the orthorhombic space group $P2_12_12_1$ was used to collect a preliminary data set. The data were processed, reduced and scaled with XDS (Kabsch, 2010). The structure of CqDVP-4 was determined by molecular replacement using Phaser (McCoy et al., 2007) by employing separate, manually constructed search models based on a sequence alignment of CqDVP-4 to proteins with known crystal structures performed with FFAS (Jaroszewski et al., 2005). For CqDVP-4, a second, higher resolution data set was collected from a crystal that was flash frozen in liquid nitrogen at Beamline 22BM (SER-CAT) of the Advanced Photon Source at the

Table 1
Data collection and refinement statistics.

	CqDVP-2	CqDVP-4
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	72.94, 73.05, 106.96	34.80, 59.68, 85.44
Resolution (Å)	46.49–2.30 (2.36–2.30)	34.74–1.87 (1.92–1.87)
I/σI	59.38 (8.92)	23.77 (3.48)
Completeness (%)	96.1 (71.8)	99.6 (96.0)
Redundancy	11.46 (7.96)	7.42 (4.85)
R-merge (%)[a]	3.1 (20.1)	3.8 (36.7)
Refinement		
Resolution (Å)	26.74–2.30 (2.36–2.30)	34.74–1.87 (1.96–1.87)
No. of reflections	24,943[b]	14,833
R_{work}/R_{free} (%)	20.46/25.47 (24.39/32.20)	22.96/25.43 (28.83/30.95)
No. of atoms		
Protein	4327	1120
Chain A	1068	
Chain B	1089	
Chain C	1069	
Chain D	1101	
Ligand (Additives)	12	9
Water	124[c]	30
B-factors (Å^2)		
Protein		
Chain A	33.08	
Chain B	30.94	
Chain C	33.94	
Chain D	34.27	54.82
Ligand (Additives)	50.82	58.39
Water	34.30	46.70
Root mean square deviations		
Bond lengths (Å)	0.011	0.009
Bond angles (°)	1.338	1.324

*Values in parentheses are for highest-resolution shell.
[a] R-merge(I) = $\sum_{hkl}(\sum_i|I_i(hkl) - <I(hkl)>|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the i-th observation of a reflection with indices (hkl), including those of its symmetry mates, and <I (hkl)> is the corresponding average intensity for all i measurements.
[b] There are 31 reflections in the 46.49–26.74 resolution range (Å).
[c] Three water molecules have alternate conformations.

Argonne National Laboratory on a MAR MX225 CCD detector. The crystal diffracted to 1.87 Å resolution and was isomorphous to the crystal used to obtain the data on the Rigaku SATURN A200 CCD detector, with cell dimensions (in Å) of $a = 34.80$, $b = 59.68$, and $c = 85.44$ and belonging to the same space group $P2_12_12_1$ (Table 1). The data were processed, reduced and scaled with XDS (Kabsch, 2010) as above. The model from the lower resolution data set was refined against the higher resolution data set using Phenix (Adams et al., 2010). The final model of CqDVP-4 was constructed by iterative manual tracing of the chain using the program Coot (Emsley et al., 2010) after each cycle of refinement with a stepwise increase in the resolution using Phenix (Adams et al., 2010). The final refinement statistics are presented in Table 1. Figures depicting the three-dimensional structures, including surface maps of electrostatic potential, were made with ChimeraX software (Goddard et al., 2018).

2.10. Glycan array and molecular docking

A glycan array was used to investigate the putative carbohydrate binding activities of CqDVP proteins in a high-throughput manner. The Glycan 300 array assesses protein binding to 300 of the most common carbohydrate structures with biological relevance from the scientific literature (https://www.raybiotech.com/glycan-array-300/); arrays to assess CqDVP-2 and CqDVP-4 carbohydrate binding were performed, and results were analyzed by RayBiotech (RayBiotech, Peachtree Corners, GA, USA). Binding was considered positive for all compounds with relative fluorescence intensities (RFI) > 100, and the 10 glycans with the highest RFI were included in the present analysis.

For molecular docking experiments, the three-dimensional structures of Gal-α-1,4-Gal-β-1,4-GlcNAc-β (G0074) and Globotetraosylceramide Gb_4 (G0026) were constructed in UCSF Chimera from components obtained from PubChem and were prepared for docking using the LigPrep procedure in Maestro (Schrödinger Release, 2020–3: Maestro, Schrödinger, LLC, New York, NY). The CqDVP protein structures were readied for docking by the Protein Preparation procedure in Maestro. Since the CqDVP-2 protein had four monomers in the crystallographic asymmetric unit, each monomer was prepared and used individually for docking experiments. Potential ligand binding sites were identified with the SiteMap procedure in Maestro. SiteMap found two potential binding sites in CqDVP-4 and either 2 or 3 binding sites in each monomer of CqDVP-2 protein. Docking of G0074 was performed to each of these sites using Glide version 6.4 (Schrödinger, LLC, New York, NY, 2015; release 2020–2). Due to the high number of rotational bonds in the ligand, Glide docking utilized expanded conformational sampling and considered a greater number of initial poses as well as post-docking poses for optimization. Post-docking minimization was performed using the 50 best scoring ligand poses (5 default), and the final number of poses reported was 20. Surface maps of the electrostatic potential of CqDVP-2 and CqDVP-4 were created using the Adaptive Poisson-Boltzmann Solver (APBS) and PDB2PQR online servers. CqDVP-2 and CqDVP-4 structures were aligned using the MatchMaker program in UCSF Chimera, and figures of the docked poses and electrostatic potential surfaces were also prepared in Chimera (Pettersen et al., 2004). Glide docking scores approximate binding energy, with more negative scores indicating greater stability. For the present analysis, scores ≤ -9 were considered, as they loosely correspond to nanomolar or greater binding.

3. Results

3.1. CqDVP orthologs are present in other Diptera

The CWRC family was first discovered through transcriptomic analysis of *Culex quinquefasciatus* salivary glands (SGs) (Ribeiro et al., 2004) and is named CWRC for its conserved cysteine (4) and tryptophan (5) residues (Fig. 1a). Within *C. quinquefasciatus,* whose genome is known, it has been observed that nearly all members of this protein family are coded by single exon genes (Arensburger et al., 2010). PSI-BLAST of CWRC proteins recovered bacterial sequences containing the trefoil domain from the ricin toxin protein. Subsequent transcriptomes have uncovered many orthologs in *Culex tarsalis* and *Corethrella appendiculata* and a single transcript in *Psorophora albipes* (Ribeiro et al., 2014, 2018; Chagas et al., 2013) (Fig. 1b). Alignment of 62 protein sequences from *Culex* and *Corethrella* shows only 1.0% identities (two C and one W), and 14.8% similarities (Fig. S1). Phylogenetic analysis (Fig. S2) shows a phylogram with many clades with poor bootstrap support at their roots. However, eleven branches have more than 90% bootstrap support, three of which have *Corethrella* and *Culex* sequences, indicating conservation of these orthologs over 200 million years. The remaining clades contain both *C. quinquefasciatus* and *C. tarsalis* sequences. We have focused our study on two CWRC family members, CqDVP-2, and CqDVP-4 (GenBank AY388544, AY388554). These proteins received their name due to the conserved amino acid sequence of aspartic acid (D), valine (V) and proline (P) following cleavage of their signal peptides (Fig. 1a).

3.2. CqDVP proteins are specific to adult female salivary glands and are secreted into vector saliva

As the *Culex tarsalis* salivary transcriptome showed CWRC transcripts to be enriched in female SGs relative to male SGs and both female and male whole body samples (Ribeiro et al., 2018; Calvo et al., 2010), we sought to compare CqDVP-2 and CqDVP-4 transcripts across tissue types and developmental stages by RT-qPCR of *C. quinquefasciatus* tissues. The results validated that as in *C. tarsalis,* CqDVP transcripts were most

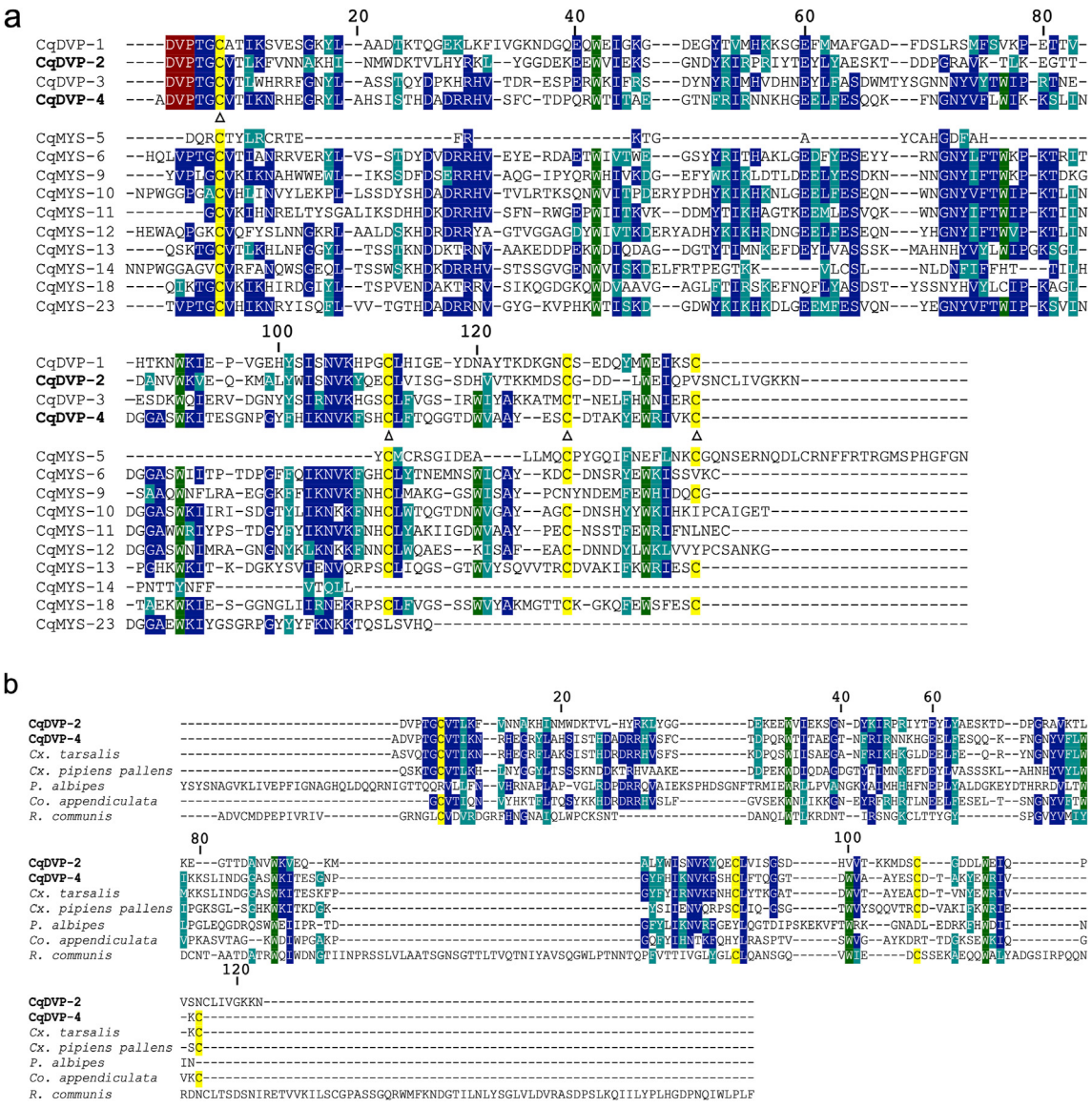

Fig. 1. Multiple sequence alignment of mature *C. quinquefasciatus* Cysteine and Tryptophan-rich (CWRC) protein family members and orthologs.
(a) Alignment of mature amino acid sequences of Cysteine and Tryptophan-rich (CWRC) protein family members (GenBank AAR18412, AAR18432, AAR18455, AAR18442, AY388551, AY388569, AY388571, AY388573, AY388546, AY388523, AY388574, AY388545, AY388531, AY388522) from *C. quinquefasciatus*. Highlighted in maroon is the conserved sequence DVP, which corresponds to four family members' starting sequence following cleavage of signal peptides. Cysteines involved in disulfide bonds are marked by triangles.
(b) Mature CqDVP-2 and CqDVP-4 amino acid sequences were aligned with CWRC orthologs from *Culex tarsalis*, *Culex pipiens pallens*, *Psorophora albipes*, *Corethrella appendiculata*, and the B-chain of the ricin toxin, *Ricinus communis* (GenBank JAV24644, ABG75915, JAA93994, JAB54960, AFH96941). Alignments were generated using Clustal Omega and the Boxshade server. Identical amino acids are shaded in dark blue, similar residues in teal, conserved tryptophan residues are colored green, and cysteines in yellow. Amino acid number corresponds to the mature CqDVP-4 sequence.

enriched in adult female head and thorax tissues, where the SGs are localized (Fig. 2). Further, Western blots showed that antibodies against CqDVP-4 recognized this protein within SGH (Fig. 2). While CqDVP-2 was not detected within SGH by Western blot, mass spectrometry analysis identified peptides from *C. quinquefasciatus* saliva that mapped to CqDVP-2 and CqDVP-4 amino acid sequences (Fig. 2, Supplementary Table 1). In addition, no cross-reactivity was found between CqDVP antibodies and the two proteins (Fig. S3); in immunofluorescence assays, both CqDVP-2 and CqDVP-4 were specifically localized within the distal lateral lobes of salivary gland tissues (Fig. 3), a finding consistent with the expression patterns of secreted proteins implicated in blood feeding from female *Aedes* and *Anopheles* mosquito SGs (Juhn et al., 2011; Arcá et al., 1999). To investigate the biological properties of these salivary proteins, mature CqDVP-2 and CqDVP-4 cDNA sequences were codon optimized for expression in *E. coli* and purified using a combination of size exclusion and ion exchange chromatography; purified recombinant proteins migrated as single bands on a Coomassie blue-stained polyacrylamide NuPAGE gel (Fig. S4). The bands' approximate size (MW) corresponded to the MW of CqDVP-2 (16.5 kDa) and CqDVP-4 (16.3 kDa), respectively (Fig. S4). The identities of the purified recombinant proteins were confirmed by N-terminal sequencing at the Research Technologies Branch, NIAID, NIH.

3.3. *The crystal structures of CqDVP proteins have a β-trefoil fold*

To study the biochemical and functional characteristics of the two CqDVP proteins, we solved the native crystal structures of both CqDVP-2 and CqDVP-4 (Fig. 4). The structures of CqDVP-2 and CqDVP-4 were

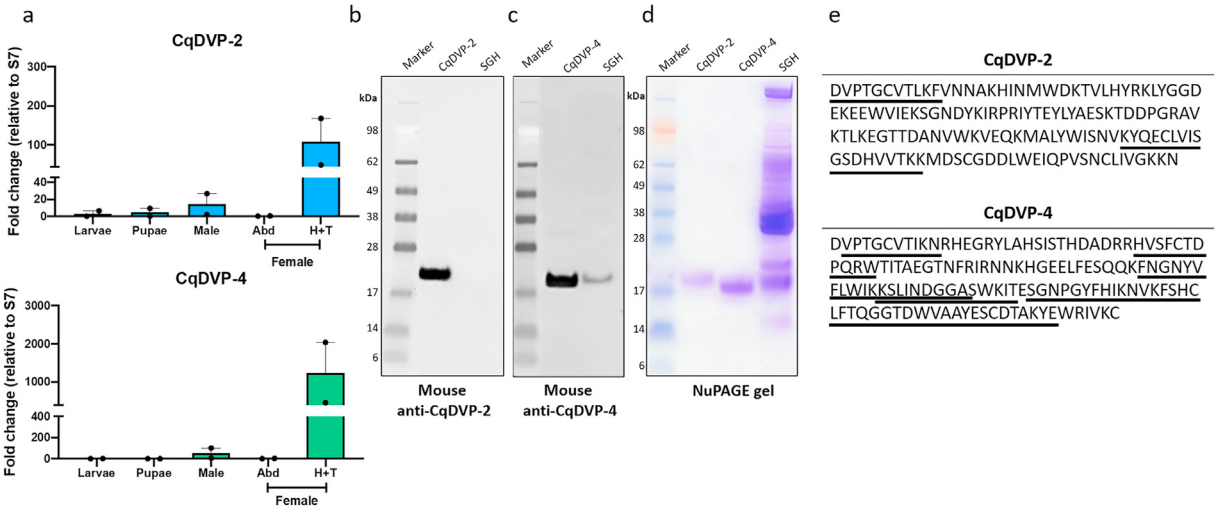

Fig. 2. CqDVP-2 and CqDVP-4 are enriched in adult female SGs and present in saliva
(a) Comparison of CqDVP-2 and CqDVP-4 transcript levels in *Culex quinquefasciatus* mosquitoes across developmental stages, tissues, and sex. Relative abundance was expressed as the fold change using the 40 S ribosomal protein S7 from *C. quinquefasciatus* as the housekeeping gene, Heads and thoraxes (H+T) containing SGs were dissected from abdomens (abd). (b) Western blots of CqDVP-2 and (c) CqDVP-4 demonstrate specific antibody recognition of these proteins and confirm the presence of CqDVP-4 in SGH. (d) Coomassie blue stained Nu-PAGE gel of recombinant proteins (2.8 μg) and the contents of five *C. quinquefasciatus* salivary glands. (e) Peptides identified by mass spectrometry of *C. quinquefasciatus* saliva corresponding to CqDVP-2 and CqDVP-4 are underlined, confirming the secretion of both CqDVP proteins into mosquito saliva. For full peptide list, see Supplementary Table 1.

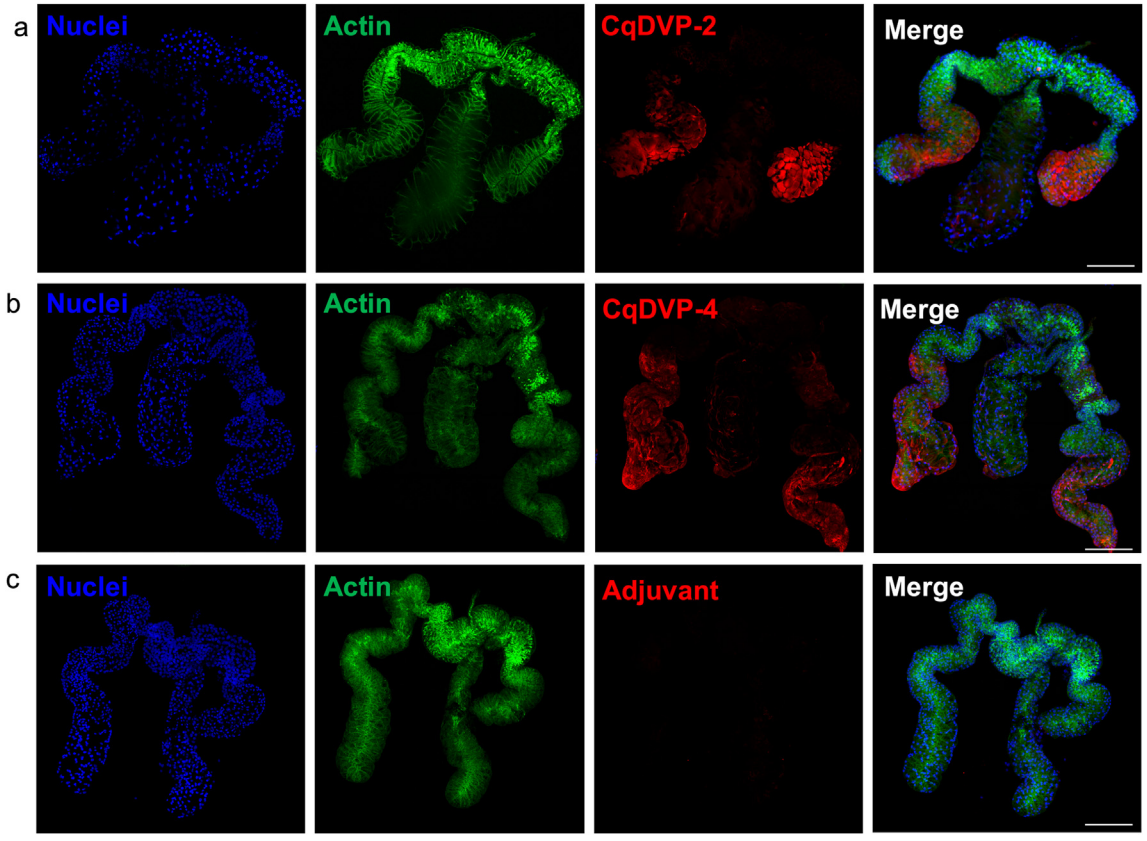

Fig. 3. CqDVP-2 and CqDVP-4 are localized in the distal lateral lobes of *C. quinquefasciatus* salivary glands Immunofluorescence of female *C. quinquefasciatus* SGs indicates specific expression of CqDVP proteins in the distal lateral lobes of these tissues. SGs were incubated with (a) anti-CqDVP-2 or (b) anti-CqDVP-4 antibodies raised in mice. (c) Glands incubated with sera from mice immunized against only Magic Mouse adjuvant were used as a control. Anti-mouse IgG Alexa Fluor 594 was used as a secondary antibody, shown in red. Actin, stained with Phalloidin 488, is shown in green, and nuclei, stained with DAPI, are in blue. Images were captured on a Leica SP8 confocal microscope at 63x. Scale bars = 100 μm.

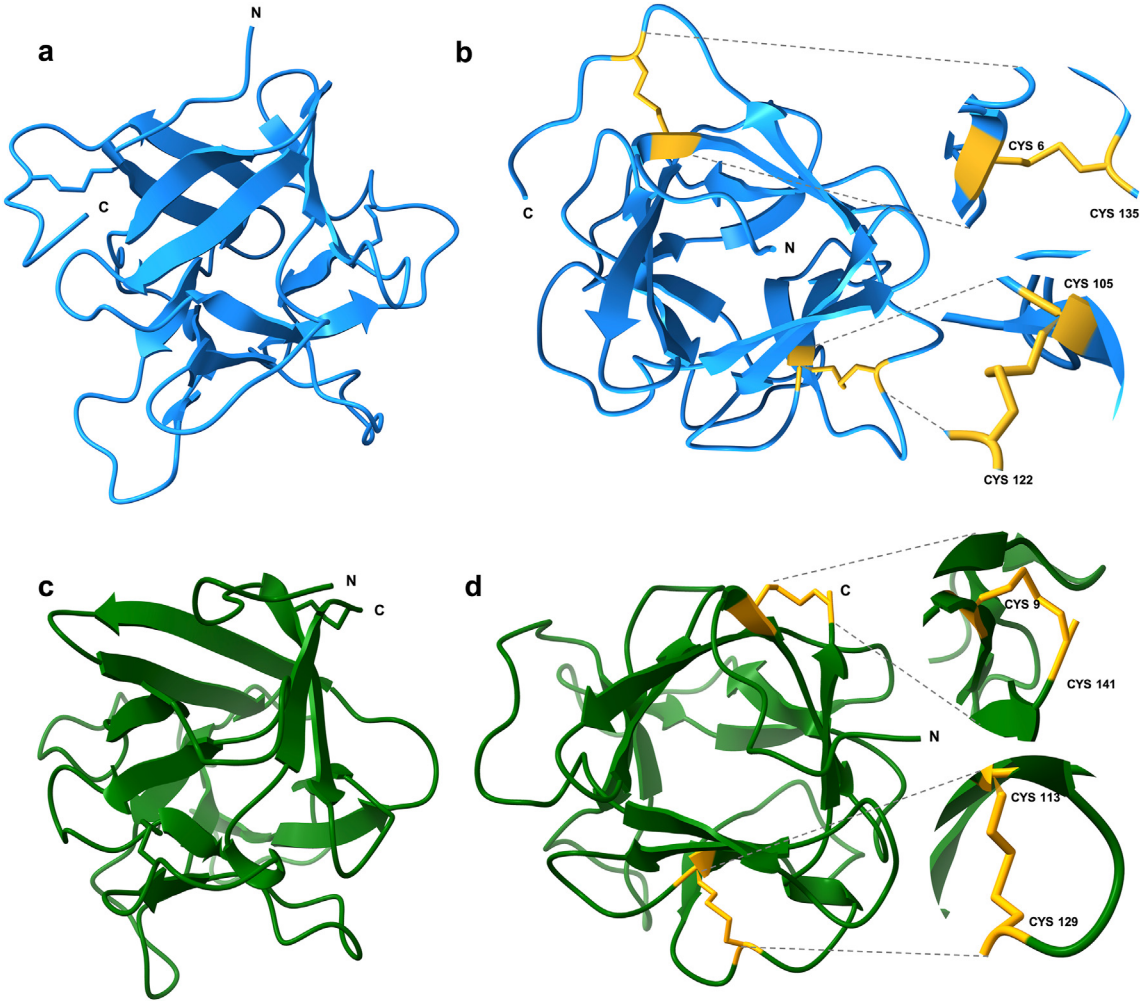

Fig. 4. The crystal structures of CqDVP-2 and CqDVP-4.
(a) Ribbon representation of CqDVP-2 β-trefoil fold structure at 2.3 Å (chain B). (b) Disulfide bonds linking cysteine residues C6 and C135, and C105 and C122 are shown in yellow. (c) The crystal structure of CqDVP-4 structure at 1.87 Å. The protein's β-trefoil fold (d) is stabilized by disulfide bonds linking cysteine residues C9 and C141, and C113 and C129, in yellow. N-terminal and C-terminal sequences are represented as N and C, respectively. Figures were produced using ChimeraX.

determined by molecular replacement; both proteins have a $P2_12_12_1$ orthorhombic space group and data were collected from crystals that diffracted to 2.3 Å, and 1.87 Å, respectively (Table 1). The coordinates of both proteins have been deposited in the Protein Data Bank with PDB IDs

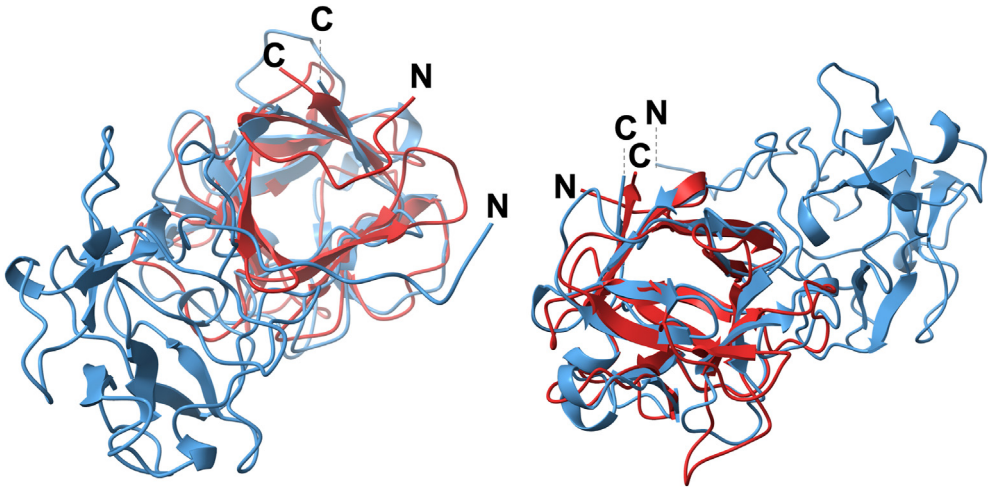

Fig. 5. The ß-trefoil fold of CqDVP-4 is similar to the ricin toxin B-lectin
Overlap of CqDVP-4 (red) and the Ricin B-lectin subunit (blue) of the ricin toxin from *Ricinus communis* (PDB ID 2AAI) highlights the proteins' similar structure. Structures were aligned based on their amino acid sequence in ChimeraX, and N and C terminal ends are labeled.

7KC8—CqDVP-2 and 7KCG—CqDVP-4. The structure of both CqDVP-2 and CqDVP-4 consist of 12 helices stabilized by 2 disulfide bonds. For CqDVP-2, disulfides link C6 and C135, C105 and C122, and for CqDVP-4, disulfide bonds link cysteines at positions C9 and C141, and C113 and C129. In addition, CqDVP-4 also has two free cysteine residues, C38 and C93. As PSI-BLAST of CWRC proteins recovered bacterial sequences containing the trefoil domain from the ricin toxin protein, we compared the structure of CqDVP-4 protein to ricin. The β-trefoil fold structure of CqDVP-4 has a pseudo three-fold axis, highly similar to that of the B-lectin from the ricin toxin (Figs. 4 and 5). The electrostatic potential map highlights the surface of CqDVP-2 has strongly negatively charged regions, whereas the surface of CqDVP-4 has a greater overall strongly positive charge (Fig. S5).

3.4. CqDVP proteins bind to terminally linked galactose and galactose derivatives

Glycan array 300 testing identified 100 carbohydrate moieties with positive binding to CqDVP-2, and 77 glycans with positive binding to CqDVP-4; a full list of the 300 glycans tested is provided (Supplementary Table 2). For all compounds with relative fluorescence intensities (RFI) > 100, binding was considered positive. The 10 glycans with the highest RFI were graphed (Fig. 6); the glycans with the highest RFI, Gal-α-1,4-Gal-β-1,4-GlcNAc-β (G0074) and Globotetraosylceramide Gb_4 (G0026), both exhibited binding to CqDVP-2 and CqDVP-4 and were used for molecular docking experiments (Supplementary Table 3).

Glide docking scores loosely approximate binding energy, with more negative scores indicating greater stability. The best-scoring docked poses of G0074 gave excellent scores with both CqDVP-2 (chain C, −11.8) and CqDVP-4 (−12.2) (Supplementary Table 3). The conformation of G0074 which returned the docking score of −11.8 to a putative binding site on the surface of CqDVP-2 is shown, as is another binding site on CqDVP-4, in which G0074 achieved a binding score of −10.2 (Fig. 6b).

While the binding sites of CqDVP-2 and CqDVP-4 in which the G0074 ligand returned the best docking scores were localized within the same

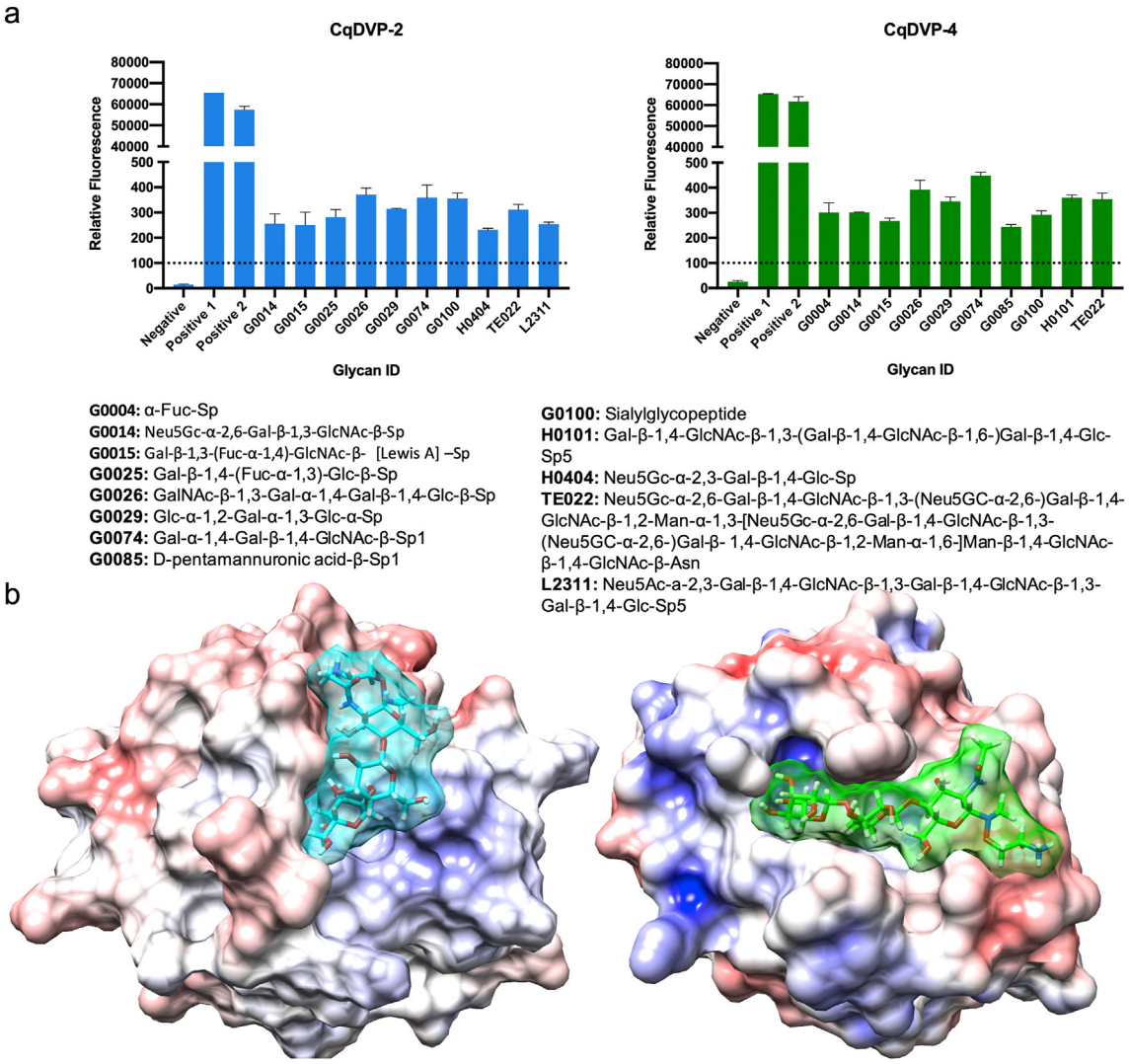

Fig. 6. CqDVP-2 and CqDVP-4 bind to carbohydrates with terminally linked galactose and galactose derivatives
(a) Glycan arrays of purified recombinant CqDVP-2 and CqDVP-4 against 300 glycans identified >10 positive hits with relative fluorescence intensities (RFI) > 100. The 10 glycans with the greatest RFI are shown above, with the full name of each glycan in the list below. Among the top 10 scoring glycans for each protein, seven of these returned high relative scores for both CqDVP-2 and CqDVP-4. These data indicate both CqDVP proteins showed positive binding to glycans with terminally linked galactose or N-acetylgalactosamine residues, notably G0074 and G0026. The best docking scores for these ligands are shown in (Supplementary Table 3). One negative and two positive controls were included in each array. (b) Surface maps of CqDVP-2 (left) and CqDVP-4 (right) with compound G0074 shown in a binding pose with a glide score of −11.8 and −10.2, respectively. The surface of both proteins is colored according to electrostatic potential, with blue representing strongly electropositive, and red, strongly electronegative surfaces, and ligand G0074 is colored in cyan (left) and green (right). The model was generated in Chimera.

general vicinity for both proteins, structural alignment revealed the position of the ligand when bound to CqDVP-2 was nearly orthogonal to its pose when bound to CqDVP-4 (Fig. 6). The differences in docked poses may be explained by local structural differences, given the differences in the electrostatic potential of the binding sites between CqDVP-2 and CqDVP-4, despite the overall similar morphologies of these proteins' backbone conformations. The four monomers identified in the unit cell of the crystal structure of CqDVP-2 yielded different Glide scores at the protein's main binding site (chain A: −8.61, chain B: −10.16, chain C: −11.83, chain D: −9.7), indicating the heterogeneity in the conformations of each monomer, and highlighting the value in performing docking experiments to each monomer separately.

4. Discussion

The salivary glands (SGs) of female mosquitoes can contain between 100 and 200 different types of salivary proteins and peptides (Arca and Ribeiro, 2018). Salivary molecules have vasodilatory, anti-platelet, and anti-coagulant activities that facilitate blood feeding (Ribeiro, 1987). In addition, saliva-induced changes at the bite site may favor pathogen transmission and increase viral dissemination by inducing leukocyte recruitment (Schneider et al., 2010). While the transcriptomes of major arthropod vector SGs have been published, between 30 and 40% of the resulting sequences return no matches in BLAST searches, representing a wealth of undescribed salivary proteins (Arca and Ribeiro, 2018).

Here, we have expanded the description of two Cysteine and Tryptophan-rich (CWRC) proteins, CqDVP-2, and CqDVP-4, previously identified by a transcriptomic analysis of *Culex quinquefasciatus* female SGs (Ribeiro et al., 2004). We have confirmed by RT-qPCR that CqDVP transcripts are enriched in female salivary glands, as are CWRC orthologs from *C. tarsalis* (Calvo et al., 2010), indicating a possible role in blood feeding and pathogen transmission, as only females bite and transmit disease. Further, we found CqDVP-2 and CqDVP-4 are specifically localized within the distal lateral lobes of SGs, a finding consistent with secreted proteins known to be important for blood feeding and pathogen transmission, including the *Culex* D7 long form proteins and *Aegyptin* (Martin-Martin et al., 2020; Juhn et al., 2011; Chagas et al., 2014). Although distinct functional compartmentalization of proteins within SG tissues has not yet been confirmed, proteins involved in nectar digestion have been found to be specifically localized within the proximal lateral lobes of SGs, and those implicated in blood feeding within the medial or distal lateral lobes (Juhn et al., 2011). Beyond CqDVP-2 and CqDVP-4 localization within a region of SGs consistent with blood feeding secreted proteins (Fig. 3), evidence from Western blots and mass spectrometry (Fig. 2) confirms that both CqDVP proteins are found in mosquito saliva, providing evidence for their presence in the human host's skin during a bite. In addition, Western blots confirmed that CqDVP-4 is present within salivary gland homogenate (SGH). CqDVP-2 was not recognized within SGH by antibodies likely because it is much less abundant compared to CqDVP-4, as *in silico* antigenicity prediction identified no differences in antigenicity potential across the two proteins (Fig. S7). Another possibility is CqDVP-2 epitopes may be lost by protein denaturation in SDS-PAGE, otherwise a low signal would also be found in histochemistry assays. Both proteins were, however, identified within the saliva by mass spectrometry (Supplementary Table 1).

Our crystallographic studies have shown that both CqDVP-2 and CqDVP-4 structures have a β-trefoil fold, a structural motif shared with the B-chain of the Ricin toxin, the portion of the toxin complex that binds to carbohydrates within the plasma membrane of host cells, including galactose (Rutenber et al., 1987). Although ricin is known to be a dimer (Rutenber et al., 1987), CqDVP proteins appear to be monomers, as observed by the single peak in the size exclusion chromatography (Fig. S4) and their electrophoretic behavior under non-reducing conditions (Ribeiro et al., 2004). CqDVP-4 in particular shares a high structural similarity with the ricin B-chain (Fig. 5). The β-trefoil fold is conserved among carbohydrate-recognition domains of plant and bacterial AB-toxins, and it is thought to have resulted from a gene triplication of a galactoside-binding peptide of bacterial origin (Rutenber et al., 1987). Given that the CWRC family is encoded by uniexonic genes, and PSI-BLAST searches exclusively return other blood feeding arthropod salivary proteins and proteins of bacterial origin, CWRC members may have originated in the horizontal transfer of carbohydrate-recognition domains from prokaryotes to a dipteran ancestor (Ribeiro et al., 2010). These newly acquired genes may have evolved beyond recognition from their common ancestor following a series of gene duplication events (Ribeiro et al., 2010; Calvo et al., 2010). Evidence of gene duplication is common across salivary genes, and it is thought to have greatly contributed to both the specificity of salivary protein function and the evolution of novel functions (Arca and Ribeiro, 2018). Horizontal gene transfer (HGT), too, has provided new tools for blood feeding arthropods; *Ornithodoros* ticks have acquired a vasodilatory hormone of vertebrate origin by HGT (Iwanaga et al., 2014), and genes originating from *Wolbachia* have been identified among salivary genes in both *Anopheles* and *Aedes* mosquitoes (Arcà et al., 2005; Korochkina et al., 2006). Horizontal transfer of a prokaryotic carbohydrate recognition domain to a common dipteran ancestor approximately 250 MYA could explain the divergence of the CWRC proteins across Culicomorpha, including frog-biting midges, *Corethrella*, and *Psorophora* mosquitoes (Fig. 1b, Fig. S2).

Phylogenetic analysis and the species occurrence of the CWRC proteins suggest that they were present in the Culicomorpha ancestor of mosquitoes and frog biting midges over 200 MYA (Fig. S1 and S2). Sialotranscriptomes of *Culex quinquefasciatus* (Ribeiro et al., 2004) and *C. tarsalis* (Ribeiro et al., 2018; Calvo et al., 2010) found more than twenty members of this family in each mosquito species. A deep sequencing sialotranscriptome of *Psorophora albipes* found a single contig belonging to this family (Chagas et al., 2013), while deep sequencing of *Aedes aegypti* failed to identify members of this family (Ribeiro et al., 2016). The single sequence found in *Psorophora* indicates it is also present among the Aedini, but in reduced numbers (Chagas et al., 2013). No sequences of this protein family have been found in salivary transcriptomes or in the genomes of Anophelines (Arcà et al., 2005). Since both *C. quinquefasciatus* and *C. tarsalis* mosquitoes feed mostly on birds (Reisen et al., 2005; Tempelis, 1975), while *Aedes aegypti* and *Anopheles* are predominantly mammal feeders (Garrett-Jones et al., 1980; Powell and Tabachnick, 2013) and *Corethrella* feeds on Amphibia (Borkent, 2008), it is possible that the genes coding for the CWRC family were acquired for its adaptive value to the ancestral Culicomorpha feeding on non-mammal vertebrates, most probably dinosaurs, which would have served as the dominant food source during this period. It appears that the family was lost on those Culicomorpha that speciated with the mammalian fauna after the extinction of the dinosaurs in the end of the Cretaceous, ~65 MYA (Axelrod and Bailey, 1968). It is thus possible that the function of this protein family may be related to some particular aspect of non-mammalian hemostasis that differs from its mammalian counterpart. Thus, the effects of these proteins on avian blood clotting, thrombocyte aggregation and vasodilation should be investigated. Within this context it is interesting to note that a salivary component from a bird feeding bat was found to activate avian blood fibrinolysis, a finding which allowed for the discovery of the avian plasminogen activation pathway (Cartwright and Hawkey, 1969).

While the physiological functions of the CqDVP proteins remain unknown, we have presented evidence both CqDVP proteins bind to carbohydrates, as their structures are characteristic of carbohydrate-binding proteins, and data from glycan arrays identified positive binding to Globotetraosylceramide Gb_4 (G0026), a globoside thought to play a role in inflammatory processes, as well as Gal-α-1,4-Gal-β-1,4-GlcNAc-β (G0074), a glycosphingolipid P^k antigen belonging to the P1PK blood group system (Zhang et al., 2019). Molecular docking experiments have identified putative binding pockets for the relevant carbohydrate ligands (Fig. 6, Fig. S6). In addition, we have confirmed the recombinant proteins are immunogenic, at least in mice, as IgG antibodies were present in high relative levels following mice immunizations (Fig. S3). These findings

suggest CqDVP proteins might play a role in modulating host immune response to mosquito bites, likely to facilitate successful blood meal acquisition. Through interactions with carbohydrates present in the plasma membranes of mast cells or other immune cells, these proteins may limit host inflammation while a mosquito feeds, reducing itch and sensation, and, as a consequence, increasing the odds of survival for the vulnerable mosquito.

Further investigation into these proteins' functions should focus on their interactions with immune cells, specifically mast cells, macrophages or neutrophils, as these cells are among the first responders to mosquito saliva following a bite. It is hypothesized that *C. quinquefasciatus* saliva contains proteins that act to antagonize mast cell degranulation and host pain receptors (Ribeiro et al., 2001). Given the abundance of both globosides and glycosphingolipids in mast cell membranes, and our evidence of CqDVPs binding to these glycans, these proteins should be studied as potential antagonists of mast cell degranulation. Similarly, CqDVP proteins may facilitate viral infection of macrophages, which are among the first cells to be exposed to saliva during a bite, contain Gb_4 globosides in their membranes, and play an important role in inflammation (Zhang et al., 2019). In addition, given *Culex* preference for blood feeding on birds, future work may investigate the effects of these proteins on avian heterophils, a type of granulocyte with important antimicrobial activity in the avian innate immune system (Harmon, 1998).

Here, we have expanded the molecular characterization of two novel proteins from the saliva of female *Culex quinquefasciatus* mosquitoes, including the first and second crystal structures of CWRC proteins, and the second and third structures of proteins from *C. quinquefasciatus* salivary glands (Martin-Martin et al., 2020). Further, we have confirmed that both CqDVP-2 and CqDVP-4 are present and relatively abundant within secreted saliva. Given that saliva is deposited into the skin during a mosquito bite, and acts as the vehicle for arbovirus inoculation, understanding what role, if any, CqDVP proteins serve in pathogen establishment is of critical importance. Our preliminary characterization of CqDVP-2 and CqDVP-4, two proteins unique to *Culex* mosquitoes, presents new insights into the salivary toolkit of *C. quinquefasciatus* mosquitoes, a lesser studied disease vector. While *C. quinquefasciatus* and other members of the *Culex* genus are globally important vectors of human and veterinary disease, these mosquitoes remain understudied relative to species belonging to *Aedes* and *Anopheles* genuses.

CRediT authorship contribution statement

Olivia Kern: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, preparation. **Paola Carolina Valenzuela Leon:** Investigation. **Apostolos G. Gittis:** Methodology, Formal analysis, Investigation. **Brian Bonilla:** Investigation. **Phillip Cruz:** Methodology, Formal analysis. **Andrezza Campos Chagas:** Investigation. **Sundar Ganesan:** Methodology, Formal analysis. **Jose M.C. Ribeiro:** Formal analysis. **David N. Garboczi:** Methodology, Formal analysis. **Ines Martin-Martin:** Conceptualization, Formal analysis, Investigation, Writing – review & editing. **Eric Calvo:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – review & editing, All authors have read and agreed to the published version of the manuscript.

Declaration of competing interest

There is no conflict of interest.

Acknowledgements

The authors thank Yonas Gebremicale, Andre Laughinghouse and Kevin Lee for their expert mosquito rearing, as well as Glenn Nardone and Lisa Renee Olano of the Research Technology Branch, NIAID, for mass spectrometry analysis. The authors also thank Andrew S. Paige for thoughtful review and comments on the manuscript. BioRender was used to generate the graphical abstract. This research was supported by the Intramural Research Program of NIH/NIAID (AI001246). PC was supported by NIAID BCBB Support Services Contract HHSN316201300006 W/HHSN27200002 to MSC, Inc.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crstbi.2021.03.001.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., et al., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66 (Pt 2), 213–221.

Arca, B., Ribeiro, J.M.C., 2018. Saliva of hematophagous insects: a multifaceted toolkit. Curr. Opinion Insect Sci. 29, 102–109. https://doi.org/10.1016/j.cois.2018.07.012.

Arcá, B., Lombardo, F., de Lara Capurro, M., della Torre, A., Dimopoulos, G., James, A.A., Coluzzi, M., 1999. Trapping cDNAs encoding secreted proteins from the salivary glands of the malaria vector *Anopheles gambiae*. Proc. Natl. Acad. Sci. U.S.A. 96 (4), 1516–1521. https://doi.org/10.1073/pnas.96.4.1516.

Arcà, B., Lombardo, F., Valenzuela, J.G., Francischetti, I.M., Marinotti, O., Coluzzi, M., Ribeiro, J.M., 2005. An updated catalogue of salivary gland transcripts in the adult female mosquito, *Anopheles gambiae*. J. Exp. Biol. 208 (Pt 20), 3971–3986. https://doi.org/10.1242/jeb.01849.

Arensburger, P., Megy, K., Waterhouse, R.M., Abrudan, J., Amedeo, P., Antelo, B., Bartholomay, L., Bidwell, S., Caler, E., Camara, F., Campbell, C.L., Campbell, K.S., Casola, C., Castro, M.T., Chandramouliswaran, I., Chapman, S.B., Christley, S., Costas, J., Eisenstadt, E., Feschotte, C., et al., 2010. Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. Science 330, 86–88. https://doi.org/10.1126/science.1191864, 6000.

Axelrod, D.I., Bailey, H.P., 1968. Cretaceous dinosaur extinction. Evolution 595–611.

Borkent, A., 2008. The frog-biting midges of the World (Corethrellidae: Diptera). Zootaxa 1804, 1–456.

Calvo, E., Sanchez-Vargas, I., Favreau-Lessard, A., Barbian, K., Pham, V., Olson, K., et al., 2010. An insight into the sialotranscriptome of the West Nile mosquito vector, *Culex tarsalis*. BMC Genom. 11, 51. https://doi.org/10.1186/1471-2164-11-51.

Cartwright, T., Hawkey, C., 1969. Activation of blood fibrinolytic mechanism in birds by saliva of vampire bat (*Diaemus youngi*). J. Physiol-London 201, 45.

Chagas, A.C., Calvo, E., Rios-Velásquez, C.M., Pessoa, F.A., Medeiros, J.F., Ribeiro, J.M.C., 2013. A deep insight into the sialotranscriptome of the mosquito, *Psorophora albipes*. BMC Genom. 14, 875. https://doi.org/10.1186/1471-2164-14-875.

Chagas, A.C., Ramirez, J.L., Jasinskiene, N., James, A.A., Ribeiro, J.M., Marinotti, O., Calvo, E., 2014. Collagen-binding protein, Aegyptin, regulates probing time and blood feeding success in the dengue vector mosquito, *Aedes aegypti*. Proc. Natl. Acad. Sci. U.S.A. 111 (19), 6946–6951. https://doi.org/10.1073/pnas.1404179111.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66 (Pt 4), 486–501.

Farajollahi, A., Fonseca, D.M., Kramer, L.D., Kilpatrick, A.M., 2011. "Bird biting" mosquitoes and human disease: a review of the role of *Culex pipiens* complex mosquitoes in epidemiology. Infect. Genet. Evol. 11 (7), 1577–1585.

Garrett-Jones, C., Boreham, P., Pant, C., 1980. Feeding habits of anophelines (Diptera: Culicidae) in 1971-78, with reference to the human blood index: a review. Bull. Entomol. Res. 70, 165–185.

Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., Ferrin, T.E., 2018. UCSF ChimeraX: meeting modern challenges in visualization and analysis. Protein Sci. 1, 14–25.

Harmon, B.G., 1998. Avian heterophils in inflammation and disease resistance. Poultry Sci. 77 (7), 972–977. https://doi.org/10.1093/ps/77.7.972.

Iwanaga, S., Isawa, H., Yuda, M., 2014. Horizontal gene transfer of a vertebrate vasodilatory hormone into ticks. Nat. Commun. 5, 3373. https://doi.org/10.1038/ncomms4373.

Jaroszewski, L., Rychlewski, L., Li, Z., Li, W., Godzik, A., 2005. FFAS03: a server for profile-profile sequence alignments. Nucleic Acids Res. 33, W284–W288.

Juhn, J., Naeem-Ullah, U., Maciel Guedes, B.A., Majid, A., Coleman, J., Paolucci Pimenta, P.F., Akram, W., James, A.A., Marinotti, O., 2011. Spatial mapping of gene expression in the salivary glands of the dengue vector mosquito. Aedes aegypti. Paras. Vectors 4, 1. https://doi.org/10.1186/1756-3305-4-1.

Kabsch, W., 2010. Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D Biol. Crystallogr. 66 (Pt 2), 133–144.

Korochkina, S., Barreau, C., Pradel, G., Jeffery, E., Li, J., Natarajan, R., Shabanowitz, J., Hunt, D., Frevert, U., Vernick, K.D., 2006. A mosquito-specific protein family includes candidate receptors for malaria sporozoite invasion of salivary glands. Cell Microbiol. 8, 163–175. https://doi.org/10.1111/j.1462-5822.2005.00611.x.

Martin-Martin, I., Chagas, A.C., Guimaraes-Costa, A.B., Amo, L., Oliveira, F., Moore, I.N., DeSouza-Vieira, T.S., Sanchez, E.E., Suntravat, M., Valenzuela, J.G., Ribeiro, J., Calvo, E., 2018. Immunity to LuloHya and Lundep, the salivary spreading factors from *Lutzomyia longipalpis*, protects against *Leishmania major* infection. PLoS Pathog. 14 (5), e1007006 https://doi.org/10.1371/journal.ppat.1007006.

Martin-Martin, I., Paige, A., Valenzuela Leon, P.C., Gittis, A.G., Kern, O., Bonilla, B., Chagas, A.C., Ganesan, S., Smith, L.B., Garboczi, D.N., Calvo, E., 2020. ADP binding by the *Culex quinquefasciatus* mosquito D7 salivary protein enhances blood feeding on

mammals. Nat. Commun. 11 (1), 2911. https://doi.org/10.1038/s41467-020-16665-z.
McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. J. Appl. Crystallogr. 40 (Pt 4), 658–674.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 13, 1605–1612.
Pingen, M., Bryden Steven, R., Pondeville, E., Schnettler, E., Kohl, A., Merits, A., et al., 2016. Host inflammatory response to mosquito bites enhances the severity of arbovirus infection. Immunity 44 (6), 1455–1469.
Powell, J.R., Tabachnick, W.J., 2013. History of domestication and spread of *Aedes aegypti*-A Review. Mem. Inst. Oswaldo Cruz 108, 11–17.
Reisen, W.K., Fang, Y., Martinez, V.M., 2005. Avian host and mosquito (Diptera: Culicidae) vector competence determine the efficiency of West Nile and St. Louis encephalitis virus transmission. J. Med. Entomol. 42, 367–375.
Ribeiro, J.M.C., 1987. Role of saliva in blood-feeding by arthropods. Annu. Rev. Entomol. 32, 463–478. https://doi.org/10.1146/annurev.en.32.010187.002335.
Ribeiro, J.M.C., Charlab, R., Valenzuela, J.G., 2001. The salivary adenosine deaminase activity of the mosquitoes Culex quinquefasciatus and Aedes aegypti. J. Exp. Biol. 204 (11), 2001–2010.
Ribeiro, J.M.C., Charlab, R., Pham, V.M., Garfield, M., Valenzuela, J.G., 2004. An insight into the salivary transcriptome and proteome of the adult female mosquito *Culex pipiens quinquefasciatus*. Insect Biochem. Mol. Biol. 34 (6), 543–563. https://doi.org/10.1016/j.ibmb.2004.02.008.
Ribeiro, J.M.C., Mans, B.J., Arcà, B., 2010. An insight into the sialome of blood-feeding Nematocera. Insect Biochem. Mol. Biol. 40 (11), 767–784.
Ribeiro, J.M.C., Chagas, A.C., Pham, V.M., Lounibos, L.P., Calvo, E., 2014. An insight into the sialome of the frog biting fly, *Corethrella appendiculata*. Insect Biochem. Mol. Biol. 44, 23–32. https://doi.org/10.1016/j.ibmb.2013.10.006.
Ribeiro, J.M.C, Martin-Martin, I., Arca, B., Calvo, E., et al., 2016. A Deep Insight into the Sialome of Male and Female *Aedes aegypti* Mosquitoes. PLoS ONE 11 (3). https://doi.org/10.1371/journal.pone.0151400.
Ribeiro, J.M.C., Martin-Martin, I., Moreira, F.R., Bernard, K.A., Calvo, E., 2018. A deep insight into the male and female sialotranscriptome of adult *Culex tarsalis* mosquitoes. Insect Biochem. Mol. Biol. 95, 1–9. https://doi.org/10.1016/j.ibmb.2018.03.001.
Rutenber, E., Ready, M., Robertus, J., 1987. Structure and evolution of ricin B chain. Nature 326, 624–626. https://doi.org/10.1038/326624a0.
Schmid, M.A., Glasner, D.R., Shah, S., Michlmayr, D., Kramer, L.D., Harris, E., 2016. Mosquito saliva increases endothelial permeability in the skin, immune cell migration, and dengue pathogenesis during antibody-dependent enhancement. PLoS Pathog. 12 (6), e1005676 https://doi.org/10.1371/journal.ppat.1005676. PMID: 27310141. PMCID: PMC4911004.
Schneider, B.S., Higgs, S., 2008. The enhancement of arbovirus transmission and disease by mosquito saliva is associated with modulation of the host immune response. Trans. R. Soc. Trop. Med. Hyg. 102 (5), 400–408. https://doi.org/10.1016/j.trstmh.2008.01.024.
Schneider, B.S., Soong, L., Coffey, L.L., Stevenson, H.L., McGee, C.E., Higgs, S., 2010. *Aedes aegypti* saliva alters leukocyte recruitment and cytokine signaling by antigen-presenting cells during West Nile virus infection. PloS One 5 (7), e11704. https://doi.org/10.1371/journal.pone.0011704.
Styer, L.M., Lim, P.-Y., Louie, K.L., Albright, R.G., Kramer, L.D., Bernard, K.A., 2011. Mosquito saliva causes enhancement of West Nile virus infection in mice. J. Virol. 85 (4), 1517.
Tempelis, C., 1975. Host-feeding patterns of mosquitoes, with a review of advances in analysis of blood meals by serology. J. Med. Entomol. 11, 635–653.
Vogt, M.B., Lahon, A., Arya, R.P., et al., 2018. Mosquito saliva alone has profound effects on the human immune system. PLoS Neglected Trop. Dis. 12 (5), e0006439 https://doi.org/10.1371/journal.pntd.0006439.
Zhang, T., de Waard, A.A., Wuhrer, M., Spaapen, R.M., 2019. The role of glycosphingolipids in immune cell functions. Front. Immunol. 10 (90) https://doi.org/10.3389/fimmu.2019.00090.