

RESEARCH ARTICLE

The Diversity of Yellow-Related Proteins in Sand Flies (Diptera: Psychodidae)

Michal Sima^{1*}, Marian Novotny², Lukas Pravda^{3,4}, Petra Sumova¹, Iva Rohousova¹, Petr Volf¹

1 Department of Parasitology, Faculty of Science, Charles University, Prague, Czech Republic, **2** Department of Cell Biology, Faculty of Science, Charles University, Prague, Czech Republic, **3** CEITEC—Central European Institute of Technology, Masaryk University, Brno, Czech Republic, **4** National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Brno, Czech Republic

* michal-sima@seznam.cz



OPEN ACCESS

Citation: Sima M, Novotny M, Pravda L, Sumova P, Rohousova I, Volf P (2016) The Diversity of Yellow-Related Proteins in Sand Flies (Diptera: Psychodidae). PLoS ONE 11(11): e0166191. doi:10.1371/journal.pone.0166191

Editor: Yara M. Traub-Csekö, Instituto Oswaldo Cruz, BRAZIL

Received: August 9, 2016

Accepted: October 24, 2016

Published: November 3, 2016

Copyright: © 2016 Sima et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by Charles University (GAUK project no. 120516/B-BIO, <http://www.cuni.cz/UKEN-1.html>, PS; UNCE 204013, <http://www.cuni.cz/UKEN-1.html>, MS). LP was supported by Ministry of Education, Youth and Sports of the Czech Republic under the project CEITEC 2020 (LQ1601) (<http://www.msmt.cz/?lang=2>, LP) Czech Science Foundation (Project Number: P506 13-05292S, <https://gacr.cz/en/>, to MS, PS, IK, PV). The funders had no role in study

Abstract

Yellow-related proteins (YRPs) present in sand fly saliva act as affinity binders of bioamines, and help the fly to complete a bloodmeal by scavenging the physiological signals of damaged cells. They are also the main antigens in sand fly saliva and their recombinant form is used as a marker of host exposure to sand flies. Moreover, several salivary proteins and plasmids coding these proteins induce strong immune response in hosts bitten by sand flies and are being used to design protecting vaccines against *Leishmania* parasites. In this study, thirty two 3D models of different yellow-related proteins from thirteen sand fly species of two genera were constructed based on the known protein structure from *Lutzomyia longipalpis*. We also studied evolutionary relationships among species based on protein sequences as well as sequence and structural variability of their ligand-binding site. All of these 33 sand fly YRPs shared a similar structure, including a unique tunnel that connects the ligand-binding site with the solvent by two independent paths. However, intraspecific modifications found among these proteins affects the charges of the entrances to the tunnel, the length of the tunnel and its hydrophobicity. We suggest that these structural and sequential differences influence the ligand-binding abilities of these proteins and provide sand flies with a greater number of YRP paralogs with more nuanced answers to bioamines. All these characteristics allow us to better evaluate these proteins with respect to their potential use as part of anti-*Leishmania* vaccines or as an antigen to measure host exposure to sand flies.

Introduction

During intake of a bloodmeal, sand flies (Diptera: Phlebotominae) and other bloodsucking insects inject saliva into the host skin. This saliva contains a mixture of various proteins, which play a major role in preventing host haemostatic and inflammatory responses of different pathways, e.g. platelet activation, coagulation, inflammation, mast cell function, and vasoconstriction (reviewed in [1]). In sand flies, the vectors of *Leishmania* protozoan parasites, these salivary proteins have been studied for decades due to their biological activities and possible

design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

use in anti-*Leishmania* vaccines (reviewed in [2]). In repeatedly bitten hosts, several salivary proteins elicit a strong antibody response, which can be utilized for the detection of exposure to sand flies in epidemiological studies [3–7].

Transcripts of yellow-related proteins (YRPs) have been present in the salivary cDNA libraries of all sand fly species tested to date [8–21]. Usually, they are found in more than one homolog, which may occur in N-glycosylated, O-glycosylated, C-glycosylated, or non-glycosylated forms [16, 20].

YRPs are known as (1) kratagonists that remove small molecule mediators of haemostasis by high affinity ligand-binding proteins, and as (2) antigens that elicit a host immune response, both antibody and cell-mediated. All sequences of sand fly YRPs contain the entire insect-specific MRJP (major royal jelly protein) domain, which defines this protein family across several insect orders/families including *Drosophila melanogaster* [22], honeybees [23], mosquitoes [24], and tse-tse flies [25]. In *Lutzomyia longipalpis*, Xu et al. [26] demonstrated that YRPs are high affinity binders of pro-haemostatic and pro-inflammatory biogenic amines, such as serotonin, histamine and catecholamines. Blocking of these small molecules by YRPs results in vasodilatation, platelet deactivation, and a decrease in vascular permeability [27, 28]. The YRP of *Phlebotomus duboscqi* was shown to have lectin-like activity and is swallowed into the mid-gut, together with saliva [29].

Sand fly salivary recombinant YRPs are the most promising antigens for measuring exposure in naturally bitten hosts, and have been the subject of large epidemiological studies [30–34]. Besides eliciting an antibody response, YRPs as well as plasmids coding these proteins of *Lutzomyia longipalpis* induce a strong delayed type hypersensitivity (DTH) reaction, which leads to protection against *Leishmania major* in vivo [26, 35] and against *L. infantum* in vitro [36]. This suggests a possible use of these proteins in an anti-*Leishmania* vaccine.

In 2011, the crystal structure of *L. longipalpis* YRP LJM11 (GenBank ACCN: AAS05318) was published as 3Q6K (Protein Data Bank ID) with a description of the ligand-binding pocket [26]. Based on this structure and available amino acid sequences obtained from GenBank, we constructed 3D models of all YRPs identified so far in sand fly sialomes. We predicted their phylogenetic relationships, glycosylation sites, surface electrostatic potentials, compared their sequences and characterized the ligand-binding tunnel. Our results show differences among individual proteins within one species as well as differences among various species. Our results may lead to a better understanding of the biological function of YRPs.

Methods

Phylogenetic analysis

Amino acid sequences of YRPs were identified in public databases at NCBI using BLAST [37] based on similarity with *Lutzomyia longipalpis* LJM11 (Protein Data Bank: 3Q6K, for the purpose of this study called 3Q6K_Llon1), the best explored protein from this group, [26]. All these analyses were performed for the sequences without a signal peptide, which was identified using SignalP 4.0 [38]. Sequences were consequently aligned using ClustalX (version 2.0) [39]. The best substitution matrix for creating a phylogenetic tree of sand fly salivary YRPs was determined in ProtTest software 2.0 [40]. TREEPUZZLE 5.2 [41] was used to create a maximum likelihood phylogenetic tree from the protein alignment using the WAG model [42] and quartet puzzling with 10000 puzzling steps. The resulting tree for all 31 proteins from 11 sand fly species in two genera was visualized in MEGA 4 [43] and rooted by the related protein from *Drosophila melanogaster* (ACCN: NP650247). Clustal Omega [44] was used with default settings to calculate a Percent Identity Matrix among all sand fly salivary YRPs.

Prediction of glycosylation

Putative N-, O-, and C- glycosylation sites for all 31 protein sequences were determined using NetNGlyc 1.0, NetOGlyc 4.0 [45], and NetCGlyc 1.0 [46] servers with default settings.

3D models construction

All proteins were modeled using 3Q6K, the only available structure of sand fly salivary YRP in Protein Data Bank (PDB, [47]), from where its PDB file and fasta sequence were downloaded. The alignment of template and target sequence was done in Clustal Omega [44] for all proteins. Scripts for Python version 2.7 (Python Software Foundation) were prepared in a txt file and ran in MODELLER [48]. Five models were calculated for each protein, with the best one chosen based on the lowest energy levels of molpdf and the DOPE score. All models were displayed and analyzed in PyMOL (The PyMOL Molecular Graphics System, Version 1.5 Schrödinger, LLC.). Electrostatic surface potentials were calculated using the APBS Tools2 plugin [49] in PyMOL.

Tunnel analysis

Tunnels in the protein structures were detected and characterized using the channel analysis tool MOLE 2.0 [50]. All proteins were superimposed using PyMOL in order to use the same settings, and therefore obtain comparable tunnels. In brief, MOLE 2.0 calculates a Delaunay triangulation/Voronoi diagram of the atomic centers. Next, tunnels are identified between every tuple of user defined end points within a protein structure using Dijkstra's algorithm. The resulting tunnels are defined by their centerline and are uniformly divided into layers. Each layer is defined by the residues lining it. A new layer starts whenever there is a change in residues lining it along its length. Additionally, MOLE 2.0 infers basic physicochemical properties for each tunnel as well as layers. These were calculated from a unique set of lining residues averaging tabulated values for hydrophobicity [51]. WebLogo 2.8.2. [52] and Clustal Omega [44] were used to visualize differences among protein tunnels.

Results

Identification of sand fly salivary YRPs

Thirty two sand fly salivary YRPs were identified based on their similarity with PDB ID: 3Q6K from *L. longipalpis* [26], the only known 3D structure of YRPs of phlebotominae sand flies. These YRPs were detected in all 13 sand fly species with a published salivary gland transcriptome: 9 *Phlebotomus* species from 5 subgenera (*Phlebotomus*, *Paraphlebotomus*, *Larrousius*, *Adlerius*, and *Euphlebotomus*), and 4 *Lutzomyia* species from 3 subgenera (*Lutzomyia*, *Nyssomyia*, and *Helcocyratomyia*). In summary, there are currently 33 different YRPs described in the Phlebotominae subfamily (Table 1). There are large differences in the number of these proteins found in various species, ranging from 1 in *P. arabicus* and *P. argentipes* to 5 in *P. sergenti*.

Phylogenetic analysis

The phylogenetic analysis was performed using the maximum likelihood phylogenetic tree of all 33 YRPs from 13 sand fly species. A higher variability in *Lutzomyia* YRPs compared to *Phlebotomus* proteins was found: two clusters from seven *Lutzomyia* YRPs were detected but four other *Lutzomyia* proteins created own branches. Twenty two *Phlebotomus* proteins clustered together in one branch with high bootstrap support. A closer phylogenetic relationship was discovered between the subgenera *Phlebotomus* (*P. papatasi* and *P. duboscqi*) and *Paraphlebotomus* (*P. sergenti*), as well as between *Larrousius* (*P. tobbi*, *P. orientalis*, *P. perniciosus*, and *P.*

Table 1. Identified sand fly salivary YRPs.

ACCN or name	Sand fly species	Subgenus	Published by	Identifier
BAM69109	<i>L. ayacuchensis</i>	<i>Helcocyrtomyia</i>	[19]	Laya1
BAM69110	<i>L. ayacuchensis</i>	<i>Helcocyrtomyia</i>	[19]	Laya2
BAM69111	<i>L. ayacuchensis</i>	<i>Helcocyrtomyia</i>	[19]	Laya3
BAM69185	<i>L. ayacuchensis</i>	<i>Helcocyrtomyia</i>	[19]	Laya4
AFP99235	<i>L. intermedia</i>	<i>Nyssomyia</i>	[17]	Lint
AAS05318, LJM11, 3Q6K	<i>L. longipalpis</i>	<i>Lutzomyia</i>	[10, 25]	3Q6K_Llon1
AAD32198, LJM17	<i>L. longipalpis</i>	<i>Lutzomyia</i>	[8]	Llon2
ABB00904, LJM111	<i>L. longipalpis</i>	<i>Lutzomyia</i>	[10]	Llon3
ANW11467	<i>L. olmeca</i>	<i>Nyssomyia</i>	[21]	Lolm1
ANW11468	<i>L. olmeca</i>	<i>Nyssomyia</i>	[21]	Lolm2
ANW11469	<i>L. olmeca</i>	<i>Nyssomyia</i>	[21]	Lolm3
ACS93501	<i>P. arabicus</i>	<i>Adlerius</i>	[14]	Para
ABA12136	<i>P. argentipes</i>	<i>Euphlebotomus</i>	[12]	Parg
AAX44093	<i>P. ariasi</i>	<i>Larrousius</i>	[11]	Pari1
AAX56360	<i>P. ariasi</i>	<i>Larrousius</i>	[11]	Pari2
ABI15938	<i>P. duboscqi</i>	<i>Phlebotomus</i>	[13]	Pdub1
ABI15941	<i>P. duboscqi</i>	<i>Phlebotomus</i>	[13]	Pdub2
ABI20172	<i>P. duboscqi</i>	<i>Phlebotomus</i>	[13]	Pdub3
AGT96460	<i>P. orientalis</i>	<i>Larrousius</i>	[20]	Pori1
AGT96461	<i>P. orientalis</i>	<i>Larrousius</i>	[20]	Pori2
AAL11051	<i>P. papatasi</i>	<i>Phlebotomus</i>	[9]	Ppap1
AAL11052	<i>P. papatasi</i>	<i>Phlebotomus</i>	[9]	Ppap2
AGE83094	<i>P. papatasi</i>	<i>Phlebotomus</i>	[15]	Ppap3
AGE83095	<i>P. papatasi</i>	<i>Phlebotomus</i>	[15]	Ppap4
ABA43049	<i>P. perniciosus</i>	<i>Larrousius</i>	[12]	Pper1
ABA43050	<i>P. perniciosus</i>	<i>Larrousius</i>	[12]	Pper2
ADJ54114	<i>P. sergenti</i>	<i>Paraphlebotomus</i>	[16]	Pser1
ADJ54115	<i>P. sergenti</i>	<i>Paraphlebotomus</i>	[16]	Pser2
ADJ54116	<i>P. sergenti</i>	<i>Paraphlebotomus</i>	[16]	Pser3
ADJ54122	<i>P. sergenti</i>	<i>Paraphlebotomus</i>	[16]	Pser4
ADJ54123	<i>P. sergenti</i>	<i>Paraphlebotomus</i>	[16]	Pser5
ADJ54079	<i>P. tobbi</i>	<i>Larrousius</i>	[16]	Ptob1
ADJ54080	<i>P. tobbi</i>	<i>Larrousius</i>	[16]	Ptob2

Sequence GenBank accession numbers (or names), sand fly species (ordered alphabetically), references, and identifier names used in this study are provided for each protein.

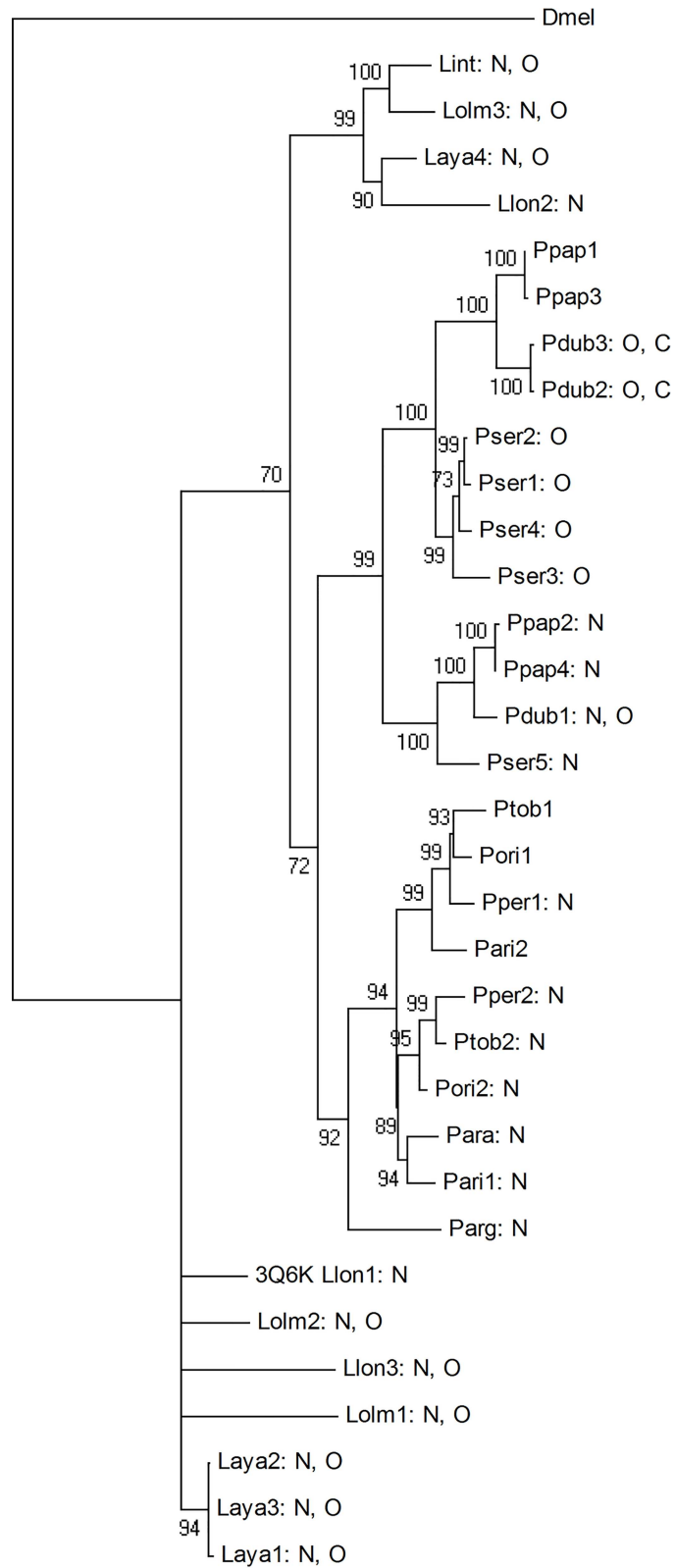
doi:10.1371/journal.pone.0166191.t001

ariasi) and *Adlerius* (*P. arabicus*) species. *Euphlebotomus* (*P. argentipes*) protein created a single branch, which is closer to species from subgenera *Larrousius* and *Adlerius* (Fig 1).

Based on the identity matrix from Clustal Omega, the similarity among all 33 YRPs (528 combinations) varied from 44 to 99% (S1 Table).

Glycosylation

For all 33 proteins, potential N-, O- and C-glycosylation sites were predicted. In 11 proteins, only N-glycosylation sites were predicted. In four YRPs of *P. sergenti* (Pser1-Pser4), only O-glycosylation sites were identified, and C-glycosylation sites were predicted for only two



0.1

Fig 1. Phylogenetic tree of 33 sand fly salivary YRPs with putative glycosylation sites marked. The maximum likelihood phylogenetic tree was created in TREEPUZZLE with WAG model using quartet puzzling with 10000 puzzling steps. For rooting, the related protein (ACCN: NP650247) from *Drosophila melanogaster* (Dmel) was used. Bootstraps with support for branching are shown. The letters N, O, and C indicate putative N-, O-, and C-glycosylation sites, respectively. Protein codes refer to Table 1.

doi:10.1371/journal.pone.0166191.g001

proteins from *P. duboscqi* (Table 2). In eleven proteins, more than one type of putative glycosylation was found.

The predicted glycosylation sites for the majority of YRPs (77% and 53% of putative N-glycans and O-glycans, respectively) share the same positions (Table 2). Importantly, proteins clustering together in the phylogenetic tree (Fig 1) had similar predicted glycosylation patterns.

Table 2. Putative glycosylation sites.

Identifier	N-glycosylation	O-glycosylation	C-glycosylation
Laya1	Asn 164, Asn 196	Ser 208	-
Laya2	Asn 164, Asn 196	Ser 113, Ser 208	-
Laya3	Asn 164, Asn 196	Ser 208	-
Laya4	Asn 11, Asn 264	Ser 262	-
Lint	Asn 11	Thr 262, Thr 263	-
3Q6K_Llon1	Asn 195	-	-
Llon2	Asn 11	-	-
Llon3	Asn 123	Ser 300	-
Lolm1	Asn 122	Thr 262	-
Lolm2	Asn 194	Thr 261	-
Lolm3	Asn 11	Thr 197	-
Para	Asn 11	-	-
Parg	Asn 11, Asn 18, Asn 307	-	-
Pari1	Asn 11, Asn 201	-	-
Pari2	-	-	-
Pdub1	Asn 11, Asn 65	Thr 367	-
Pdub2	-	Ser 208	Trp 338
Pdub3	-	Ser 208	Trp 338
Pori1	-	-	-
Pori2	Asn 11	-	-
Ppap1	-	-	-
Ppap2	Asn 11, Asn 65	-	-
Ppap3	-	-	-
Ppap4	Asn 11, Asn 65	Ser 264	-
Pper1	Asn 11	-	-
Pper2	Asn 11	-	-
Pser1	-	Ser 81, Ser 208, Thr 263	-
Pser2	-	Ser 208, Thr 263	-
Pser3	-	Ser 208, Thr 263	-
Pser4	-	Ser 208	-
Pser5	Asn 11, Asn 65, Asn 250	-	-
Ptob1	-	-	-
Ptob2	Asn 11	-	-

N-, O- and C-glycosylation sites were predicted for each protein using glycosylation servers. Numbers and three-letter abbreviations indicate the positions of predicted glycosylation and the amino acid where the glycosylation occurs (Asn—asparagine, Thr—threonine, Ser—serine, and Trp—tryptophan).—shows cases where no glycosylation sites were identified. Protein codes refer to Table 1.

doi:10.1371/journal.pone.0166191.t002

3D model analysis

The amino acid sequence alignment of 33 sand fly salivary YRPs showed high conservancy in several regions, with 57 invariant sites identified among all sequences, representing approximately 15% of the total sequence. Five out of eleven amino acids creating the ligand-binding pocket known from Xu et al. [26] were identical across all proteins. The most variable position in the ligand-binding site was the one annotated as Phe 344 in the crystal structure of YRP 3Q6K—this position is occupied by six different amino acids (His, Gln, Tyr, Phe, Met, or Lys) in our studied set of proteins. Two most common amino acids on this position are histidine and glutamine (Fig 2).

Models for 32 different proteins were created based on the template sequence 3Q6K from *L. longipalpis* (S1 Fig). All of them folded as a six-bladed beta-propeller, similarly as in 3Q6K [26]. In 3Q6K, 11 ligand-binding amino acids were identified [26]. Three of them (in the case of 3Q6K –Thr 327 with 2 bonds, Asn 342 with 2 bonds, and Phe 344 with 1 bond) play a major role in binding abilities, because of the hydrogen bonds between the protein amino acids and

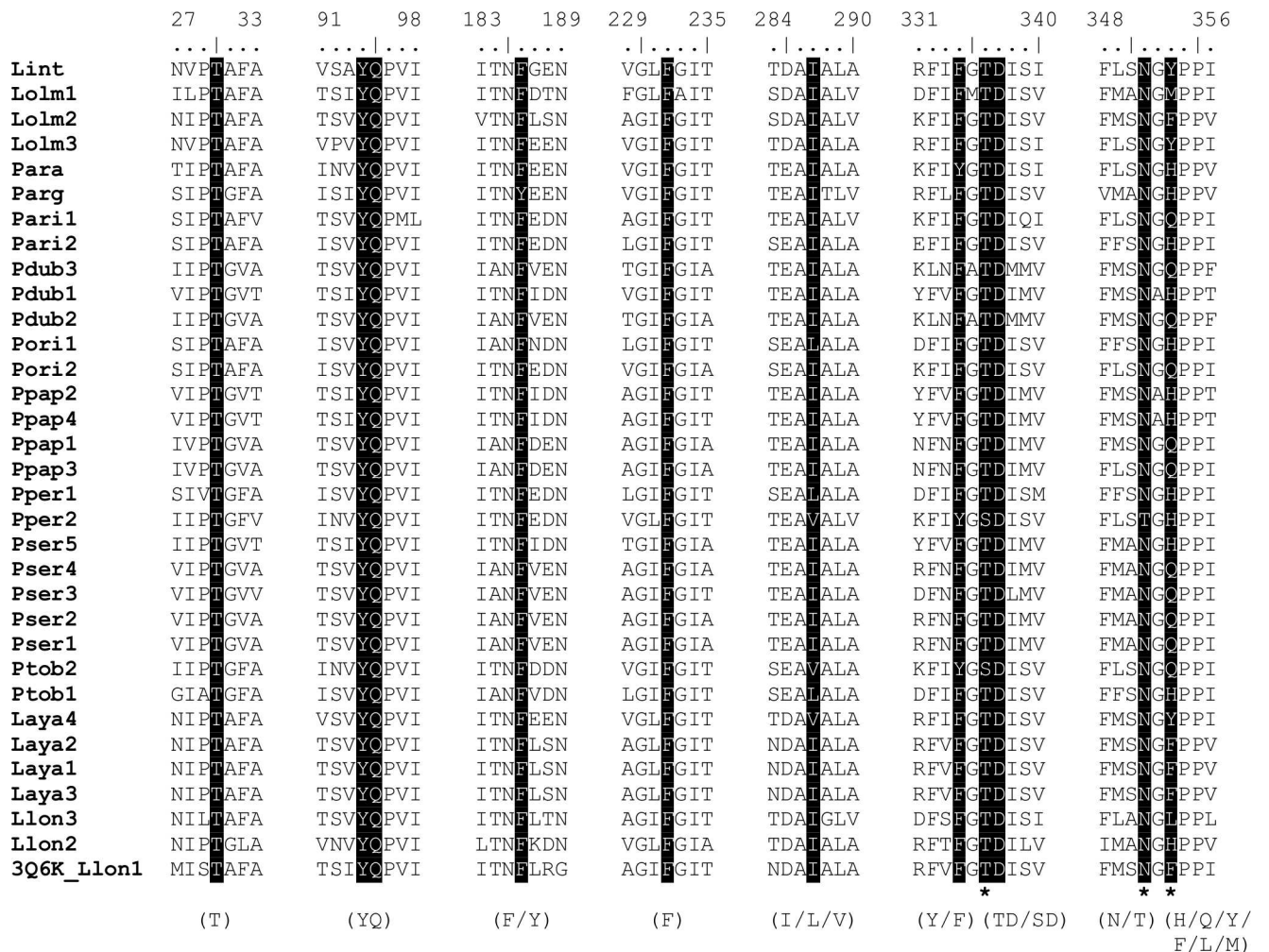


Fig 2. Amino acid alignment within the YRP ligand-binding sites. The alignment was created in ClustalX, and shows high conservancy in binding sites (letters in brackets below the sequences, slashes indicate amino acid substitutions in appropriate positions). Ligand binding sites are highlighted in black, three bold stars below the sequences represent the main binding sites where hydrogen bonds between amino acids and the ligand are predicted. Protein codes refer to Table 1.

doi:10.1371/journal.pone.0166191.g002

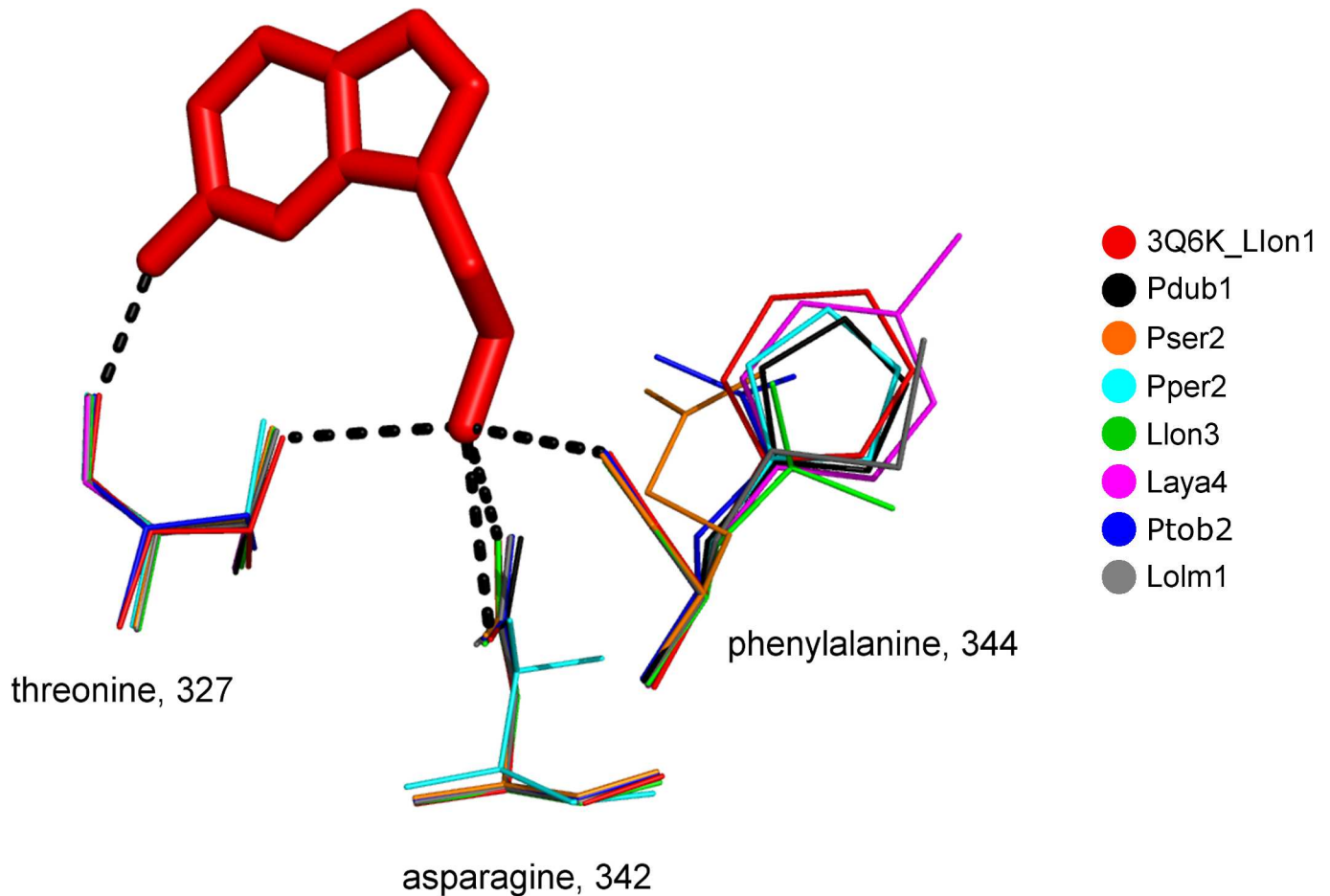


Fig 3. Hydrogen-binding between serotonin and ligand-binding amino acids in sand fly salivary YRPs. The figure shows the variability in hydrogen bonds between serotonin and 3Q6K (provided by Thr 327, Asn 342, and Phe 344) described in Xu et al. [26] and other YRP models created during this study and visualized in PyMOL. The red bold molecule symbolizes the ligand serotonin and the black dashed line show the hydrogen bonds between amino acids and serotonin in 3Q6K. Other colors indicate one representative member of one protein group. A group was defined as having at least one unique mutation from the template structure in one of the three described amino acids binding the ligand. From each group, the protein with the most abundant amino acid rotamer was chosen as being representative for this visualization. Protein codes refer to Table 1.

doi:10.1371/journal.pone.0166191.g003

the ligand. The other 8 amino acids bind serotonin by Van der Waals or hydrophobic interactions. These three amino acids were visualized in the 3D alignment of all 33 proteins. Nine different amino acids were found at three positions that create hydrogen bonds with serotonin from 3Q6K (Fig 3).

As well as in 3Q6K_Llon1, in Laya1-3 and Lolm1, the amino acids creating hydrogen bonds with serotonin are Thr 327, Asn 342, and Phe 344 (using the numbering for 3Q6K). The largest group includes 11 proteins (Llon2, Para, Parg, Pari, Pdub1, Pori2, Ppap2, Ppap4, Pper1, Pser5, and Ptob1). In comparison with the group containing 3Q6K_Llon1, there is a substitution of histidine instead of phenylalanine in the last binding position. A substitution in the same position (glutamine instead of phenylalanine) also occurs in the second biggest group, which includes 10 proteins (Pari1, Pdub2, Pdub3, Pori2, Ppap1, Ppap3, Pser1, Pser2, Pser3, and Pser4). The last group with more than one protein (Laya4, Lint, and Lolm3) has substitution in the last binding position, as well. In this case, there is tyrosine instead of phenylalanine. Pper2 has substitutions in all three positions (serine, threonine, and histidine). Because of these

substitutions (mainly due to the presence of Thr instead of Asn) these proteins are likely to have different binding ability to biamines from 3Q6K_Llon1. In the last binding positions of Llon3 and Lolm1 there are leucine and methionine, respectively, instead of phenylalanine. Pto2 contains a substitution in the first binding position, with threonine replaced by serine. The distribution of YRPs in these groups partially corresponds with branches in the phylogenetic tree (S2 Fig).

Tunnel in YRPs

In Xu et al. [26], YRP 3Q6K from *L. longipalpis* was described as a hollow barrel with two possible entrances. Here, based on the models created using other sand fly YRPs and MOLE analysis, we found a very similar hollow barrel in all proteins we studied (Table 3). It is therefore likely that all studied proteins retain the ability to bind a ligand. The differences in a length and minimal radius of the identified tunnels have limited relevance due to the fact that they are calculated on theoretical models. On contrary, differences in hydrophobicity should be less influenced by homology modeling process and could therefore indicate (where present) changes in binding affinities among paralogs within species.

The ligand-binding site for serotonin in 3Q6K is located approximately one third of the way along this tunnel (Fig 4). The comparison based on the APBS tools in PyMol of all models and the 3Q6K structure revealed that the surface electrostatic potential of the entrance closer to the binding site varies from negative to positive (S2 Fig), and the size of this entrance is smaller than the opposite entrance. In contrast, the entrance more distant from the binding site is larger and neutrally or negatively charged in all these proteins (Table 3). The distribution of the surface potentials in most of these proteins corresponds with the phylogenetic relationships (Fig 1) and predicted glycosylation sites (Table 2).

The sequence conservation of the amino acids lining the tunnels was explored using WebLogo server (Fig 5). The WebLogo was based on the multiple sequence alignment (S3 Fig) of all amino acids lining the accessible tunnels towards the ligand-binding site, as described by MOLE. The analysis showed that there are not any absolutely conserved amino acids, but both entrances to the tunnel are dominated by negatively-charged residues; however, there are proteins (e.g. Pdub1, Pari2, Ppar1, Ppap4, and Pser5) that are positively charged at least at one of the entrances. The middle section of the tunnel lining is occupied mostly by hydrophobic and small polar amino acids.

Discussion

Here we focused on the bioinformatic characterization of sand fly salivary YRPs. These proteins have high affinity binding properties for pro-haemostatic and pro-inflammatory biogenic amines [26], but they are also highly antigenic for hosts bitten by sand flies [30–34] and can be used for anti-*Leishmania* vaccine development [53]. We searched published transcriptomes and recombinant proteins and found 33 different proteins from this family among 13 species from the two genera—Old World *Phlebotomus* and New World *Lutzomyia*. We analyzed their phylogenetic relationships, predicted glycosylation sites, constructed 3D models, identified tunnels passing through these structures, described the ligand-binding site, and compared their surface electrostatic potentials. The number of detected YRPs in these 13 sand fly species varied from one in *P. arabicus* [14], *P. argentipes* [12], and *L. intermedia* [17] to five in *P. sergenti* [16]. Such interspecies variability might be partially attributed to the sensitivity of sequencing methods; however, this is not likely the case of YRPs that were analyzed in the same laboratory using the same method. For example, Anderson et al. [12] reported a single YRP in *P. argentipes* but two in *P. perniciosus*, while Rohousova et al. [16] reported two in *P. tobbi* but five in *P.*

Table 3. A comparison of tunnels in sand fly salivary YRPs.

Identificator	Length	Min. Radius	Hydrophobicity	Ligand-side	Opposite-side
Laya1	38.320	1.769	-0.13	Positive	Negative
Laya2	38.707	1.813	-0.22	Positive	Negative
Laya3	41.045	0.962	-0.17	Positive	Negative
Laya4	41.295	1.785	-0.25	Neutral (negative)	Neutral (negative)
Lint	39.194	1.539	-0.32	Negative	Negative
3Q6K_Llon1	41.245	1.892	-0.23	Positive	Neutral (negative)
Llon2	39.722	1.361	-0.07	Negative	Negative
Llon3	43.634	1.831	-0.29	Negative	Negative
Lolm1	38.595	2.055	-0.21	Negative (neutral)	Neutral
Lolm2	41.696	1.997	-0.45	Positive (neutral)	Neutral (negative)
Lolm3	43.098	1.416	-0.86	Neutral (negative)	Negative (neutral)
Para	45.294	1.229	-0.53	Negative	Negative
Parg	40.506	1.262	-0.20	Negative	Negative
Pari1	45.068	1.673	-0.02	Negative	Negative
Pari2	41.717	1.880	-0.29	Negative	Negative
Pdub1	37.292	1.744	-0.14	Negative	Neutral (negative)
Pdub2	39.912	1.433	0.03	Neutral (positive)	Neutral (negative)
Pdub3	35.967	2.021	-0.34	Negative	Neutral (positive, negative)
Pori1	42.203	1.677	-0.22	Negative	Negative
Pori2	32.368	1.824	-0.18	Negative	Neutral (negative)
Ppap1	38.163	1.834	-0.23	Negative	Neutral (positive)
Ppap2	40.508	1.780	-0.05	Neutral (negative)	Negative (neutral)
Ppap3	40.849	2.053	-0.08	Negative	Neutral (positive)
Ppap4	46.137	1.518	-0.27	Neutral	Neutral (negative, positive)
Pper1	41.652	1.419	-0.18	Negative	Negative
Pper2	41.944	1.948	-0.25	Neutral (negative)	Neutral (negative)
Pser1	50.023	2.047	-0.27	Negative	Neutral
Pser2	37.843	1.755	-0.05	Negative	Neutral
Pser3	38.013	1.757	0.02	Neutral	Negative
Pser4	31.949	1.812	0.01	Negative	Negative
Pser5	40.761	1.853	-0.01	Negative	Neutral
Ptob1	33.603	1.740	-0.31	Negative	Negative
Ptob2	34.118	1.822	-0.22	Neutral (positive)	Neutral (negative)

The length and minimum radius (in Ångströms) and hydrophobicity are shown for the tunnels of each protein. All values were calculated using MOLE. Surface electrostatic potentials around protein entry sides were calculated in PyMOL using the APBS tool. "Ligand-side" refers to the side with the shorter distance from the beginning of the tunnel to the ligand-binding site, while "Opposite-side" is the other side of the tunnel. The most common electrostatic surface potentials of both entrances to the tunnel are shown. Charges in brackets indicate that the charge around the entrance is not uniform, and small parts of the entrance are charged differently from the majority of the entrance. Protein codes refer to [Table 1](#).

doi:10.1371/journal.pone.0166191.t003

sergenti. It is also possible that occurrence of various numbers of YRPs is caused by differential gene expression. Genes with lower expression might be recorded less frequently as it was shown for other sand fly protein families [54].

Our phylogenetic analysis suggested a higher variability in *Lutzomyia* than in *Phlebotomus* YRPs. Sequence identity among YRPs ranged between 44 and 99 percent, with a majority of proteins having sequence identity around 50 percent. Further analysis divided these proteins into several branches, with closer relationships within sand flies from the same subgenera

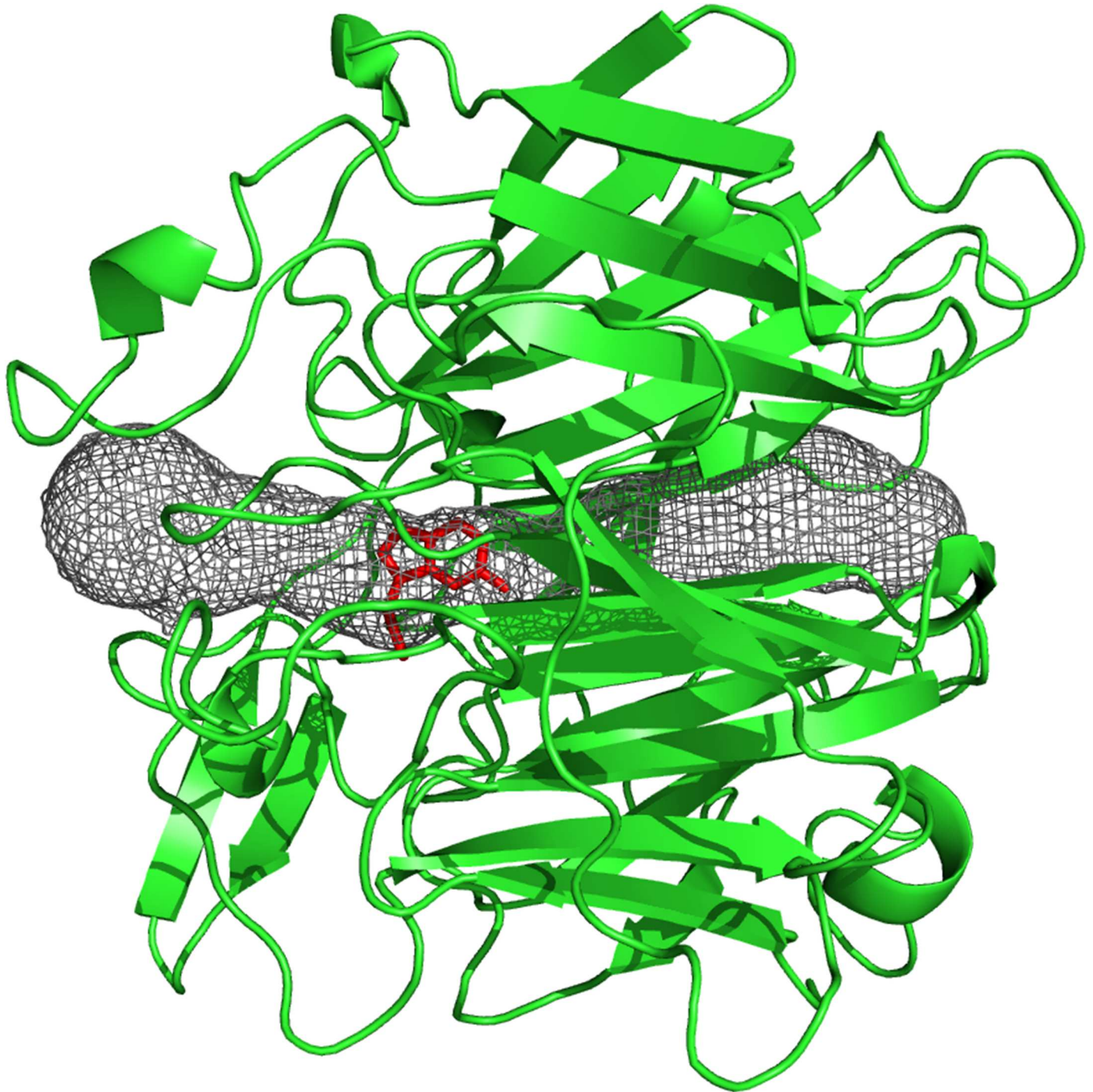


Fig 4. Visualization of the tunnel in 3Q6K_Llon1. The structure was visualized in PyMOL using the MOLE script for calculating the tunnel. The 3Q6K_Llon1 protein is drawn in green, the tunnel through this protein structure is in black mesh and the red-stick molecule represents serotonin.

doi:10.1371/journal.pone.0166191.g004

(*Phlebotomus*, *Paraphlebotomus*, and *Larrousius*). This confirmed previous analyses based on amino acid sequences [16, 20] and also corresponds with an analysis based on small subunit nuclear ribosomal DNA [55] showing two main clades—*Phlebotomus* with *Paraphlebotomus*



Fig 5. Sequence conservation of the amino acids lining the ligand-binding tunnel. The WebLogo server was used to visualize the sequence conservation of the amino acids lining the ligand-binding tunnel in 31 YRPs. The amino acids are colored according to their physical-chemical properties. LS and OS below the x-axis indicate the side closer to the ligand-binding site and the opposite side of the tunnel, respectively.

doi:10.1371/journal.pone.0166191.g005

subgenera clustering together and leaving other subgenera (e.g. *Euphlebotomus*, *Larrousius*, and *Adlerius*) in the second clade.

The biogenic function of proteins may be affected by the presence or absence of posttranslational modifications [56, 57]. For example, the antibody response against some glycoproteins can be directed strictly to the glycan part [58]. In accordance with our previous studies [16, 20, 54, 59], we showed both inter- as well as intra-species variability in the glycosylation of sand fly YRPs, which corresponded well with the phylogenetic analysis. Moreover, the same positions of putative glycosylation sites among YRPs mirrored the presumed effects of glycosylation on protein folding and stability, suggesting a highly conserved tertiary structure. It would be interesting to evaluate similarities in the antigenic potential of the proteins belonging to the same clusters.

The only protein with an experimentally solved 3D structure, LJM11 (3Q6K) from *Lutzomyia longipalpis*, has a unique six-bladed beta propeller fold that has only been identified in this protein. It is characterized by a long ligand-binding tunnel with two entrances—one on each side of the protein. Our 3D models of other YRPs showed similar structures as 3Q6K, and the tunnel is present in all studied proteins. Its length or minimum width vary between and within species but it remains uninterrupted in all modeled proteins, suggesting it is very important feature of these proteins. LJM11 is known to bind positively charged biogenic amines [26], and our structural analysis of YRPs from other sand fly species showed that in all tested sequences the ligand-binding tunnels were negatively charged and could therefore bind positively charged ligands of bioamine size. Nevertheless, we observed that the proteins and more specifically the paralogs within one species differ in the charge of tunnel entrances. Our data indicate that paralogs within one species will bind similar ligands with different affinities, and the ligands may travel different paths to the binding site according to the charge of the entrances. YRPs also differed within and among species in the length of their tunnels, the minimum radius of the tunnel and its hydrophobicity. We expect that all these parameters affect the nature and the affinity of the preferred ligand, which results in a greater binding versatility of the host bioamines.

In the 3Q6K YRP structure, the serotonin ligand is stabilized by hydrogen bonds from three amino acids in the main chain and side chain and by a number of additional hydrophobic interactions [26]. Both our sequence and structure analyses showed the diversity of this crucial site among the 33 proteins studied here. The largest diversity was observed on the position denoted in 3Q6K as Phe 344 (see Fig 3). The amino acid at this position participates in ligand-binding through a hydrogen bond from the main chain. Substitutions at this position should

not play a major role in ligand-binding, yet it is the only amino acid of the ligand-binding-site that is different among three *L. longipalis* yellow-related proteins that have been shown to have different affinities for serotonin and other tested bioamines [26]. Our models of the other YRPs suggest that at least in the case of the Phe344Gln mutation, the glutamine side chain could provide a further hydrogen bond to the ligand (serotonin). Interestingly, our studied YRPs can be divided into several groups according to their ligand-binding composition, and some groups correlate with the branches identified in phylogenetic distribution as well as with the predicted glycosylation patterns. For example, Ppap2, Ppap4, Pdub1, and Pser5 (proteins from the largest group with binding amino acids Thr-Asn-His) or Pser1-4 (proteins from the second largest group with binding amino acids Thr-Asn-Glu) are located in two branches in the phylogenetic tree and share the same N-glycosylation and O-glycosylation pattern, respectively (see S2 Fig).

The lining of the tunnel is not conserved, but the entrances to the tunnel are composed mostly of negatively charged residues and the inner part of the tunnel is lined by hydrophobic and small polar amino acids (threonines). There are however several proteins (Pdub1, Ppap4, Pser5 etc.) that have positively charged residues forming one of the entrances (typically the one closer to ligand-binding site within the tunnel)—this might allow interaction with different ligands.

In summary, we found interesting diversity among these important proteins. All 33 sand fly salivary YRPs studied here share similar folding and contain a ligand-binding tunnel. Nevertheless, modifications found among these proteins influence the charge of entrances to the tunnel, the length of the tunnel and/or its hydrophobicity. We suggest that these modifications provide sand fly species with multiple paralogs of YRPs with more nuanced answers to host bioamines. Further experiments are needed to validate these models, but we assume, that our results should allow a better understanding of the biological role of these important proteins with respect to their potential use in anti-*Leishmania* vaccines or as host exposure markers to sand flies.

Supporting Information

S1 Fig. Modeled 3D structures of sand fly yellow-related proteins. Zipped PDB files of 32 sand fly yellow-related proteins calculated in MODELLER based on 3Q6K as the template structure. Protein codes refers to Table 1.
(ZIP)

S2 Fig. Visual summary of the diversity of yellow-related proteins in sand flies. The maximum likelihood phylogenetic tree was created in TREEPUZZLE with WAG model using quartet puzzling with 10000 puzzling steps. For rooting, the related protein (ACCN: NP650247) from *Drosophila melanogaster* (Dmel) was used. Bootstraps with support for branching are shown. Protein codes refer to Table 1. Symbols preceding the protein codes indicate the most common surface electrostatic potential of the entrance to the protein tunnel closer to the ligand-binding site (LS) as shown in Table 3 (◇ positive, ○ neutral, and □ negative). Colors in symbols indicate the affiliation to protein groups with the same ligand-binding amino acids as shown in Fig 3. The letters N, O, and C indicate putative N-, O-, and C-glycosylation sites, respectively, based on Table 2.
(PDF)

S3 Fig. Alignment of amino acids creating tunnels in yellow-related proteins. For each protein, amino acids creating tunnel were determined using MOLE. Clustal Omega was used to visualize differences of these protein tunnels. In sequences, small letters indicate main chain

interaction and capital letters represent side chain interactions. Above the sequences, LS and OS describe the side closer to the ligand binding site and the opposite side of the tunnel, respectively. Protein codes refer to [Table 1](#).

(PDF)

S1 Table. Percent Identity Matrix of sand fly yellow-related proteins. Clustal Omega was used to calculate this sequence identity matrix among all identified yellow-related proteins. The similarity between the proteins is shown in percents. Different colors represent 10 percent similarity intervals. Protein codes refer to [Table 1](#).

(PDF)

Acknowledgments

We appreciate Helena Kulikova and Lenka Zitkova for excellent administrative assistance and Simon Kerdik for the help with graphical images.

Author Contributions

Conceptualization: MN.

Formal analysis: MN.

Investigation: MS.

Methodology: MS.

Software: LP.

Supervision: IR PV.

Visualization: MS MN.

Writing – original draft: MS MN LP PS.

Writing – review & editing: MN PV IR.

References

1. Ribeiro JMC, Mans BJ, Arca B. An insight into the sialome of bloodfeeding Nematocera. *Insect Biochem Mol Biol.* 2010; 40: 767–784. doi: [10.1016/j.ibmb.2010.08.002](https://doi.org/10.1016/j.ibmb.2010.08.002) PMID: [20728537](https://pubmed.ncbi.nlm.nih.gov/20728537/)
2. Abdeladhim M, Kamhawi S, Valenzuela JG. What's behind a sand fly bite? The profound effect of sand fly saliva on host hemostasis, inflammation and immunity. *Infect Genet Evol.* 2014; 28: 691–703. doi: [10.1016/j.meegid.2014.07.028](https://doi.org/10.1016/j.meegid.2014.07.028) PMID: [25117872](https://pubmed.ncbi.nlm.nih.gov/25117872/)
3. Rohousova I, Ozensoy S, Ozbel Y, Volf P. Detection of species-specific antibody response of humans and mice bitten by sand flies. *Parasitology.* 2005; 130: 493–499. PMID: [15991492](https://pubmed.ncbi.nlm.nih.gov/15991492/)
4. de Moura TR, Oliveira F, Novais FO, Miranda JC, Clarencio J, Follador I, et al. Enhanced *Leishmania braziliensis* infection following pre-exposure to sandfly saliva. *PLoS Negl Trop Dis.* 2007; 1: e84. doi: [10.1371/journal.pntd.0000084](https://doi.org/10.1371/journal.pntd.0000084) PMID: [18060088](https://pubmed.ncbi.nlm.nih.gov/18060088/)
5. Vikova M, Rohousova I, Drahota J, Stanneck D, Kruedewagen EM, Mencke N, et al. Canine antibody response to *Phlebotomus perniciosus* bites negatively correlates with the risk of *Leishmania infantum* transmission. *PLoS Negl Trop Dis.* 2011; 5: e1344. doi: [10.1371/journal.pntd.0001344](https://doi.org/10.1371/journal.pntd.0001344) PMID: [22022626](https://pubmed.ncbi.nlm.nih.gov/22022626/)
6. Gidwani K, Picado A, Rijal S, Singh SP, Roy L, Volfova V, et al. Serological markers of sand fly exposure to evaluate insecticidal nets against visceral leishmaniasis in India and Nepal: A cluster-randomized trial. *PLoS Negl Trop Dis.* 2011; 5: e1296. doi: [10.1371/journal.pntd.0001296](https://doi.org/10.1371/journal.pntd.0001296) PMID: [21931871](https://pubmed.ncbi.nlm.nih.gov/21931871/)
7. Marzouki S, Ben Ahmed M, Boussoffara T, Abdeladhim M, Ben Aleya-Bouafif N, Namane A, et al. Characterization of the antibody response to the saliva of *Phlebotomus papatasi* in people living in

- endemic areas of cutaneous leishmaniasis. *Am J Trop Med Hyg.* 2011; 84: 653–661. doi: [10.4269/ajtmh.2011.10-0598](https://doi.org/10.4269/ajtmh.2011.10-0598) PMID: [21540371](https://pubmed.ncbi.nlm.nih.gov/21540371/)
8. Charlab R, Valenzuela JG, Rowton ED, Ribeiro JMC. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly *Lutzomyia longipalpis*. *Proc Natl Acad Sci USA.* 1999; 96: 15155–15160. PMID: [10611354](https://pubmed.ncbi.nlm.nih.gov/10611354/)
 9. Valenzuela JG, Belkaid Y, Garfield MK, Mendez S, Kamhawi S, Rowton ED. Toward a defined anti-*Leishmania* vaccine targeting vector antigens: characterization of a protective salivary protein. *J Exp Med.* 2011; 194: 331–342.
 10. Valenzuela JG, Garfield M, Rowton ED, Pham VM. Identification of the most abundant secreted proteins from the salivary glands of the sand fly *Lutzomyia longipalpis*, vector of *Leishmania chagasi*. *J Exp Biol.* 2004; 207: 3717–3729. doi: [10.1242/jeb.01185](https://doi.org/10.1242/jeb.01185) PMID: [15371479](https://pubmed.ncbi.nlm.nih.gov/15371479/)
 11. Oliveira F, Kamhawi S, Seitz AE, Pham VM, Guigal PM, Fisher L, et al. From transcriptome to immune: Identification of DTH inducing proteins from a *Phlebotomus ariasi* salivary gland cDNA library. *Vaccine.* 2006; 24: 374–390. doi: [10.1016/j.vaccine.2005.07.085](https://doi.org/10.1016/j.vaccine.2005.07.085) PMID: [16154670](https://pubmed.ncbi.nlm.nih.gov/16154670/)
 12. Anderson JM, Oliveira F, Kamhawi S, Mans BJ, Reynoso D, Seitz AE, et al. Comparative salivary gland transcriptomics of sandfly vectors of visceral leishmaniasis. *BMC Genomics.* 2006; 7: 52. doi: [10.1186/1471-2164-7-52](https://doi.org/10.1186/1471-2164-7-52) PMID: [16539713](https://pubmed.ncbi.nlm.nih.gov/16539713/)
 13. Kato H, Anderson JM, Kamhawi S, Oliveira F, Lawyer PG, Pham VM, et al. High degree of conservancy among secreted salivary gland proteins from two geographically distant *Phlebotomus duboscqi* sandflies populations (Mali and Kenya). *BMC Genomics.* 2006; 7: 226. doi: [10.1186/1471-2164-7-226](https://doi.org/10.1186/1471-2164-7-226) PMID: [16952314](https://pubmed.ncbi.nlm.nih.gov/16952314/)
 14. Hostomska J, Volfova V, Mu J, Garfield M, Rohousova I, Volf P, et al. Analysis of salivary transcripts and antigens of the sand fly *Phlebotomus arabicus*. *BMC Genomics.* 2009; 10: 282. doi: [10.1186/1471-2164-10-282](https://doi.org/10.1186/1471-2164-10-282) PMID: [19555500](https://pubmed.ncbi.nlm.nih.gov/19555500/)
 15. Abdeladhim M, Jochim RC, Ben Ahmed M, Zhioua E, Chelbi I, Cherni S, et al. Updating the salivary gland transcriptome of *Phlebotomus papatasi* (Tunisian Strain): The search for sand fly-secreted immunogenic proteins for humans. *PLoS One.* 2012; 7: e47347. doi: [10.1371/journal.pone.0047347](https://doi.org/10.1371/journal.pone.0047347) PMID: [23139741](https://pubmed.ncbi.nlm.nih.gov/23139741/)
 16. Rohousova I, Subrahmanyam S, Volfova V, Mu J, Volf P, Valenzuela JG, et al. Salivary gland transcriptomes and proteomes of *Phlebotomus tobbi* and *Phlebotomus sergenti*, vectors of leishmaniasis. *PLoS Negl Trop Dis.* 2012; 6: e1660. doi: [10.1371/journal.pntd.0001660](https://doi.org/10.1371/journal.pntd.0001660) PMID: [22629480](https://pubmed.ncbi.nlm.nih.gov/22629480/)
 17. de Moura TR, Oliveira F, Carneiro MW, Miranda JC, Clarencio J, Barral-Netto M, et al. Functional transcriptomics of wild-caught *Lutzomyia intermedia* salivary glands: identification of a protective salivary protein against *Leishmania braziliensis* infection. *PLoS Negl Trop Dis.* 2013; 7: e2242. doi: [10.1371/journal.pntd.0002242](https://doi.org/10.1371/journal.pntd.0002242) PMID: [23717705](https://pubmed.ncbi.nlm.nih.gov/23717705/)
 18. Abrudan J, Ramalho-Ortigao M, O'Neil S, Stayback G, Wadsworth M, Bernard M, et al. The characterization of the *Phlebotomus papatasi* transcriptome. *Insect Mol Biol.* 2013; 22: 211–232. doi: [10.1111/imb.12015](https://doi.org/10.1111/imb.12015) PMID: [23398403](https://pubmed.ncbi.nlm.nih.gov/23398403/)
 19. Kato H, Jochim RC, Gomez EA, Uezato H, Mimori T, Korenaga M, et al. Analysis of salivary gland transcripts of the sand fly *Lutzomyia ayacuchensis*, a vector of Andean-type cutaneous leishmaniasis. *Infect Genet Evol.* 2013; 13: 56–66. doi: [10.1016/j.meegid.2012.08.024](https://doi.org/10.1016/j.meegid.2012.08.024) PMID: [23000112](https://pubmed.ncbi.nlm.nih.gov/23000112/)
 20. Vikova M, Sima M, Rohousova I, Kostalova T, Sumova P, Volfova V, et al. Comparative analysis of salivary gland transcriptomes of *Phlebotomus orientalis* sand flies from endemic and non-endemic foci of visceral leishmaniasis. *PLoS Negl Trop Dis.* 2014; 8: e2709. doi: [10.1371/journal.pntd.0002709](https://doi.org/10.1371/journal.pntd.0002709) PMID: [24587463](https://pubmed.ncbi.nlm.nih.gov/24587463/)
 21. Adeladhim M, Coutinho-Abreau IV, Townsend S, Pasos-Pinto S, Sanchez L, Rasouli M, et al. Molecular diversity between salivary proteins from New World and Old World sand flies with emphasis on *Bichromomyia olmeca*, the sand fly vector of *Leishmania mexicana* in Mesoamerica. *PLoS Negl Trop Dis.* 2016; 10: e0004771. doi: [10.1371/journal.pntd.0004771](https://doi.org/10.1371/journal.pntd.0004771) PMID: [27409591](https://pubmed.ncbi.nlm.nih.gov/27409591/)
 22. Nash WG, Kamerow HN, Merrill CR. Identification of a yellow gene-specific protein in *Drosophila melanogaster* by two-dimensional gel-electrophoresis. *Biochem Genet.* 1983; 21: 1135–1142. PMID: [6422926](https://pubmed.ncbi.nlm.nih.gov/6422926/)
 23. Hanes J, Simuth J. Identification and partial characterization of the major royal jelly protein of the honey-bee (*Apis mellifera* L.). *J Apic Res.* 1992; 31: 22–26.
 24. Johnson JK, Li J, Christensen BM. Cloning and characterization of a dopachrome conversion enzyme from the yellow fever mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol.* 2001; 31: 1125–1135. PMID: [11520691](https://pubmed.ncbi.nlm.nih.gov/11520691/)
 25. Alves-Silva J, Ribeiro JMC, Van Den Abbeele J, Attardo G, Hao R, Haines LR, et al. An insight into the sialome of *Glossina morsitans morsitans*. *BMC Genomics.* 2010; 11: e213.

26. Xu X, Oliveira F, Chang BW, Collin N, Gomes R, Teixeira C, et al. Structure and function of a “yellow” protein from saliva of the sand fly *Lutzomyia longipalpis* that confers protective immunity against *Leishmania major* infection. *J Biol Chem*. 2011; 286: 32383–32393. doi: [10.1074/jbc.M111.268904](https://doi.org/10.1074/jbc.M111.268904) PMID: [21795673](https://pubmed.ncbi.nlm.nih.gov/21795673/)
27. Oliveira MCG, Pelegrini-Da-Silva A, Parada CA, Tambeli CH. 5-HT acts on nociceptive primary afferents through an indirect mechanism to induce hyperalgesia in the subcutaneous tissue. *Neuroscience*. 2007; 145: 708–714. doi: [10.1016/j.neuroscience.2006.12.021](https://doi.org/10.1016/j.neuroscience.2006.12.021) PMID: [17257768](https://pubmed.ncbi.nlm.nih.gov/17257768/)
28. Xanthos DN, Bennett GJ, Coderre TJ. Norepinephrine-induced nociception and vasoconstrictor hypersensitivity in rats with chronic post-ischemia pain. *Pain*. 2008; 137: 640–651. doi: [10.1016/j.pain.2007.10.031](https://doi.org/10.1016/j.pain.2007.10.031) PMID: [18079061](https://pubmed.ncbi.nlm.nih.gov/18079061/)
29. Volf P, Skarupova S, Man P. Characterization of the lectin from females of *Phlebotomus duboscqi* sand flies. *Eur J Biochem*. 2002; 269: 6294–6301. PMID: [12473126](https://pubmed.ncbi.nlm.nih.gov/12473126/)
30. Teixeira C, Gomes R, Collin N, Reynoso D, Jochim RC, Oliveira F, et al. Discovery of markers of exposure specific to bites of *Lutzomyia longipalpis*, the vector of *Leishmania infantum* chagasi in Latin America. *PLoS Negl Trop Dis*. 2010; 4: e638. doi: [10.1371/journal.pntd.0000638](https://doi.org/10.1371/journal.pntd.0000638) PMID: [20351786](https://pubmed.ncbi.nlm.nih.gov/20351786/)
31. Souza AP, Andrade BB, Aquino D, Entringer P, Miranda JC, Alcantara R, et al. Using recombinant proteins from *Lutzomyia longipalpis* saliva to estimate human vector exposure in visceral leishmaniasis endemic areas. *PLoS Negl Trop Dis*. 2010; 4: e649. doi: [10.1371/journal.pntd.0000649](https://doi.org/10.1371/journal.pntd.0000649) PMID: [20351785](https://pubmed.ncbi.nlm.nih.gov/20351785/)
32. Martin-Martin I, Molina R, Rohousova I, Drahota J, Volf P, Jimenez M. High levels of anti-*Phlebotomus perniciosus* saliva antibodies in different vertebrate hosts from the re-emerging leishmaniasis focus in Madrid, Spain. *Vet Parasitol*. 2014; 202: 207–216. doi: [10.1016/j.vetpar.2014.02.045](https://doi.org/10.1016/j.vetpar.2014.02.045) PMID: [24629428](https://pubmed.ncbi.nlm.nih.gov/24629428/)
33. Kostalova T, Lestinova T, Sumova P, Vlkova M, Rohousova I, Berriatua E, et al. Canine antibodies against salivary recombinant proteins of *Phlebotomus perniciosus*: A longitudinal study in an endemic focus of canine leishmaniasis. *PLoS Negl Trop Dis*. 2015; 9: e0003855. doi: [10.1371/journal.pntd.0003855](https://doi.org/10.1371/journal.pntd.0003855) PMID: [26111018](https://pubmed.ncbi.nlm.nih.gov/26111018/)
34. Sima M, Ferencova B, Warburg A, Rohousova I, Volf P. Recombinant salivary proteins of *Phlebotomus orientalis* are suitable antigens to measure exposure of domestic animals to sand fly bites. *PLoS Negl Trop Dis*. 2016; 10: e0004553. doi: [10.1371/journal.pntd.0004553](https://doi.org/10.1371/journal.pntd.0004553) PMID: [26986566](https://pubmed.ncbi.nlm.nih.gov/26986566/)
35. Gomes R, Oliveira F, Teixeira C, Meneses C, Gilmore DC, Elnaiem DE, et al. Immunity to sand fly salivary protein LJM11 modulates host response to vector-transmitted *Leishmania* conferring ulcer-free protection. *J Invest Dermatol*. 2012; 132: 2735–2743. doi: [10.1038/jid.2012.205](https://doi.org/10.1038/jid.2012.205) PMID: [22739793](https://pubmed.ncbi.nlm.nih.gov/22739793/)
36. Collin N, Gomes R, Teixeira C, Cheng L, Laughinghouse A, Ward JM, et al. Sand fly salivary proteins induce strong cellular immunity in a natural reservoir of visceral leishmaniasis with adverse consequences for *Leishmania*. *PLoS Pathog*. 2009; 5: e1000441. doi: [10.1371/journal.ppat.1000441](https://doi.org/10.1371/journal.ppat.1000441) PMID: [19461875](https://pubmed.ncbi.nlm.nih.gov/19461875/)
37. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990; 215: 403–410. doi: [10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: [2231712](https://pubmed.ncbi.nlm.nih.gov/2231712/)
38. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*. 2001; 8: 785–786.
39. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and clustal X version 2.0. *Bioinformatics*. 2007; 23: 2947–2948. doi: [10.1093/bioinformatics/btm404](https://doi.org/10.1093/bioinformatics/btm404) PMID: [17846036](https://pubmed.ncbi.nlm.nih.gov/17846036/)
40. Abascal F, Zardoya R, Posada D. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics*. 2005; 21: 2104–2105. doi: [10.1093/bioinformatics/bti263](https://doi.org/10.1093/bioinformatics/bti263) PMID: [15647292](https://pubmed.ncbi.nlm.nih.gov/15647292/)
41. Schmidt HA, Strimmer K, Vingron M, von Haeseler A. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics*. 2002; 18: 502–504. PMID: [11934758](https://pubmed.ncbi.nlm.nih.gov/11934758/)
42. Goldman N, Whelan S. Statistical tests of gamma-distributed rate heterogeneity in models of sequence evolution in phylogenetics. *Mol Biol Evol*. 2000; 17: 975–978. PMID: [10833204](https://pubmed.ncbi.nlm.nih.gov/10833204/)
43. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol*. 2007; 24: 1596–1599. doi: [10.1093/molbev/msm092](https://doi.org/10.1093/molbev/msm092) PMID: [17488738](https://pubmed.ncbi.nlm.nih.gov/17488738/)
44. Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 2011; 7: 539. doi: [10.1038/msb.2011.75](https://doi.org/10.1038/msb.2011.75) PMID: [21988835](https://pubmed.ncbi.nlm.nih.gov/21988835/)
45. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J*. 2013; 32: 1478–1488. doi: [10.1038/emboj.2013.79](https://doi.org/10.1038/emboj.2013.79) PMID: [23584533](https://pubmed.ncbi.nlm.nih.gov/23584533/)

46. Julenius K. NetCGlyc 1.0: Prediction of mammalian C-mannosylation sites. *Glycobiology*. 2007; 17: 868–876. doi: [10.1093/glycob/cwm050](https://doi.org/10.1093/glycob/cwm050) PMID: [17494086](https://pubmed.ncbi.nlm.nih.gov/17494086/)
47. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. *Nucleic Acids Res*. 2000; 28: 235–242. PMID: [10592235](https://pubmed.ncbi.nlm.nih.gov/10592235/)
48. Sali A, Overington JP. Derivation of rules for comparative protein modeling from a database of protein-structure alignments. *Protein Sci*. 1994; 3: 1582–1596. doi: [10.1002/pro.5560030923](https://doi.org/10.1002/pro.5560030923) PMID: [7833817](https://pubmed.ncbi.nlm.nih.gov/7833817/)
49. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA*. 2001; 98: 10037–10041. doi: [10.1073/pnas.181342398](https://doi.org/10.1073/pnas.181342398) PMID: [11517324](https://pubmed.ncbi.nlm.nih.gov/11517324/)
50. Sehnal D, Svobodova Varekova R, Berka K, Pravda L, Navratilova V, et al. MOLE 2.0: advanced approach for analysis of biomacromolecular channels. *J Cheminform*. 2013; 5: 39. doi: [10.1186/1758-2946-5-39](https://doi.org/10.1186/1758-2946-5-39) PMID: [23953065](https://pubmed.ncbi.nlm.nih.gov/23953065/)
51. Cid H, Bunster M, Canales M, Gazitua F. Hydrophobicity and structural classes in proteins. *Protein Eng Des Sel*. 1992; 5: 373–375.
52. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: A sequence logo generator. *Genome Res*. 2004; 14: 1188–1190. doi: [10.1101/gr.849004](https://doi.org/10.1101/gr.849004) PMID: [15173120](https://pubmed.ncbi.nlm.nih.gov/15173120/)
53. Fiuza JA, Dey R, Davenport D, Abdeladhim M, Meneses C, Oliveira F, et al. Intradermal immunization of *Leishmania donovani* centrin knock-out parasites in combination with salivary protein LJM19 from sand fly vector induces a durable protective immune response in hamsters. *PLoS Negl Trop Dis*. 2016; 10: e0004322. doi: [10.1371/journal.pntd.0004322](https://doi.org/10.1371/journal.pntd.0004322) PMID: [26752686](https://pubmed.ncbi.nlm.nih.gov/26752686/)
54. Rohousova I, Volfova V, Nova S, Volf P. Individual variability of salivary gland proteins in three *Phlebotomus* species. *Acta Trop*. 2012; 122, 80–86. doi: [10.1016/j.actatropica.2011.12.004](https://doi.org/10.1016/j.actatropica.2011.12.004) PMID: [22197318](https://pubmed.ncbi.nlm.nih.gov/22197318/)
55. Aransay AM, Scoulica E, Tselentis Y, Ready PD. Phylogenetic relationships of Phlebotomine sandflies inferred from small subunit nuclear ribosomal DNA. *Insect Mol Biol*. 2000; 9: 157–168. PMID: [10762423](https://pubmed.ncbi.nlm.nih.gov/10762423/)
56. Hofinger ESA, Spickenreither M, Oschmann J, Bernhardt G, Rudolph R, Buschauer A. Recombinant human hyaluronidase Hyal-1: insect cells versus *Escherichia coli* as expression system and identification of low molecular weight inhibitors. *Glycobiology*. 2007; 17: 444–453. doi: [10.1093/glycob/cwm003](https://doi.org/10.1093/glycob/cwm003) PMID: [17227790](https://pubmed.ncbi.nlm.nih.gov/17227790/)
57. Reddy VB, Kounga K, Mariano F, Lerner EA. Chrysoptin is a potent glycoprotein IIb/IIIa fibrinogen receptor antagonist present in salivary gland extracts of the deerfly. *J Biol Chem*. 2000; 275: 15861–15867. PMID: [10821849](https://pubmed.ncbi.nlm.nih.gov/10821849/)
58. Atta A, Colossi R, Sousa-Atta MLB, Jeronimo SMB, Nascimento MDSB, Bezerra GF, et al. Antileishmanial IgG and IgE antibodies recognize predominantly carbohydrate epitopes of glycosylated antigens in visceral leishmaniasis. *Mem Inst Oswaldo Cruz*. 2004; 99: 525–530. PMID: [15543418](https://pubmed.ncbi.nlm.nih.gov/15543418/)
59. Volf P, Tesarova P, Nohynkova E. Salivary proteins and glycoproteins in phlebotomine sandflies of various species, sex and age. *Med Vet Entomol*. 2000; 14: 251–256. PMID: [11016431](https://pubmed.ncbi.nlm.nih.gov/11016431/)