Research Article

Identification of *Lutzomyia longipalpis* Odorant Binding Protein Modulators by Comparative Modeling, Hierarchical Virtual Screening, and Molecular Dynamics

Isis Bugia Santana,1 Franco Henrique Andrade Leite2,3 and Manoelito Coelho Santos Junior2,3

1Laboratório de Modelagem Molecular (LMM), Universidade Estadual de Feira de Santana, Feira de Santana-Ba 44031-460, Brazil
2Departamento de Saúde, Universidade Estadual de Feira de Santana, Feira de Santana-Ba 44031-460, Brazil
3Programa de Pós-Graduação em Biotecnologia, Universidade Estadual de Feira de Santana, Feira de Santana-Ba 44031-460, Brazil

Correspondence should be addressed to Franco Henrique Andrade Leite; fhenrique@uefs.br

Received 19 May 2018; Accepted 22 July 2018; Published 24 September 2018

Academic Editor: Fabio Polticelli

Visceral leishmaniasis (VL) is the second most important vector-borne disease in the world. It is transmitted by *Lutzomyia longipalpis* in America; therefore, controlling the vector is essential to prevent the disease, especially using traps with chemical attractants. It is known that odorant binding proteins (OBPs) act at the first odor selection level, so *in silico* methodology was used to identify putative vector chemical modulators based on OBPs on known ligand structures. Therefore, 3D structures of *L. longipalpis* OBP were predicted through different comparative modeling methods. The best model was subjected to molecular dynamics studies. Then, a hierarchical virtual screening approach filtered OBP modulator-like compounds from ZINC12 biogenic database based in global chemical space, using principal components from ChemGPS-NP server. Such compounds then were evaluated and ranked according to their affinity with the OBP orthosteric site by molecular docking in DOCK 6.7. The compounds were scored by Grid Score function and top five ranked poses had their intermolecular complex interactions analyzed in PLIP server. Most ligands in the top of the rank were lysophospholipids, which could potentially interact with the OBP hydrophobic pocket through Phe72, Tyr76, Ilec79, Ala87, Lys88, Asp92, Phe61, Leu75, Trp113, His120, and Phe122 residues and H-bonding with His120 and Phe122. Next, compounds in the top of the rank were evaluated by 50 ns MD and the results showed that the phosphate group of these compounds could set a salt bridge with His110. Additionally, Tyr76, Ala87, Met91, Trp113, and Phe122 were important to hydrophobic interactions with the ligand. These results highlight the importance of accurate assessments such as MD studies in order to analyze the docking results in the identification of new odorant modulators.

1. Introduction

Vector-borne diseases are one of the major and most concerning human epidemic issues to be solved. They affect more than half of the world population and cause more than a million deaths per year [1]. Visceral leishmaniasis (VL) is an important vector-borne disease manifested through systemic infections caused by parasites belonging to genus *Leishmania*. It is transmitted by the bites of female phlebotomine *Lutzomyia longipalpis*, only [2]. Estimates assume that 500,000 new people get VL in a yearly basis and that 60,000 of them die due to the disease [3, 4]. The main vector-control strategy applied to prevent and combat vector-borne diseases rely on using insecticides such as dichlorodiphenyltrichloroethane (DDT) [5]. However, damages to the environment and to human health are caused by DDT accumulation restraint due to the application of these insecticides [6–8].

Chemical ecology is an alternative approach to insecticide application, besides being a nonpolluting procedure. This
approach is based on volatile chemical compounds used to disturb and manage the chemical odorant sensing of insect vectors in order to modulate their behavior [9–11]. The odorant binding proteins (OBPs) are the first molecular filters favoring insects’ sensitivity because they act in odor selection [11–13]. These proteins are responsible for the selective capture and transport of small hydrophobic compounds (odors) into the sensillar lymph, for an aqueous solution that fills the sense organs into membrane olfactory receptors in the neuron cell and for carrying pheromones into insects glands [14, 15]. Therefore, OBPs structures can guide the identification of putative-behavior modulator compounds according to their affinity with the transporter binding site [16, 17].

Furthermore, there are some advantages on dealing with OBPs because they are extracellular soluble proteins, and their structure is lesser complex than that of olfactory receptors with seven transmembrane domains (7TM). So, these molecular targets have been used in a new method known as chemical reverse ecology since it uses the protein-target structure rather than the structure of known similar compounds to develop chemicals that are attractive and/or repellent to insects by using computational tools [16, 17].

Computational tools allow the fast and cheaper assessment to big virtual compound libraries and help identifying bioactive/modulative molecules. Overall, this virtual screening (VS) process may be guided by the structure of the protein binding site or by chemical similarity to the structures of known modulators [18]. Hierarchical virtual screening (HVS) methods use all the information about the structure (from protein and ligands) to strategically filter compound libraries.

Thereby, the main aim of the present study was to search for putative chemical modulators of Lutzomyia longipalpis behavior by using the structure of an OBP found in the transcriptome of the insect’s male sexual organs (RAAP-BAR014D07) and also in the cDNA of its antennae (BAM011G02) [19, 20]. This work rose from previous computational studies focused in L. longipalpis OBPs [21] in which 3D OBP comparative models were generated through different methods, and the best ones were chosen based on their overall quality (validation). Then, they were optimized through minimization procedures and subjected to molecular dynamics (MD) simulations. Finally, these comparative models were used to screen a virtual library of natural compounds, which was previously filtered by ligand-based virtual screening, i.e., a scheme composed of different hierarchical steps.

2. Materials and Methods

2.1. Comparative Modeling: Template Selection. The peptide sequence of Lutzomyia longipalpis OBP4 (RAAP-BAR014D07) was subjected to the SignalIP v. 4.1 server [22] in order to identify and further remove the signal peptide. Template selection was achieved by using a sequence of similar searches at PSI-BLAST [23] against sequences of 3D structures experimentally solved and deposited in the Protein Data Bank (PDB) repository [24]. Only proteins presenting aligned sequences with at least 45% identity, structure with 2 Å resolution, and 20% R-in complex with ligand (holo structure) were taken into consideration [25–27].

2.2. Comparative Modeling: Construction of L. longipalpis OBP 3D Model. The target-template sequence alignment was subjected to the SWISS-MODEL server (https://swissmodel.expasy.org/) [28], to the MODELLER software [29, 30], and to the I-TASSER server [31] (http://zhanglab.ucsb.edu/I-TASSER/) for structure prediction methods applied to rigid body assembly, spatial restraints, and threading, respectively.

2.3. Comparative Modeling: L. longipalpis OBP Model Validation. The OBPs models early developed from different prediction methods were visually examined in the UCSF Chimera 1.9.8 [30] and Discovery Studio Visualizer 4.1 software [32]. They had their steric quality estimated through Ramachandran plots generated in the PROCHECK v.3.5.4 software [33]. Energy quality examinations were conducted through ANOLEA in order to corroborate the overall quality of the models [34] (http://melalab.org/anolea/); global quality estimations were performed through QMEAN6 [35] and Z-SCORE [36]. The best-quality model was chosen to optimize the procedures through minimization cycles and molecular dynamics (MD) simulation of the model in the complex (with template bound ligand in holo crystal structure). It was done to reproduce solvation effects such as ligand-protein intermolecular interactions [37].

2.4. Molecular Dynamics Simulations. The calculations described below were performed in GROMACS v.5 with GROMOS 53a6 force field, in a NPT ensemble [38]. Firstly, the protonation states of early-model ionizable residues were calculated in PROPKA in order to mimic the pH (pH 7.5) of the sensillar lymph [39]. Next, the OBP model was complexed with the ligand found in the template structure, and the system was solvated in 7496 SPC-E water models [40] (without the addition of counterions) in a dodecahedron crystal cell of 71.92 Å lengths (A, B, and C) with each angle set at 60° (alpha, beta, and gamma). The protein was placed in the center of the box (13 Å in relation to box vectors), and the energy in the system was minimized through two steepest descent algorithm (SD) routines: first, all protein heavy atoms were restrained to their initial positions using 200 kJ/mol force and after that, without restraints. LINCS algorithm restricted all the atom bond movements in the protein [41]. Then, the system was further minimized through conjugated gradient (CG) algorithm, without any restraint. Verlet integrator [42] was used to force the system to reach the equilibrium when the system was heated from 0 to 300 K at a constant pressure through short MD simulation (1 ns). MD simulation procedures were performed within periodic boundary conditions assuring that the atoms would remain inside the simulation box. Long-range bonds (at maximum 13 Å) of each atom were taken into account,
and the electrostatic interactions were treated through the particle mesh ewald (PME) method [43]. An extensive MD simulation (100 ns) was conducted under the aforementioned conditions without any restraint.

The choice was made for a representative structure to be used in the structure-based virtual screening procedures (docking). It was done by analyzing frames from every 10 ps of simulation after the system reached stability and by subsequently selecting the central frame of the most populated cluster of system coordinates according to the RMSD values at cutoff 2 Å of RMSD for cluster differentiation based on the GROMOS method.

2.5. Ligand-Based Virtual Screening. Ligands required for ligand-based screening were selected from crystallographic holo structures (protein + ligand) of mosquito OBPs, because of their evolutionary and structural link to phlebotominae OBPs, since there is no phlebotomine OBP structure solved through experimental methods such as crystallography and NMR (nuclear magnetic resonance) deposited in the protein data bank. The biogenic (natural) chemical compound data bank of ZINC12 [44] was filtered according to its structural similarity to previously selected ligands. Therefore, similarity was based on the spatial chemical coordinates (from axes PC1, PC2, and PC3) collected in the CHEMGPS-NP Web [45] through Euclidean distance calculations. The spatial location of biogenic chemical structures was compared to the location of mosquito OBP ligands, but only the closest cluster to the OBP ligand structures was selected for the second screening filter known as molecular docking (receptor-based virtual screening).

2.6. Structure-Based Virtual Screening. Lutzomyia longipalpis OBP4 was the protein structure used in docking procedures; it was obtained through comparative modeling and optimized through MD calculations. The adopted chemical structure database was the one derived from a previously ligand-based filtered biogenic set of compounds. The protein orthosteric site was predicted through the superpositioning of template and homologue modeled structures in the Discovery Studio 4.5 [32]. The position of the backbone ligand complexed in the template structure was used to define a central point. Thus, the docking site was composed of amino acid residues located 8 Å from the central point. The protein structure was prepared in the UFSC CHIMERA 1.9.8 [30]. First, hydrogen atoms and the Gasteiger charge were added to the PDB protein file. A copy of this charged file (without the hydrogens) was used to find the molecular surface (dms file), which was used for sphere generation in all the protein cavities. The spheres within 8.0 Å of the ligand complexed structure were selected, and the docking calculation regions were delimited. Subsequently, grid generation was conducted by using the charged protein file and the selected spheres. The docking calculations were performed in the DOCK 6.7 package [46].

The binding energy of all docked compounds was scored through the Grid Score. The top virtual hit compounds were these compounds complex interactions calculated in the PLIP server (https://projects.biotec.tu-dresden.de/plip-web/plip), which is a fully automated protein-ligand interaction profiler [47]. PLIP generated 3D interaction maps depicting the OBP orthosteric site and the docked ligand structures presenting the best energy values.

2.7. MD Simulation for Complex. The top virtual hit compound classified after hierarchical virtual screening procedures was subjected to assess molecular dynamics studies with the Lutzomyia longipalpis OBP model in silico under the same conditions adopted to previous MD studies implemented for OBP optimization. All the complex interactions were analyzed by using 3D interaction maps generated in PLIP [47].

3. Results and Discussion

The local alignment (BLAST) of target sequences without signal peptide amino acids allowed choosing the Culex quinquefasciatus OBP1 holo structure complexed with an oviposition pheromone crystalized at 8.2 pH, (PDB ID 3OGN-resolution: 1.3 Å; R-Value Free: 0.174) as the best template for homology modeling. The alignment had 61% of overall sequence identity and 57% orthosteric site conservation at 98% alignment coverage and 3e− 49 e-value (Figure 1). These identity values (overall and binding site) corroborate the extremely well conserved secondary structure folding the OBP family; on the other hand, they highlight the specificity of the olfactory chemical sense modulation. Sequence identity and similarity levels in the target-template alignment were satisfactory for homology prediction, since it only presented a six-residue region poorly conserved in the alignment (in red) and one noncorresponding additional residue at the beginning of the target sequence (Figure 1). The OBP target had the cysteine signature of the protein family, as well as presented the same typical C-terminal ionizable residues (HIS110, ASP117, and HIS120) found in OBPs that capture and release ligands according to pH medium variations [48] (Figure 1). The different strategies used to generate comparative models provided three 3D OBP models. Energetic parameters of the I-Tasser model presented the best overall value depending on the quality of the taken measurements; however, it was the model recording less residues in favorable Ramachandran regions. The MODELLER model with the lowest % of high energy had the best Ramachandran result at one steric clash in a flexible zone of the structure away from the orthosteric site. The model deriving from the SWISS-MODEL had energy and steric values raging between the values recorded for the other two mentioned strategies; therefore, the model resulting from MODELLER due to the adoption of the restriction method allowed the best steric evaluation results: 96.4% of the residues in favorable regions of the Ramachandran plot. The global MODELLER measurement was also satisfactory, this model recorded normalized discrete optimized protein energy (zDOPE) score (−1.94), and it...
indicated that the atom pair distances in the model comply the distances recorded for a large sample of known protein structures; therefore, it was the model of choice (Table 1).

The structural alignment between the homology model and the template (3OGN) OBPs showed C-α root mean square deviation (RMSD) 0.179 Å, which was within the range of experimental models deriving from the same protein (0.5 Å), and it indicated the high quality of the early-built homology model [49]. The built model presented all the common characteristics of insect odorant binding proteins, namely, globular shape formed by secondary alpha-helices structures presenting a hydrophobic cavity and a tail at the c-terminal chain. This tail helps to enclose the ligand in the cavity (Figure 2).

The protonation prediction of OBP4 amino acid residues indicates that all acid residues and histidines were deprotonated at pH = 7.5 and that the basic residues were protonated. It is worth noticing that both tail termini histidines (HIS120 and HIS110) could be protonated in more acidic environments (at pH 6.5), and it reinforced the suggested mechanism of releasing odor through pH dropping close to the olfactory neuron membrane.

3.1. Model Refinement. The system solvated at the bound state (holo) was subjected to minimization runs and to 150 ns of MD simulation, but it did not require any charge to be neutralized. The RMSD trajectory of all 150 ns recorded standard deviation (4.9). The last 50 ns reached stability path at standard deviation 1.52 and mean value 3.7 Å (Figure 3).

The overall structure had good packing at the mean radius of gyration value 13.9 Å (standard deviation 1.11) when the stabilized period was taken into account, and it met the values found for others OBPs [50] (Figure 4). The overall energy had inconspicuous change, fact that confirmed a more stable state for the model. Furthermore, the residues presenting prominent RMS fluctuation (RMSF larger than 2 Å) were mainly located in flexible regions such as the coils and alpha-helices extremities, however, only in residues close to the cavity of the binding site (alpha-helices 5 and 6) (Figure 5).

Structure clustering within the trajectory was achieved according to the RMSD variation in the last 50 ns in 11 clusters. The largest cluster comprised 2332 structure frames from which the representative structure at 139.09 ns was retrieved. The optimized protein model selected for virtual screening showed satisfactory quality, it recorded 99.1% of residues in favored regions of the Ramachandran plot and only 0.9% (one residue) of residues in disallowed regions (Figure 6).

3.2. Ligand-Based Virtual Screening. A tiny fraction of all 120 million compounds available for virtual Screening [51] was selected, namely, 119,624 molecules of natural compounds from the biogenic zinc database [44]. Roughly 60,000 of these compounds were very close to the chemical space of typical OBP ligands (decanal, indol, PEG6, mosquito pheromone oviposition (MOP), and picaridin); therefore, they composed a database for the next step: the receptor-based virtual screening (Figure 7).

3.3. Structure-Based Virtual Screening. The docking results of ligands in the orthosteric OBP site were scored through the Grid Score and ranked by number; the top five compounds are shown in Table 2. All molecules were lysophospholipids and phospholipids missing one of their two O-acyl chains; they are commonly found in human metabolome since they are used for plant communication.

Structure clustering within the trajectory was achieved according to the RMSD variation in the last 50 ns in 11 clusters. The largest cluster comprised 2332 structure frames from which the representative structure at 139.09 ns was retrieved. The optimized protein model selected for virtual screening showed satisfactory quality, it recorded 99.1% of residues in favored regions of the Ramachandran plot and only 0.9% (one residue) of residues in disallowed regions (Figure 6).

3.2. Ligand-Based Virtual Screening. A tiny fraction of all 120 million compounds available for virtual Screening [51] was selected, namely, 119,624 molecules of natural compounds from the biogenic zinc database [44]. Roughly 60,000 of these compounds were very close to the chemical space of typical OBP ligands (decanal, indol, PEG6, mosquito pheromone oviposition (MOP), and picaridin); therefore, they composed a database for the next step: the receptor-based virtual screening (Figure 7).

3.3. Structure-Based Virtual Screening. The docking results of ligands in the orthosteric OBP site were scored through the Grid Score and ranked by number; the top five compounds are shown in Table 2. All molecules were lysophospholipids and phospholipids missing one of their two O-acyl chains; they are commonly found in human metabolome since they are used for plant communication.
residues (HIS120 and PHE120) could interact with h-donor ligand tails in h-bonds. Moreover, the terminal loop residue ALA124 was subtly facing the hydrophobic cavity in contrast to the 3OGN template, and it could allow greater anchoring for longer chain ligands.

A stabilized MD trajectory was reached when the Lutzomyia longipalpis OBP complex presenting top virtual hit compound (Z40165350) was evaluated (∼1.37 Å) (Figure 8(a)). A representative structure, which evidenced that the hydrogen bonds previously identified through molecular docking did not stayed in place, was chosen from this trajectory. However, the h-bond between the nitrogen of PHE122 and the oxygen of ligand appeared to be more consistent (Figure 8(b)). A salt bridge was also detected between HIS110 and the ligand phosphate group in addition to hydrophobic interactions between residues TYR76, ALA87, MET91, TRP113, and PHE122 and the aliphatic chain of ligand.

### Table 1: Overall evaluation of the 3D comparative models generated from three different methods.

<table>
<thead>
<tr>
<th>Comparative modeling software</th>
<th>% aminoacides in Ramachandran plot regions</th>
<th>ET (E/kT)</th>
<th>% aa (high energy)</th>
<th>Qmean6</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Favorable</td>
<td>Allowed</td>
<td>Gen. allowed</td>
<td>Nonallowed</td>
<td></td>
</tr>
<tr>
<td>SWISS-MODEL</td>
<td>91.9</td>
<td>8.1</td>
<td>0</td>
<td>0</td>
<td>−557</td>
</tr>
<tr>
<td>I-Tasser</td>
<td>87.4</td>
<td>12.6</td>
<td>0</td>
<td>0</td>
<td>−598</td>
</tr>
<tr>
<td>MODELLER</td>
<td><strong>96.4</strong></td>
<td>2.7</td>
<td>0</td>
<td>0.9</td>
<td>−457</td>
</tr>
</tbody>
</table>

**Figure 2:** The best 3D OBP model built from MODELLER.

**Figure 3:** Root mean square deviation (RMSD) of the 150 ns trajectory in MD simulation.
Figure 4: Gyration radius.

Figure 5: Root mean square fluctuation (RMSF) of productive phase from 100 ns to 150 ns of MD simulation.

Figure 6: Steric validation, Ramachandran plot, of the optimized OBP model hierarchical virtual screening (VS).
Figure 7: Ligand-based screening of odor-like compounds (blue circles: OBP ligands; green dots: biogenic compounds from zinc12 database; red dots: selected compounds similar to OBP ligands in the chemical space).

Table 2: Top virtual hits selected from structure-based virtual screening.

Top-ranked 1 (ZINC ID: 000040165350 PubChem annotation ID: 53480942) grid score energy: -98.73 KJ/mol

Top-ranked 2 (ZINC ID: 000040165373 PubChem annotation ID: 52925143) grid score energy: -97.82 KJ/mol
Understanding the molecular mechanisms responsible for the olfaction modulation in insects is a complex and long path capable of revolutionizing the manner we deal with pollinators, pests, and disease vectors in general. The effect of neglect and technological backwardness on the control of disease vectors has been devastating in recent years. Vector control is extremely difficult to perform when it comes to leishmaniasis; therefore, strategies focused on the olfactory modulation of the vector’s behavior could help developing new tools to prevent this disease.

The use of odor binding proteins (OBPs) as important molecular target in studies focused on vectors’ olfactory modulation has been gaining more space in chemical ecology. Some of these proteins have already been identified for Lutzomyia longipalpis; however, the three-dimensional structure of these molecules has never been investigated, and there is no information about their expression or interaction with ligands. Hence, this is the first study showing detailed aspects of the structural biology of L. longipalpis OBPs through comparative modeling techniques. The validation of the quality of the built models allowed inferring that MODELLER is very useful to predict 3D models of high-quality OBPs and made it possible prioritizing the best refinement model (OBP4), including the use of molecular dynamics simulations. The final structure, which is the most stable one and corresponds to the production of the MD trajectory, is expected to be the closest to the molecular bioactive conformation of OBP.
The use of the hierarchical approach applied to virtual screening has allowed systematically, efficiently, and roughly filtering 120,000 natural compounds by selecting compounds in the chemical space similar to the known OBP/filtering 120,000 natural compounds by selecting compounds in the chemical space similar to the known OBP

Data Availability

Data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors thank for academic support of Graduate Program in Biotechnology, Universidade Estadual de Feira de Santana, which allowed the development of “Comparative Modeling and Hierarchical Virtual Screening for Identification of OBPs Modulators of Lutzomyia Longipalpis” (dissertation) that includes the data shown on this manuscript.

References
