

Insect Biochemistry and Molecular Biology

2D finger-printing and molecular docking studies identified potent mosquito repellents targeting Odorant Binding Protein 1

--Manuscript Draft--

Manuscript Number:	IB-D-22-00330R1
Article Type:	Full Length Article
Keywords:	Mosquito repellent; Odorant-binding protein; OBP; Virtual screening; Molecular docking; Molecular Dynamics; Behavioral bioassay
Corresponding Author:	Spyros E Zographos, Dr National Hellenic Research Foundation Institute of Chemical Biology GREECE
First Author:	Panagiota G.V. Liggri, Dr
Order of Authors:	Panagiota G.V. Liggri, Dr Alfonso Pérez-Garrido, Dr Katerina E. Tsitsanou, Dr Kalarickal V Dileep, Dr Antonios Michaelakis, Dr Dimitrios P. Papachristos, Dr Horacio Pérez-Sánchez, Dr Spyros E Zographos, Dr
Abstract:	Abstract Mosquito Odorant-Binding Proteins (OBPs) involved in the initial steps of the olfactory signal transduction cascade, constitute molecular targets for the design of novel pest control agents. Herein, ten compounds known to be active against mosquitoes and/or displaying a binding affinity for AgamOBP1 were used as queries in an in silico 2D-similarity screening of over 96 million chemical samples. Further filtering of the acquired hits on the basis of toxicity, vapor pressure, and commercial availability resulted in 120 unique molecules that were subjected to molecular docking studies against OBP1. For seventeen potential OBP1-binders, the free energy of binding (FEB) and mode of interaction with the protein were further estimated by molecular docking leading to the selection of eight molecules exhibiting the highest similarity with their parental compounds and favorable energy values. The in vitro determination of their binding affinity to AgamOBP1 and the evaluation of their repellent activity against female mosquitoes revealed three molecules with enhanced repellent properties: (a) a novel DEET-like repellent with lower volatility and higher binding affinity for OBP1 than DEET, (b) a highly active repellent that represents a new scaffold and (c) a potent, highly volatile repellent that found to be a strong DEET-site binder.
Suggested Reviewers:	James Devillers, Dr Professor, Scientific Information Processing Centre j.devillers@ctis.fr https://www.researchgate.net/scientific-contributions/James-Devilleers-39617853 Offmann Bernard, Dr Professor, Nantes University Faculty of Sciences and Technologies bernard.offmann@univ-nantes.fr http://www.bernardoffmann.fr Ramanathan Sowdhamini, Dr Professor, Tata Institute of Fundamental Research Library mini@ncbs.res.in https://www.ncbs.res.in/faculty/mini

	Ana Claudia do Amaral Melo, Dr Professor, Federal University of Rio de Janeiro anamelo@iq.ufrj.br https://www.iq.ufrj.br/docentes_ver_antiga/ana-claudia-do-amaral-melo/
Response to Reviewers:	

2D finger-printing and molecular docking studies identified potent mosquito repellents targeting Odorant Binding Protein 1

Panagiota G.V. Liggri^{1,5}, Alfonso Pérez-Garrido², Katerina E. Tsitsanou¹, Kalarickal V. Dileep³ Antonios Michaelakis⁴, Dimitrios P. Papachristos⁴, Horacio Pérez-Sánchez^{2*}, Spyros E. Zographos^{1*}

¹Institute of Chemical Biology, National Hellenic Research Foundation, 48 Vassileos Constantinou Avenue, 11635 Athens, Greece

²Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), Universidad Católica de Murcia (UCAM), 30107 Spain

³Laboratory for Computational and Structural Biology, Jubilee Centre for Medical Research, Jubilee Mission Medical College and Research Institute, Thrissur, Kerala-680005, India

⁴Benaki Phytopathological Institute, Department of Entomology and Agricultural Zoology, 8 S Delta Str. 14561 Kifissia, Athens, Greece

⁵Department of Biochemistry and Biotechnology, University of Thessaly, Biopolis, 41500 Larissa, Greece

Running title: Potent mosquito repellents targeting AgamOBP1

Correspondence to: Horacio Pérez-Sánchez; e-mail: hperez@ucam.edu (Corresponding author for inquiries related to the computational studies) and Spyros E. Zographos; Institute of Chemical Biology, National Hellenic Research Foundation, 48, Vassileos Constantinou Avenue, 11635 Athens, Greece. Tel.: +302107273850, e-mail: sez@eie.gr

Abstract

Personal protection measures against the mosquitoes like the use of repellents constitute valuable tools in the effort to prevent the transmission of vector-borne diseases. Therefore, the discovery of novel repellent molecules which will be effective at lower concentrations and provide a longer duration of protection remains an urgent need.

Mosquito Odorant-Binding Proteins (OBPs) involved in the initial steps of the olfactory signal transduction cascade have been recognized not only as passive carriers of odors and pheromones but also as the first molecular filter to discriminate semiochemicals, hence serving as molecular targets for the design of novel pest control agents. Among the three-dimensional structures of mosquito OBPs solved in the last decades, the OBP1 complexes with known repellents have been widely used as reference structures in docking analysis and molecular dynamics simulation studies for the structure-based discovery of new molecules with repellent activity.

Herein, ten compounds known to be active against mosquitoes and/or displaying a binding affinity for *Anopheles gambiae* AgamOBP1 were used as queries in an *in silico* screening of over 96 million chemical samples in order to detect molecules with structural similarity. Further filtering of the acquired hits on the basis of toxicity, vapor pressure, and commercial availability resulted in 120 unique molecules that were subjected to molecular docking studies against OBP1. For seventeen potential OBP1-binders, the free energy of binding (FEB) and mode of interaction with the protein were further estimated by molecular docking simulations leading to the selection of eight molecules exhibiting the highest similarity with their parental compounds and favorable energy values. The *in vitro* determination of their binding affinity to AgamOBP1 and the evaluation of their repellent activity against female *Aedes albopictus* mosquitoes revealed that our combined ligand similarity screening and OBP1 structure-based molecular docking successfully detected three molecules with enhanced repellent properties. A novel DEET-like repellent with lower volatility (8.55×10^{-4} mmHg) but a higher binding affinity for OBP1 than DEET (1.35×10^{-3} mmHg). A highly active repellent molecule that is predicted to bind to the secondary Icaridin (sIC)-binding site of OBP1 with higher affinity than to the DEET-site and, therefore, represents a new scaffold to be exploited for the discovery of binders targeting multiple OBP sites. Finally, a third potent repellent exhibiting a high degree of volatility was found to be a strong DEET-site binder of OBP1 that could be used in slow-release formulations.

Keywords: Mosquito repellent, Odorant-binding protein, OBP, Virtual screening, molecular docking, Molecular Dynamics, Behavioral bioassay

Abbreviations: AgamOBP1, *Anopheles gambiae* odorant binding protein 1; OR, Odorant Receptor; DEET, N,N-diethyl-3-methylbenzamide; Icaridin, butan-2-yl 2-(2-hydroxyethyl)piperidine-1-carboxylate; MOP, (5R,6S)-6-acetoxy-5-hexadecanolide; 1-NPN, N-phenyl-1-naphthylamine; sIC-binding site; second Icaridin-binding site

1. Introduction

Several strategies have been developed for controlling mosquitoes including long-lasting insecticidal nets (LLINs) and indoor residual spraying (IRS) recommended by WHO (WHO, 2015). However, these methods do not provide protection in an outdoor environment, while their wide use has led to increased outdoor bite rates as a result of the reduced indoor biting and resting activity of mosquito populations (Keita et al., 2021; Meyers et al., 2016; Russell et al., 2011). As a consequence, personal protection measures against mosquitoes such as the application of repellents, constitute valuable tools in order to minimize human-mosquito contacts, hence preventing the transmission of vector-borne diseases. Therefore, the discovery of novel molecules which would protect from mosquito bites for prolonged periods and in lower concentrations is more than a necessity.

So far, several Quantitative structure-activity relationship (QSAR) and Artificial Neural Network (ANN) modeling studies have tried to correlate repellency with the structural and physicochemical characteristics of active compounds (Devillers, 2018; Katritzky et al., 2008; Oliferenko et al., 2013). Although they have provided key molecular features to be employed in the design of novel repellents, the obtained information on the interactions governing the recognition of repellents by their protein targets is relatively limited (Zographos et al., 2018).

The discovery of novel bioactive compounds is considered a remarkably challenging field as a small fraction of scaffolds that possess the desired physicochemical and biological properties should be retrieved from the vast chemical space estimated to comprise approximately 10^{60} organic molecules below 500 Da (Polishchuk et al., 2013), a number too large to be screened experimentally. Today several computational methods and cheminformatics tools help scientists not only to accelerate the discovery of drugs and other agents but also to overcome costly and time-consuming procedures. Depending on the information available on the target macromolecule, “ligand-based” and “structure-based” approaches are being increasingly employed for lead compound discovery. Pharmacophore modeling, virtual screening, molecular docking, and molecular dynamic simulations approaches are employed to identify promising parent structures. Prediction of protein-ligand and protein-protein interactions at the molecular level combined with structure-activity relationships (SAR) results provide valuable information for further lead optimization.

The success of any attempt to discover a bioactive compound largely depends on the selection of a suitable macromolecular target associated with the biochemical pathway of interest. In the quest for new molecules with repellent action against mosquitoes and other disease vectors, insects' olfactory system constitutes the target system of choice, and proteins involved in odor recognition are considered valuable molecular targets for the design of novel pest control agents. In mosquitoes, the perception of volatile semiochemicals is carried out in hair-like projections called sensilla, present on the surface of the antenna and the maxillary palps. Within the sensilla lymph, olfactory proteins such as odorant-binding proteins (OBPs), odorant receptors (ORs) and

ionotropic receptors (IRs), chemosensory proteins (CSPs), and sensory neuron membrane proteins (SNMPs) have been found to play essential roles in the chemoreception of odorants (Cassau and Krieger, 2021; Leal, 2013; Liu et al., 2020; Pelosi et al., 2018; Venthur and Zhou, 2018). The lack of three-dimensional structural information concerning OR-ligand complexes is the main obstacle to using them as targets in structure-based approaches. In contrast, OBPs are the most abundantly expressed genes products, present in large numbers in the lymph. Several studies carried out in various insect species have shown that these proteins do not only act as passive carriers of odorants and pheromones but are also involved in the discrimination of semiochemicals, functioning as the first molecular filter of the selection (Chang et al., 2015; Leal, 2005; Yin et al., 2015).

OBPs are low molecular weight proteins produced in large amounts in bacteria and can be highly purified, thus facilitating their crystal structural determination. So far, 66 crystal structures of insect OBPs in apo forms and complexes with a variety of ligands have been solved and deposited in the Protein data bank, contributing to the elucidation of the odor detection mechanism. Among them, 23 structures belong to OBPs of the mosquito species *Anopheles gambiae*, *Aedes aegypti*, *Culex quinquefasciatus*. Notably, the 3D structure of OBP1, a protein well conserved among these species, has been widely exploited in the last decade for the structure-based development of novel repellents. Structural studies have shown that AgamOBP1 from *An. gambiae* is able to bind the widely-used repellent DEET (Tsitsanou et al., 2012) and the natural repellent 6-methyl-5-heptene-2-one (6-MH) (Murphy et al., 2013) utilizing the same binding site, while it can accommodate two molecules of the known repellent Icaridin (Drakou et al., 2017). Exploration of those structures revealed unique key features such as the location, conformation, orientation, and the interactions of ligands with the protein. Moreover, molecular dynamics (MD) simulations of the AgamOBP1 complexes with DEET and 6-MH that were performed in order to investigate ligands' binding capacity, provided new insights into the molecular recognition of repellents by this protein, further assisting the generation of repellent pharmacophore models (Tzotzos et al., 2018). The AgamOBP1-DEET complex was used as a reference structure in docking analysis and MD simulations for the identification of plant-derived volatiles stabilized upon their binding to DEET's site (Affonso et al., 2013; Gaddaguti et al., 2016), as well as in the pharmacophoric and structural similarity search for synthetic molecules (Neto et al., 2020; Okoli et al., 2021) and natural compounds derived from essential oils (da Costa et al., 2019). Most recently, the AgamOBP1-Icaridin complex was utilized for the generation of an "olfactophore" model in order to be employed in the detection of novel scaffolds with repellent activity (Kritsi et al., 2022).

The use of OBP structures in homology modeling and molecular docking screening studies for the discovery of active semiochemicals has also been applied to other insect pests like the codling moth *Cydia pomonella* (L.) (Liu et al., 2016a), the oriental fruit fly *Bactrocera dorsalis* (Kamala Jayanthi et al., 2014), the Chilean scarab beetle *Hylamorphia elegans* (González-González et al., 2016), the polyphagous mirid bug *Adelphocoris lineolatus* (Sun et al., 2017), the leaf beetle

Ambrostoma quadriimpressum (Wang et al., 2017), and the jujube bud weevil *Pachyrhinus yasumatsui* (Hong et al., 2022).

In our previous study, a two-step virtual screening of a chemical library of 42,755 synthetic molecules (Key Organics; KO) was performed to discover novel bioinspired synthetic repellents that exhibit lower volatility and higher binding affinity for the OBP1 protein compared to their plant-derived parental compounds. Such molecules could provide enhanced mosquito repellent protection time for the user (Thireou et al., 2018). Continuing this work, six of the above KO compounds and four molecules of plant origin found to bind to AgamOBP1 according to fluorescence and NMR experiments (*Unpublished results*), were used as query ligands in an *in silico* screening study of a database containing over 96 million chemical samples in order to detect molecules with structural similarity. Subsequent filtering of the hits on the basis of toxicity and vapor pressure, and *in silico* evaluation of their binding ability to OBP1, resulted in seventeen potential OBP1-binders. For eight of them that exhibit the highest similarity with their parental compounds, binding free energy and mode of interaction with the OBP1 were estimated by MD calculations. Finally, seven out of the eight compounds that are commercially available were further evaluated for their *in vitro* binding affinity to AgamOBP1, and their repellent activity against female *Ae. albopictus* mosquitoes.

2. Materials and Methods

2.1. Creation of Library

Computational studies were performed to suggest novel mosquito repellents. Initially, the shape and pharmacophores of ten molecules that are known to be active against mosquitoes and/or to exhibit a binding affinity for AgamOBP1 (Ren et al., 2005; Thireou et al., 2018) were generated using ROCS (US EPA. [2017]) and LigandScout (Estrada, 1996) and chosen for our studies. Overall 96 million chemical samples that are distributed in two datasets, the Namiki Shoji (4.5 million molecules) and the Aldrich Market Select Chemical structure database (91.5 million molecules), were used for screening. For the Namiki Shoji dataset screening was performed using shape and pharmacophoric features. At the same time in order to screen the Aldrich Market Select Chemical structure database we used the iResearch™ System (<https://www.chemnavigator.com/cnc/products/iResearchSys.asp>). In both cases, molecules showing $\geq 70\%$ structural similarity to the ten known active leads were selected for our studies. Additional filtering based on the availability and cost of the obtained compounds was also performed.

2.2. Toxicity and Vapor Pressure Filtering

Removal of unwanted compounds at an early stage of processing makes the experimental or virtual screening significantly more efficient. Therefore, FILTER 2.1.1 (OpenEye Scientific Software,

Santa Fe, NM, USA) (Ren et al., 2005) was applied to reject molecules with undesirable features such as toxic moieties, a high likelihood of covalent binding to the target protein, and/or dyes and aggregators that may interfere with the experimental assay. Aggregators often lead to false positive results as these small-molecule aggregates can sequester some proteins, giving the appearance of being active inhibitors. Similarly, colored molecules such as dyes can appear to be inhibitors by interfering with colorimetric or fluorometric assays or due to non-specific binding to the target protein (see FILTER documentation https://docs.eyesopen.com/applications/omega/theory/filter_theory.html). Further, the vapor pressures of the selected compounds were estimated using the program MPBPWIN v.1.44 as implemented in the EPI suite (US EPA. [2017]). Based on the structure of a given chemical, the MPBPVP model estimates vapor pressure from various physico-chemical equations which use as input data, measured or calculated boiling points derived from group contribution QSAR methods. The final "suggested" vapor pressure estimation values were converted into volatility classes according to Spicer *et al.* 2002 (Spicer et al., 2002) : $VP < 10^{-7}$ mmHg = nonvolatile; VP between 10^{-7} and 10^{-1} mmHg = semi volatile; VP between 10^{-1} and 380 mmHg = volatile; $VP > 380$ mmHg = very volatile. Therefore, compounds with a vapor pressure of 7.5×10^{-7} mm Hg or greater, are more likely to be volatile. Based on that, molecules with estimated vapor pressures $\leq 7.5 \times 10^{-7}$ mmHg were excluded from the studies.

2.3. Molecular docking studies

The crystal structures of AgamOBP1 in complex with Icaridin (PDB ID: 5EL2) (Drakou et al., 2017) and DEET (PDB ID: 3N7H) (Tsitsanou et al., 2012) served as receptors in the current study. Both consist of two protein molecules in the asymmetric unit with a structural difference ≤ 0.3 Å. Two different binding sites (DEET-site and Icaridin (sIC)-site) were reported for AgamOBP1 (Drakou et al., 2017; Tsitsanou et al., 2012), however, in this study, we were only interested in the DEET binding site. While DEET binds in the same conformation to both subunits of the protein dimer, the Icaridin molecule adopts two different orientations at the DEET-site. Hence, we prepared grids (radius 10 Å) in these three different conformations (named Rec-1= 5EL2-A chain; Rec-2=5EL2-B chain, and Rec-3=3N7H-A chain; **Fig. S1**) and performed docking studies. Before setting up the grids, we prepared the selected protein molecules (i.e. monomers A and B for 5EL2 and monomer A of 3N7H) using the prep wizard of Schrödinger Maestro by deleting all crystallographic water molecules, adding hydrogen atoms to the polar groups and setting up the appropriate protonation states. An energy minimization with OPLS3 force field (Estrada, 1996) was also performed with an RMSD cutoff of 0.30 Å to refine the structure. We used an extra precision (XP) docking method and corresponding docking scores to screen the ligands. Molecules that were consistently scored at the top towards all three receptors were of our interest for molecular dynamics simulation (MD) studies.

2.4. Molecular dynamics simulations

MD simulations for 500 ns were performed using Desmond on the top 10% of the most stable AgamOBP1-ligand complexes to assess the binding stability of the proposed compounds. The binding energy was also calculated using Molecular Mechanics/Generalized Born Surface Area (MM-GBSA) from the trajectory. The AgamOBP1-ligand complexes were prepared using prep wizard as previously explained and soaked into an orthorhombic water box that contained TIP3P water molecules. The size of the box was set in such a way to include the entire protein-ligand complex. Also, the system was neutralized by applying the respective number of Na⁺ or Cl⁻ atoms. All MD simulations were performed using an NPT ensemble at default values of pressure and temperature, i.e. 1.01325 bar and 300 K respectively. We recorded the trajectory and energy at every 500 ps and 1.20 ps respectively, for 500 ns simulation.

2.5. Structural Interpretation

A QSAR model was developed in order to investigate the structural implications of the molecules in the activity under study. Although it is known that a low number of compounds doesn't let the generation of a robust QSAR model, this approach may provide a provisional idea of the sub-structural motifs that are related to the activity. For this purpose, the model was developed using a TOPological Substructural MOlecular DEscriptors (TOPSMODE) approach (Perez-Garrido et al., 2009). The selected descriptors have been used in various works as they are able to reveal relations between the activity and structure of chemical compounds (Perez-Garrido et al., 2014; Perez-Garrido et al., 2010; Perez-Garrido et al., 2012). These descriptors are weighted with several properties: standard bond distance (μ^{Std}); standard bond dipole moments (μ^{Dip} , $\mu^{\text{Dip}2}$); as well as contributions from the following atomic properties: hydrophobicity (μ^{Hyd}); polar surface area (μ^{Pols}); polarizability (μ^{Pol}); molar refractivity (μ^{Mol}); van der Waals radii (μ^{vdW}); Gasteiger–Marsilli charges (μ^{Gas}); atomic masses (μ^{Ato}); solute excess molar refraction ($\mu^{\text{Ab-R}2}$); solute dipolarity/polarizability ($\mu^{\text{Ab-}\pi\text{H}2}$); effective hydrogen-bond basicity ($\mu^{\text{Ab-}\Sigma\beta\text{o}}$, $\mu^{\text{Ab-}\Sigma\beta\text{H}}$) and solute gas hexadecane partition coefficient ($\mu^{\text{Ab-log L16}}$). We calculated a total 226 descriptors which were related to the activity value (**Table 1**) by Multilinear Regression analysis. The best model with no more than two descriptors was chosen based on a high value of both, coefficient of determination (R^2) and coefficient of determination of leave-one-out cross-validation (Q^2) and low values of Root Mean Squared Error (RMSE).

2.6. Bacterial expression and purification of recombinant AgamOBP1 protein

The pET22b(+)-AgamOBP1 expression plasmid was transformed into *E. coli* Origami B (DE3) competent cells (Novagen) and expressed as the full-length AgamOBP1 protein lacking its signal peptide (residues 20-144 of UniProt entry A0A1U7F4W2) as previously described (Drakou et al., 2017; Tsitsanou et al., 2012). The protein was purified by a series of chromatographic steps using the HiTrapQ FF and Resource Q ion-exchange columns (Cytiva) and the size-exclusion Superdex 75 column (Cytiva) according to the protocol described by Tsitsanou et al. (2012) (Tsitsanou et

al., 2012). The final protein fractions were dialyzed against 10 mM Tris-HCl (pH 8.0) buffer and concentrated to ~15 mg/ml.

2.7. Fluorescent measurements

The emission spectra of the fluorescence probe 1-NPN were recorded on a FluoroLog spectrofluorometer (HORIBA Scientific) at 25 °C, in a 90° right angle configuration, using a 1cm light path quartz cuvette. Slit widths were set at 5 nm for both excitation and emission. The AgamOBP1 protein was diluted in a buffer containing 20 mM Tris-HCl, pH 8.0, and 100 mM NaCl to a final volume of 1 ml. Fresh stock solutions of 1-NPN and the tested ligands were prepared in methanol or DMSO 100% and used for titration studies. The final solvent content was not greater than 1.5%. The probe was excited at 337 nm, and emission spectra were recorded from 350 to 500 nm. Maximal fluorescence intensities in the presence of AgamOBP1 were observed in the emission region of 402 - 410 nm after the subtraction of the background spectra (1-NPN and ligand in buffer without protein). Fluorescence data for 1-NPN and for each ligand were collected from two independent experiments (n = 2).

2.8. Kd determination of AgamOBP1-ligand complexes

The affinity of 1-NPN for AgamOBP1 was determined by titration of 2 µM protein solution with the fluorescent probe at concentrations ranging from 0 to 24 µM. The affinity of the examined ligands for AgamOBP1 was indirectly determined from the displacement of 1-NPN by the ligand. The AgamOBP1 was assayed at a final concentration of 2 µM and a fixed concentration of 1-NPN (10 µM). The probe was allowed to bind to AgamOBP1 prior to the addition of aliquots of each ligand in order to achieve a final concentration range from 0 to 50 µM. The Kd of 1-NPN and the Ki of ligands were calculated using the non-linear regression data analysis program GraphPad Prism 8 (www.graphpad.com).

The Kd of 1-NPN for AgamOBP1 was found to be $4.29 \mu\text{M} \pm 0.94$ from a saturation curve derived by fitting the data ($R^2=0.97$) to a one-site binding equation,

$$Y = B_{\text{max}} * X / (K_d + X) + NS * X + \text{Background},$$

assuming specific binding ($NS=0$). In the case of the tested ligands, their Ki values were calculated from competition dose-response curves for binding to one or two sites. For the one site binding model, the data were fitted to the equation

$$Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + 10^{(X - \log EC_{50})}),$$

where $\log EC_{50} = \log(10^{\log Ki * (1 + \text{Hot} \mu\text{M} / \text{Hot} K_d \mu\text{M})})$, using as $\text{Hot} \mu\text{M}$ the concentration of 1-NPN in the competition assays (10 µM) and as $\text{Hot} K_d \mu\text{M}$ the Kd of 1-NPN (4.29 µM). Respectively, for the two-site binding model, a modified equation (Two sites - Fit Ki) was used to determine the Ki values of the ligands for each site.

2.9. Rearing of *Aedes albopictus* in the laboratory

The present studies focused on *An. gambiae* AgamOBP1 whose structure and interactions with known repellent molecules are well known. Nevertheless, the discovered compounds were evaluated for their repellent properties against the Asian tiger mosquito *Ae. albopictus*. AgamOBP1 (Uniprot: A0A1U7F4W2) shares a high degree of sequence identity (>80%) with its homologous proteins *Ae. aegypti* AaegOBP1 (Uniprot: Q6Y2R8), *Cx. quinquefasciatus* CquiOBP1 (Uniprot: Q8T6I2) and AalbOBP39 (Uniprot: M4W733) from *Ae. albopictus* (Deng et al., 2013), with the exception of their N-terminal signal sequences. After the removal of signal peptides, the mature AgamOBP1 and AalbOBP39 proteins have 82% identity and a sequence similarity of 92%. The critical residues involved in hydrogen bonds and vdW interactions with DEET and Icaridin in both DEET and sIC-binding sites of AgamOBP1 are conserved in AalbOBP39, while Met89 and Val125 are substituted by a Leucine and an Isoleucine, respectively, with very similar properties (**Fig. S2**). The major difference between the ligand-interacting aminoacids of the two proteins appears in Ser79 of AgamOBP1. However, its contribution to Icaridin's binding to the sIC-binding site is probably minor as it participates in only 2 out of the 27 vdW contacts (Drakou et al., 2017). Therefore, binders of AgamOBP1 displaying repellent activity are expected to have the same effect on *A. albopictus* mosquitoes as well as on other mosquito species. The sequence alignment was performed by Clustal Omega online software and alignment plot was created by ESPript v3.0 (Robert and Gouet, 2014).

Adults of *Ae. albopictus* mosquitoes were obtained from a laboratory colony which was maintained in the Benaki Phytopathological Institute (Kifissia, Greece) at 25±2°C, 80% relative humidity, and photoperiod of 16:8-h light/dark (L/D) (Giatropoulos et al., 2015; Kythreoti et al., 2021). Plastic beakers with 100 ml of water and strips of moistened filter paper were provided in the cage for oviposition. The eggs were kept moist for a few days and then placed in the pans for hatching. The larvae were reared in tap water-filled cylindrical enamel pans, about 400 larvae in each pan, and fed in excess with powdered fish food (JBL Novo Tom 10 % Artemia) until the emergence of adults. Adult mosquitoes were often collected with a mouth aspirator and transferred to the rearing cage. Females were fed with fresh chicken blood using the Hemotek® blood-feeding system.

2.10. Repellent activity bioassay

The assessment of the repellent activity of the tested compounds was based on the human landing counts (Giatropoulos et al., 2012). The study was conducted into a cage (33×33×33 cm) with a 32×32 mesh containing 100 adult mosquitoes (sex ratio 1:1) 5 to 10-days old, starved for 12 h, at 25±2°C, and 70–80% relative humidity. Control treatments without the components or with DEET were also included in the repellency tests as standards (control and positive control, respectively). Each treatment was repeated eight times and four human volunteers were used.

2.11. Determination of repellency index

The landing numbers were converted to repellency index (RI $\pm$ SE) using the following equation

$$RI = 1 - (T/C) = (C - T) / C,$$

where C is the number of landings in the control and T the number of landings in the compound treatment.

Statistical analysis

The effects of tested compounds on the *Ae. albopictus* female landings were estimated using the Kruskal–Wallis test (Kruskal and Wallis, 1952). When significant differences were detected, Mann–Whitney U tests with Bonferroni correction were carried out for pair-wise among treatments (Mann and Whitney, 1947).

3. Results and Discussion

3.1. Computational screening identified putative disruptors of mosquito host-seeking behavior

The shape, pharmacophores, and 2D-fingerprints of ten compounds (**Fig. 1A-J** and **Table S1**) known to be active against mosquitoes and/or displaying a binding affinity for AgamOBP1 were used as queries to screen the Namiki Shoji and Aldrich Market Select Chemical structure databases. Among these, six were KO compounds that have been previously identified as AgamOBP1 binders by *in vitro* fluorescence competition experiments and shown to possess repellent activity against *An. gambiae* mosquitoes, with the exception of the strongest but non-volatile binder KO8 (Thireou et al., 2018). The set of leads was completed with four plant-derived compounds of unknown repellent activity, terpinen-4-ol, chrysanthenyl acetate, yomogi alcohol, and artemisia alcohol, that was found to bind to AgamOBP1 according to fluorescence binding assays, and ^1H STD and 2D ^1H - ^{15}N HSQC NMR studies (Dima, 2015).

Most of the query molecules are hydrophobic in nature. Compounds such as yomogi alcohol, artemisia alcohol, and terpinen-4-ol contain one hydrogen bond donor attached to their hydrophobic core, while a hydrogen bond acceptor is present in chrysanthenyl acetate. The molecules KO2, KO3, KO4, KO8, KO9, and KO11 are structurally similar, containing an aromatic ring and a hydrogen bond acceptor. In addition, in compounds KO8 and KO9, an acidic group (marked as -ve ionizable area) is also included in their structures (**Fig. 1A-J**).

Our chemical substructure and similarity screening, and the subsequent elimination of duplicates helped us to select 152 unique molecules out of the 184 hits that are structurally similar (similarity cutoff $\geq 70\%$) to the lead compounds (**Fig. 1K**). Due to budget constraints, we restricted to molecules that were commercially available in at least 10 mg amount and cost less than 100€. Queries KO3 and KO4 retrieved most of the outputs i.e. 44 and 43 molecules respectively, while yomogi alcohol, artemisia alcohol, terpinen-4-ol, and KO11 ≤ 9 molecules (**Table S1**). Toxicity-

based filtering led to the elimination of another 32 molecules with undesired properties such as toxic functionalities, a high likelihood of covalent binding to the target protein, and interference with the experimental assay (see Materials & Methods subsection 2.2). Furthermore, we estimated the vapor pressures of the molecules, and those with $\leq 7.5 \times 10^{-7}$ mmHg (16 molecules) were excluded from the studies (Spicer et al., 2002). Finally, a dataset of 120 molecules (**Table S2**) was subjected to molecular docking towards the three different receptors (Rec-1 to -3).

3.2. Putative mosquito host-seeking disruptors have energetically favorable interactions with AgamOBP1

Hits from the computational screening were subjected to the XP docking method against the three receptors. The top 30 % of the highest scoring (docking score ≥ -7.00 kcal/mol) molecules towards each receptor (i.e. 31 molecules of each) was taken for further analysis (**Fig. 1L**). 17 molecules (**Table S4**) that were consistently scored high towards all three receptors were of our interest, as we hypothesized that they may exhibit a potential binding affinity for AgamOBP1 *in vitro*. These ligands were subsequently subjected to MM-GBSA analysis based on the docked poses. Their calculated binding energies for the receptors were in the range of -29 to -90 kcal/mol (**Fig. 2A**). Out of the 17 molecules (ligand numbering is according to the main list that contains 120 molecules), Lig-86 exhibits a high structural similarity to the KO3 compound. Likewise, Lig-11, Lig-13, Lig-16, and Lig-103, Lig-111, Lig-118, Lig-119 are structurally similar to KO11 and chrysanthenyl acetate, respectively. Their high structural resemblance to the bioactive query molecules and their favorable binding energies towards AgamOBP1 prompt us to choose these eight ligands (i.e. Lig-11, Lig-13, Lig-16, Lig-86, Lig-103, Lig-111, Lig-118, and Lig-119) (**Table 2**) for further studies. Their binding modes at the DEET-binding site are shown in **Fig. 2 (B to I)**. We found that the MM-GBSA binding free energies obtained for Lig-119, Lig-111, and Lig-16 were higher than -80 kcal/mol placing them within the top 3 positions among the 17 molecules. Although the binding energy of Lig-86 (-29 kcal/mol) is lower than that of the other ligands, we included it in our MD simulations and further *in vitro* studies due to its high structural similarity with the query molecule KO3. Top binding energies predicted for the selected ligands are as follows; Lig-11 (-66 kcal/mol); Lig-13 (-74 kcal/mol); Lig-16 (-86 kcal/mol); Lig-86 (-29 kcal/mol); Lig-103 (-72 kcal/mol); Lig-111 (-87 kcal/mol); Lig-118 (-83 kcal/mol) and Lig-119 (-89 kcal/mol).

3.3. MD simulations provide additional insight towards the putative mosquito olfactory disruptors' binding to the DEET-site

The MD simulation studies for 500 ns suggested that all tested ligands except Lig-86 are stable in the DEET-binding site without much displacement from their initial docked pose, which further justifies our selection of ligands. RMSD of the selected eight protein-ligand complexes is shown in **Fig. S3**. We found that Lig-86 deviated 3.8 Å from its initial docked position and tended to shift out of the binding cavity. Though a hydrogen bond between its carboxyl group and the aromatic side chain of Trp114 helped initially the benzyl moiety to orient in such a way in order to favor

the development of weak π - π stacking interactions with this residue, these interactions were completely abolished during the MD simulations. Since the DEET-binding site mainly consists of hydrophobic residues, we assume that the carboxyl group of Lig-86 may participate in unfavorable interactions within this site and therefore tends to be displaced out of the binding cavity of DEET. The pattern of MM-GBSA values calculated from the trajectory and docked poses are in perfect agreement with each other. Among all ligands Lig-86 scored the lowest (-39 ± 2.53 kcal/mol). Conversely, Lig-111, 118 and 119 scored the most i.e. -92 ± 7.12 , -89 ± 4.20 and -97 ± 5.00 kcal/mol respectively. The MM-GBSA values of all ligands were shown in **Fig. S4**.

3.4. The binding modes of the selected host-seeking disruptors

Since molecular docking and subsequent MD simulations suggested a possible binding for the 17 selected ligands (**Table S3**) at the DEET binding site, we further critically investigated the binding modes of the eight top-scored molecules (Lig-11, 13, 16, 86, 103, 111, 118 and 119). The DEET binding site is formed by several non-polar residues such as Leu73, Leu76, His77, Leu80 from $\alpha 4$ helix; Ala88, Met89, Met91, Gly92 from $\alpha 5$; Trp114 and Leu96 from $\alpha 6$ as well as Lys93', Arg94' and Leu96' from the other subunit (**Fig. S1**). Similar to the DEET molecule, all the investigated ligands mainly participate in van der Waals interactions at the binding site. However, the carbonyl oxygen groups of Lig-16 and Lig-119 are also involved in the formation of hydrogen bonds by accepting a hydrogen atom from the NE1 atom of Trp114 and the ND1 atom of His77 residue, respectively (**Fig. 2**). We observed that during MD simulations, these hydrogen bonds were retained in more than 90 % of the intermediate frames. In contrast, no hydrogen bond with protein residues was observed in the cases of Lig-111 and Lig-118 upon their binding to AgamOBP1. In the docked poses, the oxygen atom of Lig-111 is 3.9 Å away from the NE1 of Trp114 and therefore, unable to form a hydrogen bond as Lig-16 does. At the same time, the oxygen of Lig-118 is pointed towards the aromatic side chain of Trp114. We noticed that the hydrophobic ends of most of the top-scored ligands enter deep into the hydrophobic channel. Upon forming a complex with AgamOBP1, the solvent-accessible surface area (SASA) for these molecules is reduced, and the buried surface area for Lig-16, Lig-111, Lig-118, and Lig-119 was calculated to be 95%, 82%, 93%, and 91%, respectively. In the study by Tsitsanou et al (2012) (Tsitsanou et al., 2012), it was reported a 90% decrease in the SASA of DEET upon its binding. It has also been suggested that the hydrogen bond formed between Trp114 and DEET via a water molecule is critical for DEET recognition. The binding energies of Icaridin and DEET molecules are -64 and -76 kcal/mol, respectively, suggesting a weaker binding affinity for AgamOBP1 than that of the top-scored ligands (binding energies for Lig-16: -86 kcal/mol, Lig-111: -87 kcal/mol, and Lig-119: -89 kcal/mol).

As was revealed from the AgamOBP1 complex structures with DEET and Icaridin, both ligands form one direct hydrogen bond with a water molecule which in turn interacts with the NE1 nitrogen of Thr114 and the backbone carbonyl of Cys95 (Drakou et al., 2017; Tsitsanou et al., 2012) (**Fig. S1**). This conserved water is proposed by the authors to play an important role in ligand recognition and should be taken into account during design processes (Anderson, 2003; Nandi et al., 2009; Neidle, 2021). Novel ligands that make additional contacts with the water or incorporate it into

their structures may exhibit an improved binding affinity due to the increased number of direct protein-ligand interactions. Since we didn't include this invariant water molecule in our docking calculations, we further investigated whether or not this water is replaced by the investigated ligands. To do this, we superimposed our docked poses on the DEET and Icaridin bound structures. We found that Lig-16 and Lig-86 could form hydrogen bonds with the structural water molecule. On the contrary, the orientation of the Lig-13 is such that it could displace the water molecule from the binding site. Concerning the other ligands, Lig-103, Lig-111, Lig-118, and Lig-119, they do not appear able to make any interactions with the conserved water (**Fig. S5 A to H**).

3.5. Structural Interpretation

The best QSAR model was obtained for activity values with 20 $\mu\text{g}/\text{cm}^2$. Following, we show the best QSAR model equation with standardized coefficient and statistical parameters (**Table 1**):

Table 1 comes about here

(**Table 1**: Statistical parameters of the QSAR model development)

We extracted the bond contribution (**Fig. S6**) based on the model coefficient and the results showed positive contributions of an ester moiety, a well-known group present in many repellent substances (Gouck et al., 1957; McCabe et al., 1954). Furthermore, a negative influence is observed in the areas around these ester groups and in the central area of the molecules, indicating that an increase in the volume of the molecule is accompanied by a decrease in the repellent activity of the compound. Although the model is not free from chance correlation, the interactions obtained by docking present some coincidences with the contributions extracted from the QSAR model in Lig-13, Lig-119, and Lig-16 (**Fig. S6**). Lig-13 presents several hydrophobic interactions (Ala88, Leu80 and Leu76) that agree with the contributions extracted from the QSAR model developed with the TOPSMODE descriptors. We can see something similar in Lig-119 with 3 hydrophobic interactions (Leu80 and two with Trp114) and Lig-16 with one hydrogen bond (Trp114) and one hydrophobic (Trp114) interaction. The negative effect of molecular volume in terms of molecular refractivity was also observed in other mosquito species (Suryanarayana et al., 1991). A direct correlation between vapor pressure (VP) values (or $-\log(\text{VP})$) and repellent activity has not been found in these compounds for any of the three concentrations.

Fig. 1 comes about here

Fig. 2 comes about here

Table 2 comes about here

Table 2. The final hits selected for MD simulations and *in vitro* studies. 2D structures of the hits and the parental.

3.6. *In vitro* binding affinity determination of the selected docking hits for AgamOBP1

With the exception of compound Lig-11 which was not commercially available, the binding capacity of the other seven ligands for AgamOBP1 has been evaluated by a fluorescence competition assay. Lig-86 and Lig-118 showed no binding affinity for the protein. Both ligands were not able to displace the fluorescent probe 1-NPN resulting in only 2% fluorescence reduction at 50 μM of ligand concentration. The remaining five ligands display a high affinity for AgamOBP1, exhibiting dissociation constants (K_i) between 0.04 and 1.25 μM (**Fig. 3** and **Table 3**). These results indicate that our molecular docking modeling method was quite successful in proposing strong binders of AgamOBP1.

The fluorescence analysis suggested that Lig-16 and Lig-13 bind to two protein sites, a high affinity binding site (0.04 and 0.08 μM , respectively) and a second of lower affinity (7.7 and 12.8 μM , respectively) (**Fig. 3** and **Table 3**). Both ligands show structural similarity to the KO11 lead compound bearing a furan aromatic ring and an ethyl acrylate group. Interestingly, in our previous modeling study, we predicted that KO11 may bind to both DEET and sIC-binding sites. KO11 was predicted to occupy the position of the structural water thus displacing it from the DEET-binding site. While this change has been found to be associated with reduced repellency in the cases of other compounds (Zographos et al., 2018), KO11 exhibits a significant repellent activity possibly due to its additional binding to the sIC-binding site, where it interacts with the N-terminal Val125 similar to Icaridin (Thireou et al., 2018) (**Fig. S1**). Likewise, we calculated the binding energies of these two ligands for the sIC-binding site. The MM-GBSA values for Lig-16 (-98 kcal/mol) and Lig-13 (-60 kcal/mol) at the sIC-site suggest that Lig-16 has a higher affinity for the sIC rather than the DEET-site (-86 kcal/mol). Conversely, in the case of Lig-13, the DEET-site is the primary site of binding (-74 kcal/mol). The fluorescence analysis cannot discriminate between the low and the high affinity binding sites. Nevertheless, by combining fluorescence data with the MM-GBSA binding energy predictions we can propose that Lig-16 binds preferentially to the sIC-site with high affinity (0.04 μM), whereas the high affinity site for Lig-13 corresponds to the DEET-binding site (0.08 μM). Although the sIC-binding site was not exploited in the molecular docking calculations of this study, the binding of Lig-13 and Lig-16 to this site could be a generalized feature of this class of furan compounds that is worth further investigation, as it may provide an alternative new strategy for future OBP1-based screening efforts.

Furthermore, based on our predictions, upon their binding to the DEET-binding site (Rec-1 to 3), the furan moieties of these ligands are placed closer to the bend of the L-shaped tunnel of the monomer (**Fig. 2C** and **2E**) by adopting different orientations. In particular, Lig-16 places its furan oxygen and bromo group at the entrance of a previously defined complementary pocket, which in the CquiOBP1-MOP complex (PDB id: 3OGN) accommodates the acetyl ester branch of MOP pheromone (Mao et al., 2010). Noticeably, we have previously proposed that this empty cavity lined by Phe123, Ile87, Met91, Met84, and Tyr10 residues could be exploited in the design of novel ligands (Tsitsanou et al., 2012). In the case of Lig-13, its furan ring lies almost in the same plane as the one of Lig-16 at an angle of approximately 180 degrees, orienting its furan ring oxygen

atom towards the $\alpha 4$ helix (**Fig. 2E**). The acrylate group of both Lig-13 and Lig-16 is involved in a direct hydrogen bond with the side chain of Trp114. Also, Lig-13, through its second acrylate moiety, participates in an additional hydrogen bond with the backbone of His77 as Icaridin does upon binding to the DEET-binding site (**Fig. S1**). Nevertheless, the binding modes of these ligands indicate that only Lig-13 is able to displace the structural water molecule from its position in the DEET-binding site. However, the observed interaction of the known DEET-site binders with Trp114 through this water molecule is not disrupted in the case of Lig-13 but replaced by the direct hydrogen bond.

Compounds Lig-119 and Lig-103 bearing a methyl acetate-substituted oxabicyclo [3.3.1] nonene system exhibit a structural resemblance to the chrysanthenyl acetate, though they differ in the number of ring atoms (nine atoms instead of seven), the methyl acetate group (instead of acetate), and the positions of methyl (methyl and isopropyl for Lig-103) moieties. In contrast to Lig-16 and Lig-13, fluorescent data analysis indicated binding to a single site, with Lig-119 and Lig-103 displaying the highest (0.042 μM) and the lowest (1.25 μM) affinity for the protein, respectively, among all the tested compounds (**Fig. 3** and **Table 3**). Both ligands are predicted to bind to the bend of the L-shaped channel (**Fig. 2H** and **2I**) by adopting, however, different conformations. While Lig-119 positions its oxane oxygen at the entrance of the complementary pocket (*vide supra*), Lig-103 orients this atom in the opposite direction towards the $\alpha 4$ helix, in order to bring its methyl cyclohex-1-ene group in front of the pocket's entrance (**Fig. 2I**). Furthermore, by superimposing the binding poses of the two ligands, it was observed that there is a $\sim 3\text{\AA}$ difference between their oxane oxygen atoms. Most likely, Lig-103 adopts this configuration in order for its isopropyl group to be accommodated better within the binding cavity. The different interactions of both the oxane oxygen and methyl-acetate group of the above ligands with the protein residues may explain the lower affinity of Lig-103 compared to Lig-119.

Finally, Lig-118 and Lig-111 are monocyclic compounds having a methylcyclohex-1-ene ring conjugated to a butyrate or isobutyrate group, respectively. Interestingly, despite their structural similarity, Lig-111 showed high affinity for AgamOBP1 (0.06 μM), while the butyrate analog Lig-118 did not displace the 1-NPN probe at any of the tested concentrations (**Fig. 3** and **Table 3**). Upon further analysis of their binding conformations, it has been found that these ligands do not displace the structural water from its location in the DEET-binding site. Although both ligands share similar binding modes, the strain energy for Lig-111 (1 kcal/mol) is much lesser than for Lig-118 (10 kcal/mol). In addition, differences in the ligands' binding energies were also observed. The optimum internal strain energy is crucial for the favorable binding of ligands. The high strain energy observed for Lig-118 could be reflected in the poor *in vitro* binding capability of the ligand.

Fig. 3 comes about here

3.7 Dose dependent effect of the selected host-seeking disruptors on *Aedes albopictus* mosquitoes.

Six out of the eight ligands identified by the molecular docking calculations (Lig-16, Lig-13, Lig-118, Lig-111, Lig-119 and Lig-103) were tested in behavioral bioassays against *Ae. albopictus*, an aggressive mosquito vector. We were not able to test Lig-86 because it was completely insoluble in bioassay compatible solvents such as DCM or ethanol, whereas Lig-11 was not commercially available.

We found that female landings were significantly affected by Lig-119 in all four doses evaluated for repellency (Dose 40 $\mu\text{g}/\text{cm}^2$: Chi-squared= 50.776, df= 2, $p<0.0001$; Dose 20 $\mu\text{g}/\text{cm}^2$: Chi-squared= 52.394, df= 6, $p<0.0001$; Dose 10 $\mu\text{g}/\text{cm}^2$: Chi-squared= 50.981, df= 5, $p<0.0001$; Dose 4 $\mu\text{g}/\text{cm}^2$: Chi-squared= 44.861, df= 6, $p<0.0001$) (**Fig. 4** and **Table 3**).

At doses between 40 and 10 $\mu\text{g}/\text{cm}^2$, Lig-119 displayed landing inhibition equal to or superior to that of DEET (100% to 96.6%), while it showed strong repellency though slightly weaker than the one of DEET (81.8% versus 92.2% of DEET) at the lowest test dose of 4 $\mu\text{g}/\text{cm}^2$.

The second strong repellent among the compounds was Lig-13 which provided full protection at the higher doses of 40 and 20 $\mu\text{g}/\text{cm}^2$ (landing inhibition 100%) and a considerable reduction in landings at 10 $\mu\text{g}/\text{cm}^2$ (81.5%), whereas at the lowest tested dose of 4 $\mu\text{g}/\text{cm}^2$ it retained a weak activity (33%).

Lig-111 significantly affected the number of landings at doses of 40 to 10 $\mu\text{g}/\text{cm}^2$ (89.1% to 51.8%, respectively), but it displayed a markedly weaker activity (8.3%) at the lowest dose.

Finally, Lig-16 and Lig-118 demonstrated weak activities in all four evaluated doses. Nevertheless, it is worth mentioning that Lig-16 displayed an anomalous behavior. At the lowest dose of 4 $\mu\text{g}/\text{cm}^2$ it showed repellent activity (39.3%) comparable to that of Lig-13 (33%), as was expected due to their structural similarity. However, strangely enough, and in contrast to its cognate ligand, the repellent action of Lig-16 eliminates at higher doses. A plausible explanation could be that Lig-16 acts as an attractant beyond a critical concentration, and this activity prevails over repellency. Such dose-dependent behavior is not unprecedented. For example, indole was found to act as a repellent at the lowest dose of 0.05 μg tested in adults of the common green bottle fly *Lucilia sericata* and as an attractant at the highest dose (5 μg) (Liu et al., 2016b). Also, the volatile odorant Sulcatone (6-methyl-5-hepten-2-one) has been reported not only as a repellent but also as a human host attractant for *Ae. aegypti* mosquitoes (McBride et al., 2014). Nevertheless, this hypothesis can only be confirmed by a suitable mosquito attractant bioassay in future studies.

Conclusion

In conclusion, a combined ligand-based similarity screening and OBP1-based molecular docking approach was applied for the rational discovery of new repellents. As a result, we identified Lig-119 as a novel DEET-like repellent and, Lig-13 and Lig-111 as highly active repellent molecules. Specifically, Lig-119, representing a new molecular scaffold, exhibits a high affinity for the target OBP1 protein while featuring lower volatility (8.55×10^{-4} mmHg) than DEET (1.35×10^{-3} mmHg). Therefore, Lig-119 could provide longer protection times and higher spatial coverage than DEET when compared at equimolar concentrations. This hypothesis regarding the combined effect of

binding affinity and volatility of Lig-119 on the duration and spatial extent of protection, remains to be confirmed in future behavioral assays and field studies. Nevertheless, such experiments that require the synthesis or the purchase of large quantities of compounds of interest, are beyond the scope and scale of this study.

Interestingly, Lig-13 was found to bind in two different sites on AgamOBP1. Given its high binding affinity (0.08 μ M) and low calculated MM-GBSA binding energy (-74 kcal/mol) for the DEET-binding site, we suggest that its activity derives from the targeting of the DEET-site rather than the sIC-site. Such behavior is not unexpected. In our previous work, we showed that the potent synthetic repellent Icaridin is capable of binding to both DEET and sIC binding sites (Drakou et al., 2017). Therefore, the identification of new compounds that are able to bind in multiple protein sites could significantly increase the repertoire of available “olfactophore” features to be used in screening protocols. Hence, the Lig-13 scaffold could be further exploited and optimized for the discovery of better repellents. Regarding our hypothesis that Lig-16 may act as an attractant at high concentration doses in contrast to its structurally similar compound Lig-13, we should clarify that the OBPs are transport proteins that cannot discriminate between repellent and attractant molecules. The physiological activity of compounds, either repellency or attraction of mosquitoes, is eventually defined by the Odorant Receptors, so the type of activity can be ex-post evaluated only in behavioral bioassays.

Furthermore, we previously suggested that a ligand with high *in vitro* affinity and specificity for female antenna-biased OBPs is more likely to display strong repellent action (Thireou et al., 2018; Zographos et al., 2018). Indeed, studies of plant volatile compounds discovered through the OBP-based screening of bioactive essential oils (Kröber et al., 2018) revealed that the most potent of the identified repellent molecules act directly on the Odorant Receptor co-receptor (ORco) of the OR/ORco heterodimeric complex (Kythreoti et al., 2021; Tsitoura et al., 2015). The same OBP-based screening study also showed that essential oil components with no or low affinity for any of the tested OBPs lack of a strong repellent action (Kröber et al., 2018). Based on these results, the observed striking difference in the repellency of the structurally similar ligands Lig-118 (no active) and Lig-111 (third more active compound) could be attributed to the inability of Lig-118 to bind to AgamOBP1. The differences in ligands binding affinity for AgamOBP1 could also account, in part at least, for the weak repellent activity of Lig-103 compared to its structurally related analog Lig-119. However, in that case, the most profound differentiation between Lig-119 (most active) and Lig-103 (inactive) lies in their different predicted binding modes. We have previously suggested that molecules that adopt biologically relevant binding modes are more likely to be specifically recognized by the binding sites of the cognate ORs and initiate a biological response (Thireou et al., 2018; Zographos et al., 2018).

Finally, an important factor to be considered in the screening process of novel repellents is the volatility of the candidate ligands (**Table 2**). On the one hand, the low volatility of a compound exhibiting a high efficacy against mosquitoes, such as Lig-119 (8.55×10^{-4} mmHg), can significantly enhance the duration of protection. On the other hand, high volatility may negatively affect the repellent activity of a ligand as a result of its rapid evaporation e.g. Lig-111 ($1.17 \times 10^{-}$

² mmHg), thus reducing or eliminating its time of protection. Nevertheless, such compounds are of merit, as they can be used in slow-release formulations or devices.

Fig. 4 comes about here

Table 3 comes about here

Table 3. Dissociation constants (K_i) and Repellency Indices (RI) of the tested ligands.

Ethics statement

The laboratory strain of *Ae. albopictus* used in this study was established using mosquito eggs collected from ovitraps from different areas in Greece. The collection areas were public and not privately owned or protected. Mosquito egg collections from the field did not involve endangered or protected animal species. Consequently, the establishment of the laboratory mosquito strain did not require a specific permit. The repellence studies abide by the Declaration of Helsinki principles. The Ethics Committee of Benaki Phytopathological Institute concluded that the current study was implemented in accordance with the Ethics Code for Research.

Acknowledgments

This study was supported by European Regional Development Fund of the EU and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation-call RESEARCH, CREATE, INNOVATE (QFytoTera, T1EDK-00996). This work was also supported by a fellowship to P.G.V.L. by the Hellenic State Scholarships Foundation and the action “Support of human research resources through doctoral research” funded by the “Operational Programme Education and Lifelong Learning” co-funded by the European Social Fund (ESF) and National Resources, MIS5000432.

References

- US EPA. [2017]. Estimation Programs Interface Suite™ for Microsoft® Windows, v 4.11. United States Environmental Protection Agency, Washington, DC, USA.
- Affonso, R.D., Guimaraes, A.P., Oliveira, A.A., Slana, G.B.C., Franca, T.C.C., 2013. Applications of Molecular Modeling in the Design of New Insect Repellents Targeting the Odorant Binding Protein of *Anopheles gambiae*. *Journal of the Brazilian Chemical Society* 24, 473-482.
- Anderson, A.C., 2003. The process of structure-based drug design. *Chem Biol* 10, 787-797.
- Cassau, S., Krieger, J., 2021. The role of SNMPs in insect olfaction. *Cell Tissue Res* 383, 21-33.
- Chang, H.T., Liu, Y., Yang, T., Pelosi, P., Dong, S.L., Wang, G.R., 2015. Pheromone binding proteins enhance the sensitivity of olfactory receptors to sex pheromones in *Chilo suppressalis*. *Sci Rep* 5, 13093. <https://doi.org/10.1038/srep13093>.
- da Costa, K.S., Galucio, J.M., da Costa, C.H.S., Santana, A.R., Dos Santos Carvalho, V., do Nascimento, L.D., Lima, E.L.A.H., Neves Cruz, J., Alves, C.N., Lameira, J., 2019. Exploring the Potentiality of Natural Products from Essential Oils as Inhibitors of Odorant-Binding Proteins: A Structure- and Ligand-Based Virtual Screening Approach To Find Novel Mosquito Repellents. *ACS Omega* 4, 22475-22486.
- Deng, Y., Yan, H., Gu, J., Xu, J., Wu, K., Tu, Z., James, A.A., Chen, X., 2013. Molecular and functional characterization of odorant-binding protein genes in an invasive vector mosquito, *Aedes albopictus*. *PLoS One* 8, e68836. <https://doi.org/10.1371/journal.pone.0068836>
- Devillers, J., 2018. 2D and 3D structure–activity modelling of mosquito repellents: a review. *SAR and QSAR in Environmental Research* 29, 693-723.
- Dima, E., 2015. Isolation of volatile metabolites with possible insect-repellent activity and evaluation of their binding potency for the mosquito odorant binding proteins (OBPs), School of Health Sciences, DEPARTMENT OF PHARMACY. National and Kapodistrian University of Athens., Athens. <https://pergamos.lib.uoa.gr/uoa/dl/object/1316347>, p. 124.
- Drakou, C.E., Tsitsanou, K.E., Potamitis, C., Fessas, D., Zervou, M., Zographos, S.E., 2017. The crystal structure of the AgamOBP1*Icaridin complex reveals alternative binding modes and stereo-selective repellent recognition. *Cell Mol Life Sci* 74, 319-338.
- Estrada, E., 1996. Spectral moments of the edge adjacency matrix in molecular graphs .1. Definition and applications to the prediction of physical properties of alkanes. *Journal of Chemical Information and Computer Sciences* 36, 844-849.
- Gaddaguti, V., Venkateswara Rao, T., Prasada Rao, A., 2016. Potential mosquito repellent compounds of *Ocimum* species against 3N7H and 3Q8I of *Anopheles gambiae*. *3 Biotech* 6, 26. <https://doi.org/10.1007/s13205-015-0346-x>
- Giatropoulos, A., Papachristos, D.P., Kimbaris, A., Koliopoulos, G., Polissiou, M.G., Emmanouel, N., Michaelakis, A., 2012. Evaluation of bioefficacy of three Citrus essential oils against the dengue vector *Aedes albopictus* (Diptera: Culicidae) in correlation to their components enantiomeric distribution. *Parasitol Res* 111, 2253-2263.

- Giatropoulos, A., Papachristos, D.P., Koliopoulos, G., Michaelakis, A., Emmanouel, N., 2015. Asymmetric Mating Interference between Two Related Mosquito Species: *Aedes (Stegomyia) albopictus* and *Aedes (Stegomyia) cretinus*. *PLoS One* 10, e0127762. <https://doi.org/10.1371/journal.pone.0127762>
- González-González, A., Palma-Millanao, R., Yáñez, O., Rojas, M., Mutis, A., Venthur, H., Quiroz, A., Ramírez, C.C., 2016. Virtual Screening of Plant Volatile Compounds Reveals a High Affinity of *Hylamorphia elegans* (Coleoptera: Scarabaeidae) Odorant-Binding Proteins for Sesquiterpenes From Its Native Host. *Journal of insect science* 16, 30. <https://doi.org/10.1093/jisesa/iew008>
- Gouck, H.K., Hall, S.A., Smith, C.N., Gilbert, I.H., 1957. Repellency of Homologous Series of Cyclohexane Aliphatic Acids and Amides. *J Econ Entomol* 50, 175-177.
- Hong, B., Chang, Q., Zhai, Y., Ren, B., Zhang, F., 2022. Functional Characterization of Odorant Binding Protein PyasOBP2 From the Jujube Bud Weevil, *Pachyrhinus yasumatsui* (Coleoptera: Curculionidae). *Front Physiol* 13, 900752. <https://doi.org/10.3389/fphys.2022.900752>
- Kamala Jayanthi, P.D., Kempraj, V., Aurade, R.M., Kumar Roy, T., Shivashankara, K.S., Verghese, A., 2014. Computational reverse chemical ecology: Virtual screening and predicting behaviorally active semiochemicals for *Bactrocera dorsalis*. *BMC Genomics* 15, 209.
- Katritzky, A.R., Wang, Z., Slavov, S., Tsikolia, M., Dobchev, D., Akhmedov, N.G., Hall, C.D., Bernier, U.R., Clark, G.G., Linthicum, K.J., 2008. Synthesis and bioassay of improved mosquito repellents predicted from chemical structure. *Proc Natl Acad Sci U S A* 105, 7359-7364.
- Keita, M., Doumbia, S., Sissoko, I., Toure, M., Diawara, S.I., Konate, D., Sodio, A.B., Traore, S.F., Diakite, M., Doumbia, S.O., Sogoba, N., Krogstad, D.J., Shaffer, J.G., Coulibaly, M.B., 2021. Indoor and outdoor malaria transmission in two ecological settings in rural Mali: implications for vector control. *Malar J* 20, 127.
- Kritsi, E., Liggi, P.G.V., Stamati, E.C.V., Tsitsanou, K.E., Zographos, S.E., Michaelakis, A., Papachristos, D., Zoumpoulakis, P., 2022. A Combined Computational Methodology for the Discovery of Hit Compounds with Putative Insect Repellency Properties. *ChemMedChem* 17, e202200271. <https://doi.org/10.1002/cmdc.202200271>
- Kröber, T., Koussis, K., Bourquina, M., Tsitoura, P., Konstantopoulou, M., Awolola, T.S., Danie, F.R., Qiao, H., Pelosi, P., Iatrou, K., Guerin, P.M., 2018. Odorant-binding protein-based identification of natural spatial repellents for the Africa malaria mosquito *Anopheles gambiae*. *Insect Biochemistry and Molecular Biology* 96, 36-50.
- Kruskal, W.H., Wallis, W.A., 1952. Use of Ranks in One-Criterion Variance Analysis. *J Am Stat Assoc* 47, 583-621.
- Kythreoti, G., Sdralia, N., Tsitoura, P., Papachristos, D.P., Michaelakis, A., Karras, V., Ruel, D.M., Yakir, E., Bohbot, J.D., Schulz, S., Iatrou, K., 2021. Volatile allosteric antagonists of mosquito odorant receptors inhibit human-host attraction. *J Biol Chem* 296, 100172.

- Leal, W.S., 2005. Pheromone reception. *Chemistry of Pheromones and Other Semiochemicals* 11 240, 1-36.
- Leal, W.S., 2013. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol* 58, 373-391.
- Liu, G., Xuan, N., Rajashekar, B., Arnaud, P., Offmann, B., Picimbon, J.F., 2020. Comprehensive History of CSP Genes: Evolution, Phylogenetic Distribution and Functions. *Genes (Basel)* 11.
- Liu, J., Tian, Z., Zhang, Y., 2016a. Structure-based discovery of potentially active semiochemicals for *Cydia pomonella* (L.). *Sci Rep* 6, 34600. <https://doi.org/10.1038/srep34600>
- Liu, W.Q., Longnecker, M., Tarone, A.M., Tomberlin, J.K., 2016b. Responses of *Lucilia sericata* (Diptera: Calliphoridae) to compounds from microbial decomposition of larval resources. *Anim Behav* 115, 217-225.
- Mann, H.B., Whitney, D.R., 1947. On a Test of Whether One of 2 Random Variables Is Stochastically Larger Than the Other. *Ann Math Stat* 18, 50-60.
- Mao, Y., Xu, X., Xu, W., Ishida, Y., Leal, W.S., Ames, J.B., Clardy, J., 2010. Crystal and solution structures of an odorant-binding protein from the southern house mosquito complexed with an oviposition pheromone. *Proc Natl Acad Sci U S A* 107, 19102-19107.
- McBride, C.S., Baier, F., Omondi, A.B., Spitzer, S.A., Lutomiah, J., Sang, R., Ignell, R., Vosshall, L.B., 2014. Evolution of mosquito preference for humans linked to an odorant receptor. *Nature* 515, 222-227.
- Mccabe, E.T., Barthel, W.F., Gertler, S.I., Hall, S.A., 1954. Insect Repellents .3. N,N-Diethylamides. *Journal of Organic Chemistry* 19, 493-498.
- Meyers, J.I., Pathikonda, S., Popkin-Hall, Z.R., Medeiros, M.C., Fuseini, G., Matias, A., Garcia, G., Overgaard, H.J., Kulkarni, V., Reddy, V.P., Schwabe, C., Lines, J., Kleinschmidt, I., Slotman, M.A., 2016. Increasing outdoor host-seeking in *Anopheles gambiae* over 6 years of vector control on Bioko Island. *Malar J* 15, 239.
- Murphy, E.J., Booth, J.C., Davrazou, F., Port, A.M., Jones, D.N., 2013. Interactions of *Anopheles gambiae* odorant-binding proteins with a human-derived repellent: implications for the mode of action of n,n-diethyl-3-methylbenzamide (DEET). *J Biol Chem* 288, 4475-4485.
- Nandi, T.K., Bairagya, H.R., Mukhopadhyay, B.P., Sekar, K., Sukul, D., Bera, A.K., 2009. Conserved water-mediated H-bonding dynamics of catalytic Asn 175 in plant thiol protease. *J Biosci* 34, 27-34.
- Neidle, S., 2021. Structured Waters Mediate Small Molecule Binding to G-Quadruplex Nucleic Acids. *Pharmaceuticals (Basel)* 15(1), 7. <https://doi.org/10.3390%2Fph15010007>
- Neto, M.F.d.A., Santos, C.B.R.d., Magalhães-Junior, J.T., Leite, F.H.A., 2020. Identification of novel *Aedes aegypti* odorant-binding protein 1 modulators by ligand and structure-based approaches and bioassays. *Journal of Biomolecular Structure and Dynamics*, 40, 117-129
- Okoli, B.J., Ladan, Z., Mtunzi, F., Hosea, Y.C., 2021. Vitex negundo L. Essential Oil: Odorant Binding Protein Efficiency Using Molecular Docking Approach and Studies of the Mosquito Repellent. *Insects* 12, 1061

- Oliferenko, P.V., Oliferenko, A.A., Poda, G.I., Osolodkin, D.I., Pillai, G.G., Bernier, U.R., Tsikolia, M., Agramonte, N.M., Clark, G.G., Linthicum, K.J., Katritzky, A.R., 2013. Promising *Aedes aegypti* repellent chemotypes identified through integrated QSAR, virtual screening, synthesis, and bioassay. *PLoS One* 8, e64547. <https://doi.org/10.1371/journal.pone.0064547>
- Pelosi, P., Iovinella, I., Zhu, J., Wang, G., Dani, F.R., 2018. Beyond chemoreception: diverse tasks of soluble olfactory proteins in insects. *Biol Rev Camb Philos Soc* 93, 184-200.
- Perez-Garrido, A., Giron-Rodriguez, F., Helguera, A.M., Borges, F., Combes, R.D., 2014. Topological structural alerts modulations of mammalian cell mutagenicity for halogenated derivatives. *Sar Qsar Environ Res* 25, 17-33.
- Perez-Garrido, A., Helguera, A.M., Cordeiro, M.N.D.S., Escudero, A.G., 2009. QSPR Modelling With the Topological Substructural Molecular Design Approach: beta-Cyclodextrin Complexation. *Journal of Pharmaceutical Sciences* 98, 4557-4576.
- Perez-Garrido, A., Helguera, A.M., Lopez, G.C., Cordeiro, M.N.D.S., Escudero, A.G., 2010. A topological substructural molecular design approach for predicting mutagenesis end-points of alpha, beta-unsaturated carbonyl compounds. *Toxicology* 268, 64-77.
- Perez-Garrido, A., Helguera, A.M., Ruiz, J.M.M., Rentero, P.Z., 2012. Topological sub-structural molecular design approach: Radical scavenging activity. *European Journal of Medicinal Chemistry* 49, 86-94.
- Polishchuk, P.G., Madzhidov, T.I., Varnek, A., 2013. Estimation of the size of drug-like chemical space based on GDB-17 data. *J Comput Aided Mol Des* 27, 675-679.
- Ren, H., Wang, L., Bennett, M., Liang, Y., Zheng, X., Lu, F., Li, L., Nan, J., Luo, M., Eriksson, S., Zhang, C., Su, X.D., 2005. The crystal structure of human adenylate kinase 6: An adenylate kinase localized to the cell nucleus. *Proc Natl Acad Sci U S A* 102, 303-308.
- Robert, X., Gouet, P., 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research* 42, W320-W324.
- Russell, T.L., Govella, N.J., Azizi, S., Drakeley, C.J., Kachur, S.P., Killeen, G.F., 2011. Increased proportions of outdoor feeding among residual malaria vector populations following increased use of insecticide-treated nets in rural Tanzania. *Malar J* 10, 80.
- Spicer, C.W., Gordon, S.M., Kelly, T.J., Holdren, M.W., Mukund, R., 2002. Hazardous Air Pollutant Handbook: Measurements, Properties, and Fate in Ambient Air, 1st ed. CRC Press. p.240. <https://doi.org/10.1201/9781420032352>
- Sun, L., Wang, Q., Yang, S., Wang, Q., Zhang, Z., Khashaveh, A., Zhang, Y.J., Guo, Y.Y., 2017. Functional analysis of female-biased odorant binding protein 6 for volatile and nonvolatile host compounds in *Adelphocoris lineolatus* (Goeze). *Insect molecular biology* 26, 601-615.
- Suryanarayana, M.V.S., Pandey, K.S., Prakash, S., Raghuv eeran, C.D., Dangi, R.S., Swamy, R.V., Rao, K.M., 1991. Structure Activity Relationship Studies with Mosquito Repellent Amides. *Journal of Pharmaceutical Sciences* 80, 1055-1057.
- Thireou, T., Kythreoti, G., Tsitsanou, K.E., Koussis, K., Drakou, C.E., Kinnarsley, J., Krober, T., Guerin, P.M., Zhou, J.J., Iatrou, K., Eliopoulos, E., Zographos, S.E., 2018. Identification of

- novel bioinspired synthetic mosquito repellents by combined ligand-based screening and OBP-structure-based molecular docking. *Insect Biochem Mol Biol* 98, 48-61.
- Tsitoura, P., Koussis, K., Iatrou, K., 2015. Inhibition of *Anopheles gambiae* Odorant Receptor Function by Mosquito Repellents. *Journal of Biological Chemistry* 290, 7961-7972.
- Tsitsanou, K.E., Thireou, T., Drakou, C.E., Koussis, K., Keramioti, M.V., Leonidas, D.D., Eliopoulos, E., Iatrou, K., Zographos, S.E., 2012. *Anopheles gambiae* odorant binding protein crystal complex with the synthetic repellent DEET: implications for structure-based design of novel mosquito repellents. *Cell Mol Life Sci* 69, 283-297.
- Tzotzos, G., Iley, J.N., Moore, E.A., 2018. New insights on repellent recognition by *Anopheles gambiae* odorant-binding protein 1. *PLoS One* 13, e0194724. <https://doi.org/10.1371/journal.pone.0194724>
- Venthur, H., Zhou, J.J., 2018. Odorant Receptors and Odorant-Binding Proteins as Insect Pest Control Targets: A Comparative Analysis. *Front Physiol* 9, 1163.
- Wang, Y., Jin, Y., Chen, Q., Wen, M., Zhao, H., Duan, H., Ren, B., 2017. Selectivity and ligand-based molecular modeling of an odorant-binding protein from the leaf beetle *Ambrostoma quadriimpressum* (Coleoptera: Chrysomelidae) in relation to habitat-related volatiles. *Sci Rep* 7, 15374. <https://doi.org/10.1038/s41598-017-15538-8>
- WHO, 2015. Indoor residual spraying: an operational manual for indoor residual spraying (IRS) for malaria transmission control and elimination., 2nd edition ed, https://apps.who.int/iris/bitstream/handle/10665/177242/9789241508940_eng.pdf;sequence=1.
- Yin, J., Choo, Y.M., Duan, H., Leal, W.S., 2015. Selectivity of odorant-binding proteins from the southern house mosquito tested against physiologically relevant ligands. *Front Physiol* 6, 56.
- Zographos, S.E., Eliopoulos, E., Thireou, T., Tsitsanou, K.E., 2018. OBP-structure-aided repellent discovery: An emerging tool towards the prevention of mosquito-borne diseases, in: Devillers, J. (Ed.), *Computational design of chemicals for the control of mosquitoes and their diseases*. CRC press book series "QSAR in Environmental and Health Sciences" pp. 65-105.

Tables

Table 1. Statistical parameters of the QSAR model development

$$P = 5.58 \cdot 10^{-8} \cdot \mu(15)^{\text{Dip}} - 56.21 \cdot \mu(1)^{\text{Ab-log L16}} + 281.4$$

R²	RMSE	RSS	s	F	Q² (LOO-CV)	RMSE (LOO-CV)
0.94	10.19	623.55	14.42	23.03	0.77	19.80

Table 2. The final hits selected for MD simulations and *in vitro* studies. 2D structures of the hits and the parental.

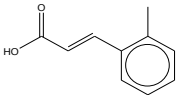
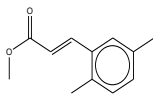
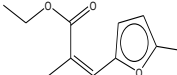
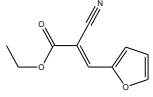
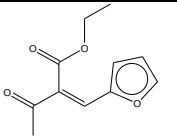
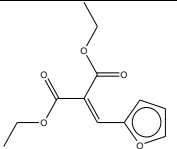
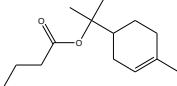
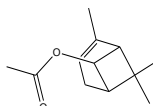
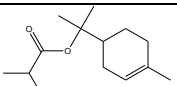
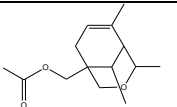
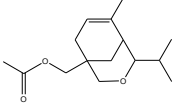
Lig-no	Name	Hit Structure	MW (Da)	VP (mmHg)	Lead	Lead Structure	VP (mmHg)
Lig-86	2-Methylcinnamic acid (predominantly trans)		162.19	4.74×10^{-4}	KO3 Methyl (2E)-3-(2,5-dimethylphenyl) prop-2-enoate		5.19×10^{-3}
Lig-16	ethyl (2Z)-3-(5-bromofuran-2-yl)-2-methylprop-2-enoate		259.10	1.83×10^{-3}	KO11 Ethyl 2-Cyano-3-(2-furanyl)acrylate		5.87×10^{-4}
Lig-11	Ethyl (Z)-2-(furan-2-ylmethylene)-3-oxobutanoate		208.21	1.57×10^{-3}			
Lig-13	diethyl 2-(2-furylmethylene)malonate		238.24	3.92×10^{-3}			
Lig-118	Terpinyl Butyrate		224.34	5.58×10^{-3}	Chrysanthenyl acetate [(1S,5R)-4,6,6-trimethyl-7-bicyclo[3.1.1]hept-3-enyl] acetate		3.19×10^{-2}
Lig-111	Terpinyl Isobutyrate		224.34	1.17×10^{-2}			
Lig-119	(4,6,9-trimethyl-3-oxabicyclo[3.3.1]non-6-en-1-yl)methyl acetate		238.33	8.55×10^{-4}			
Lig-103	(4-isopropyl-6-methyl-3-oxabicyclo[3.3.1]non-6-en-1-yl)methyl acetate		252.35	4.63×10^{-4}			

Table 3. Dissociation constants (K_i) and Repellency Indices (RI) of the tested ligands.

<i>Ligand</i>	Binding energy (kcal/mol)		K _i , μM		Repellency Index (x100)			
			Hi	Lo	4 μg/cm ²	10 μg/cm ²	20 μg/cm ²	40 μg/cm ²
DEET			Not tested		92.2	96.5	96.8	99.8
Site	DEET	sIC						
Lig-16	-86	-98	0.040	7.71	39.3	8.1	0.2	11.0
Lig-13	-74	-60	0.082	12.84	33.0	81.5	100	100
Lig-119	-89	nd	0.042		81.8	96.6	99.3	100
Lig-103	-72	nd	1.246		12.4	8.0	11.4	22.2
Lig-111	-87	nd	0.060		8.3	51.8	76.8	89.1
Lig-118	-83	nd	No binding ¹		11.1	16.1	25.2	27.6
Lig-86	-29	nd	No binding ^{1,2}		Not tested ²			
Lig-11	-66	nd	Not tested ³		Not tested ³			

¹2% fluorescence reduction at 50 μM ligand.

² Soluble only in DMSO. Insoluble in bioassay compatible solvents DCM or EtOH.

³ Not commercially available

Captions to illustrations

Fig. 1: Filtering of active hits. Shape and pharmacophoric features of (A) yomogi alcohol, (B) artemisia alcohol, (C) terpinen-4-ol, (D) chrysanthenyl acetate, (E) KO11, (F) KO9, (G) KO8, (H) KO4, (I) KO2 and (J) KO3. Important pharmacophoric features were also marked. K: Number of hits retrieved from each query. L: Venn diagram suggesting the top 30% of hits against Rec-1, 2 and 3. The numbers of predicted molecules that are common for all the targets are also mentioned.

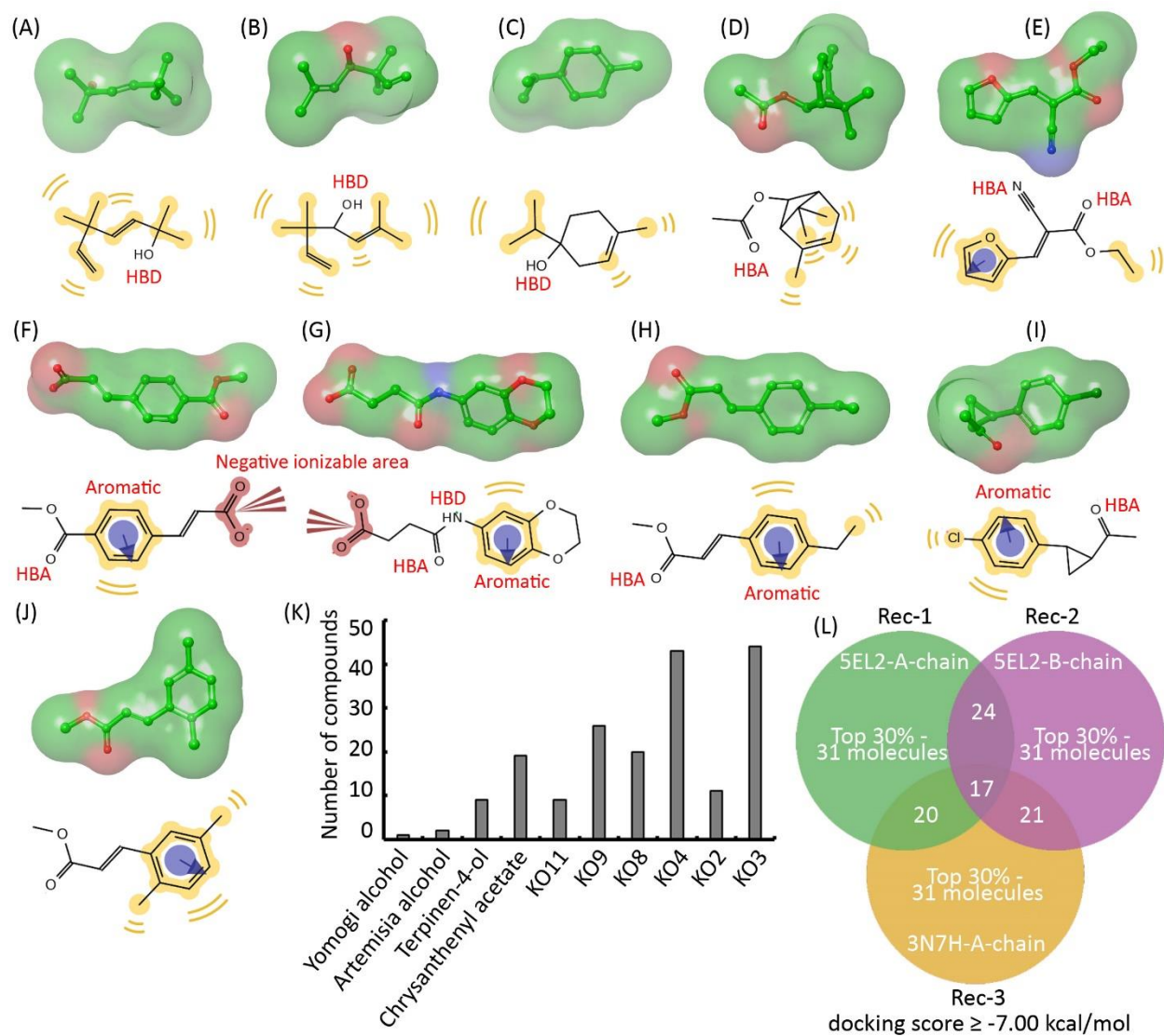
Fig. 2: Binding of molecule towards AgamOBP1. A: Heat map indicating MM-GBSA for all the selected 17 molecules. B-I: Cartoon representation (Left) and 2D protein ligand interaction diagram (Right) showing binding modes and interactions of Lig-11 (B), Lig-13 (C), Lig-16 (D), Lig-86 (E), Lig-103 (F), Lig-111 (G), Lig-118 (H) and Lig-119 (I) to the DEET binding site of AgamOBP1. (Left) For figure simplicity the cartoon representations shown only the side chains of selected protein residues bordering the binding sites. The hydrogen bonds are represented as dashed lines. (Right) In the schematic diagram of Protein-Ligand interactions all the protein residues within a 4Å distance from the ligand are shown as spheres colored by residue type; Polar (cyan), Positive charged (violet), Hydrophobic (green), Cysteine (dark grey) and Glycine (light grey). The hydrogen bonds are represented as purple arrows and the π - π stacking interactions as green lines. Atoms exposed to the bulk solvent are marked with a surrounding gray circle.

Fig. 3. Comparative %fluorescence reduction of AgamOBP1/1-NPN solutions titrated with various concentrations of the selected ligands. Data of Lig-16 and Lig-13 best fit the 2-site model. R squared (R^2) provide the goodness-of-fit.

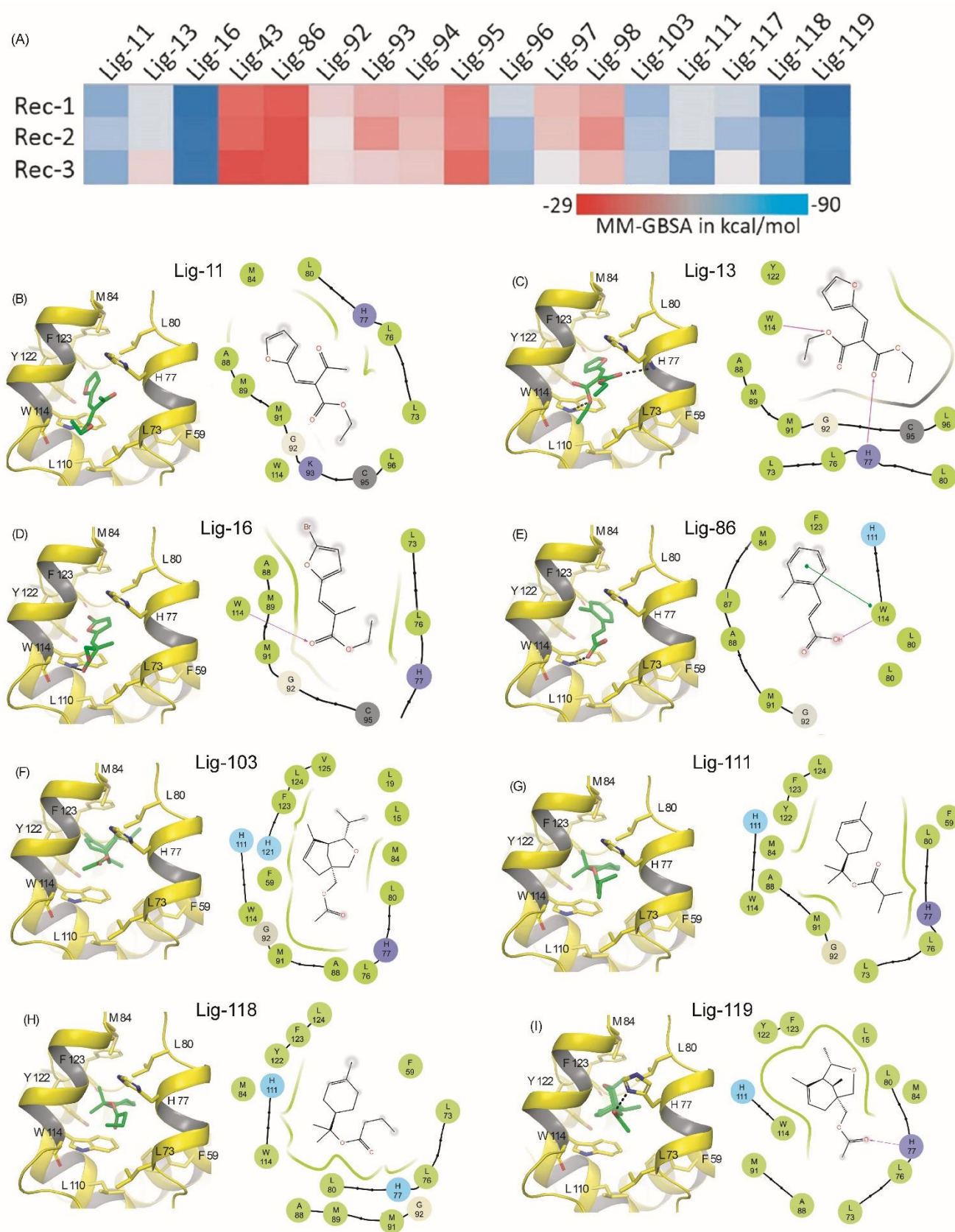
Fig. 4. Box plots depicting Repellency Index of *Ae. albopictus* female mosquitoes after exposure to a dose of A) 40 $\mu\text{g}/\text{cm}^2$, B) 20 $\mu\text{g}/\text{cm}^2$, C) 10 $\mu\text{g}/\text{cm}^2$ and D) 4 $\mu\text{g}/\text{cm}^2$ of the tested material for five minutes. The different letters indicate significant differences among treatments ($P < 0.05$, Mann–Whitney U test with Bonferroni correction).

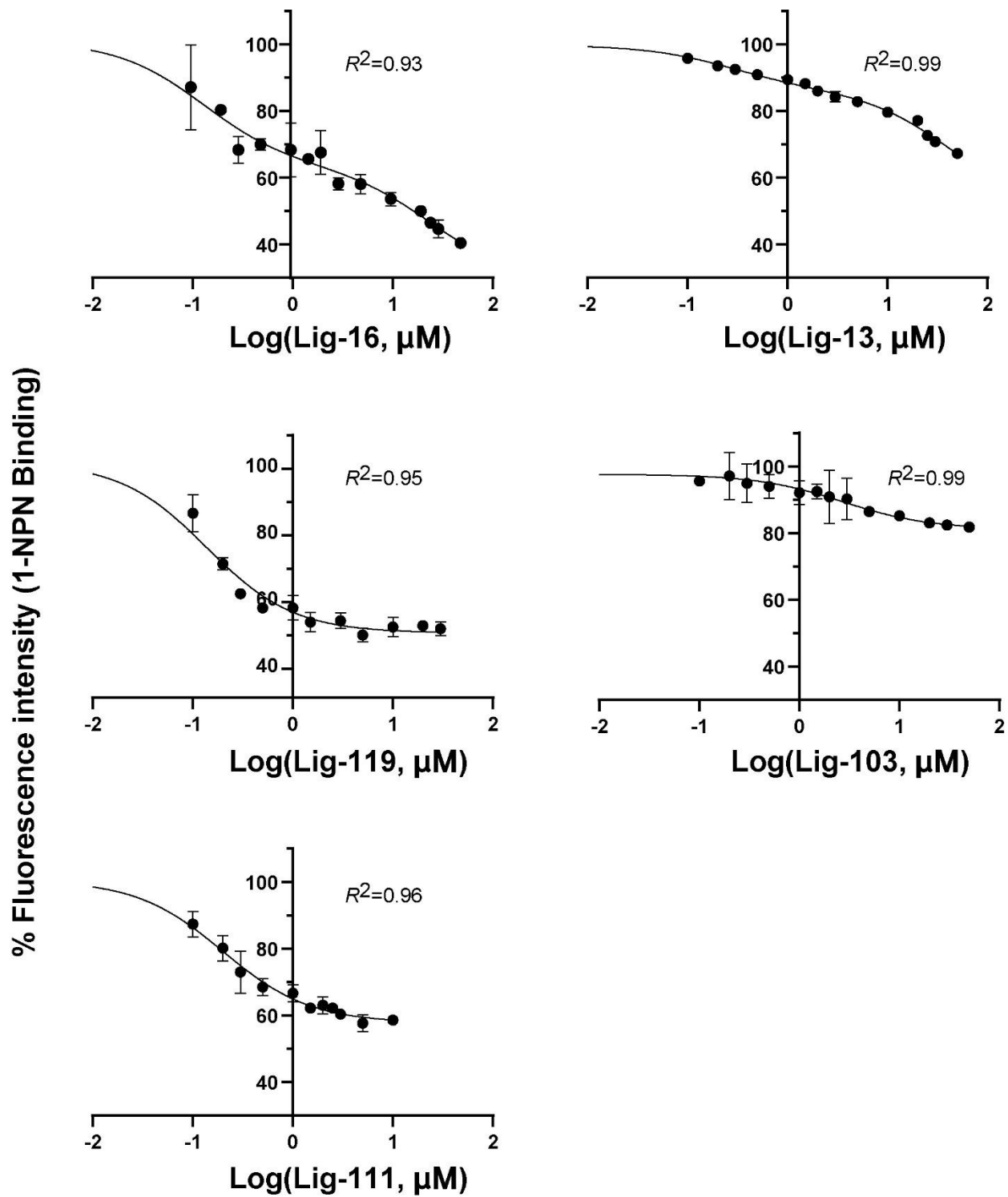
Illustrations

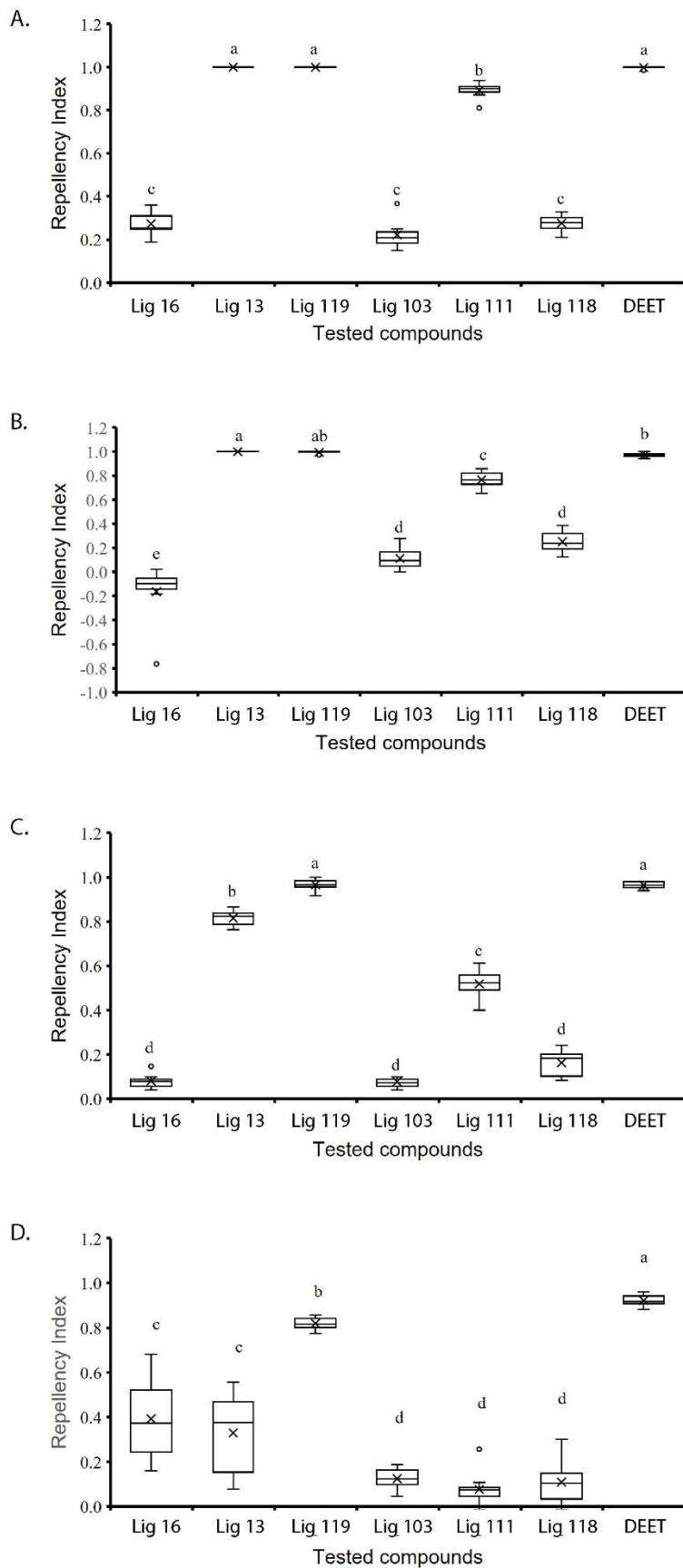
Insect Biochem. Mol. Biol., Liggri et al., **Fig. 1**



Insect Biochem. Mol. Biol., Liggri et al., Fig. 2







Supplemental Material

2D finger-printing and molecular docking studies identified potent mosquito repellents targeting Odorant Binding Protein 1

Panagiota G.V. Liggri^{1,5}, Alfonso Pérez-Garrido², Katerina E. Tsitsanou¹, Kalarickal V. Dileep³ Antonios Michaelakis⁴, Dimitrios P. Papachristos⁴, Horacio Pérez-Sánchez^{2*}, Spyros E. Zographos^{1*}

¹Institute of Chemical Biology, National Hellenic Research Foundation, 48 Vassileos Constantinou Avenue, 11635 Athens, Greece

²Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), Universidad Católica de Murcia (UCAM), 30107 Spain

³Laboratory for Computational and Structural Biology, Jubilee Centre for Medical Research, Jubilee Mission Medical College and Research Institute, Thrissur, Kerala-680005, India

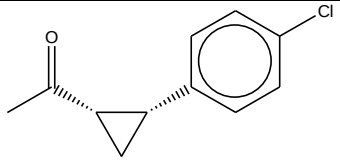
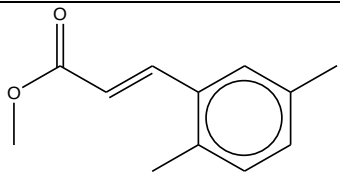
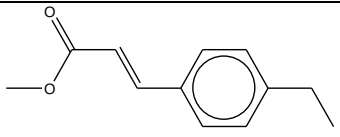
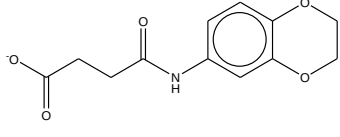
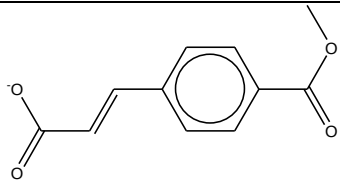
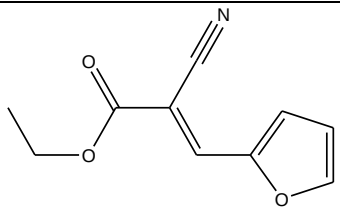
⁴Benaki Phytopathological Institute, Department of Entomology and Agricultural Zoology, 8 S Delta Str. 14561 Kifissia, Athens, Greece

⁵Department of Biochemistry and Biotechnology, University of Thessaly, Biopolis, 41500 Larissa, Greece

Running title: Potent mosquito repellents targeting AgamOBP1

Correspondence to: Horacio Pérez-Sánchez; e-mail: hperez@ucam.edu (Corresponding author for inquiries related to the computational studies) and Spyros E. Zographos; Institute of Chemical Biology, National Hellenic Research Foundation, 48, Vassileos Constantinou Avenue, 11635 Athens, Greece. Tel.: +302107273850, e-mail: sez@eie.gr

Table S1. The structures of the lead compounds and the number of acquired screening hits

Lead Code/ name	Lead structure	No of Hits
KO2		11
KO3		44
KO4		43
KO8		20
KO9		26
KO11		8

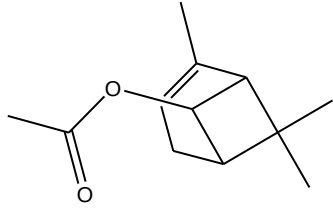
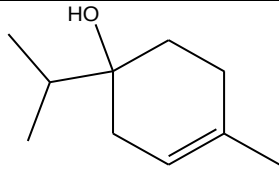
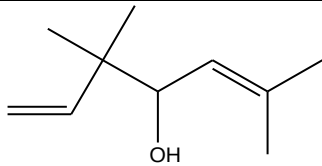
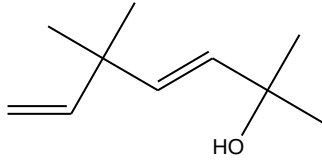
Chrysanthenyl acetate		19
Terpinen-4-ol		9
Artemisia alcohol		2
Yomogi alcohol		2
Hits Total (Unique)		184 (152)

Table S2. Final dataset of 120 molecules that are selected for the studies.

Index	SMILES
1	<chem>OC(C)(C)\C=C\CC=C</chem>
2	<chem>O[C@@]1(CCC(=CC1)C)C(C)C</chem>
3	<chem>OC([C@H]1CCC(=CC1)C)(C)C</chem>
4	<chem>OC1(CCCCC1)CC=C</chem>
5	<chem>OC(CC1=CCCCC1)(C)C</chem>
6	<chem>OC1CC(CC(=C1)C)(C)C</chem>
7	<chem>O[C@@H]1C[C@H](CC=C1C)C(O)(C)C</chem>
8	<chem>OCC(C1CCC(=CC1)C)C</chem>
9	<chem>OC(C1CCC(=CC1)C)(CCC=C(C)C)C</chem>
10	<chem>[o]1c(ccc1)\C=C\C(=O)OCCOC(C)C</chem>
11	<chem>[o]1c(ccc1)\C=C(/C(=O)OCC)\C(=O)C</chem>
12	<chem>[o]1c(ccc1)\C=C\C(=O)OCC</chem>
13	<chem>[o]1c(ccc1)C=C(C(=O)OCC)C(=O)OCC</chem>
14	<chem>[o]1c(ccc1)\C=C(\C)/C(=O)O</chem>
15	<chem>[o]1c(ccc1)\C=C(/C)\C(=O)O</chem>
16	<chem>Brc1[o]c(cc1)\C=C(\C)/C(=O)OCC</chem>
17	<chem>[o]1c(ccc1)\C=C\C(=O)OC(C)C</chem>
18	<chem>OC(=O)c1c(cccc1)\C=C\C(=O)O</chem>
19	<chem>O(C)C(=O)c1ccc(cc1)\C=C\C(=O)O</chem>
20	<chem>O(Cc1ccc(cc1)C(=O)OC)C</chem>
21	<chem>O(C)C(=O)\C=C\c1ccc(cc1)\C=C\C(=O)OC</chem>
22	<chem>OC(=O)c1ccc(cc1)\C=C\C(=O)O</chem>
23	<chem>O(C)C(=O)c1ccc(cc1)CO</chem>
24	<chem>N(CC)C(=O)\C=C\c1ccc(cc1)C(=O)OC</chem>
25	<chem>O(CC)C(=O)c1ccc(cc1)C(=O)OCC</chem>
26	<chem>O(C)C(=O)c1ccc(cc1)C(=O)OC</chem>
27	<chem>O(C\C=C\c2ccccc2)C(=O)c1ccc(cc1)O</chem>
28	<chem>O(C)C(=O)c1ccc(cc1)C(=O)O</chem>
29	<chem>O(C)C(=O)\C=C\c1cccc1</chem>
30	<chem>O(C)C(=O)c1cc(ccc1)C(=O)OC</chem>
31	<chem>O(C)C(=O)c1ccc(cc1)C</chem>
32	<chem>O(C)C(=O)\C=C\c1cc(ccc1)C</chem>
33	<chem>N(c1cc2c(cc1)OCCO2)C(=O)CCC</chem>
34	<chem>N(c1cc2c(cc1N)OCCO2)C(=O)C</chem>
35	<chem>N(c1cc2c(cc1)OCCO2)C(=O)C(C)(C)C</chem>
36	<chem>N1C(CCC1)C(=O)Nc2cc3c(cc2)OCCO3</chem>
37	<chem>N(c2cc3c(cc2)OCCO3)C(=O)C1C(C1)C</chem>
38	<chem>N(c1cc2c(cc1)OCCO2)C(=O)COC</chem>
39	<chem>N(c1cc2c(cc1)OCCO2)C(=O)C(CC)CC</chem>
40	<chem>NCC(=O)Nc1cc2c(cc1)OCCO2</chem>

41 N(CC(=O)Nc1cc2c(cc1)OCCO2)C
 42 N(c2cc3c(cc2)OCCO3)C(=O)C1CCC1
 43 **N(c1cc2c(cc1)OCCO2)C(=O)CC(=O)C**
 44 N3(C(=O)CCC3=O)c1cc2c(cc1)OCCO2
 45 N(c1cc2c(cc1)OCCO2)C(=O)C(C)C
 46 N(c1cc2c(cc1)OCCO2)C(=O)CC(C)(C)C
 47 N(c2cc3c(cc2)OCCO3)C(=O)C1CCCC1
 48 N(c1cc2c(cc1)OCCO2)C(=O)C(CCC)C
 49 N(c2cc3c(cc2)OCCO3)C(=O)C1OCCC1
 50 N(c1cc2c(cc1)OCCO2)C(=O)C=C(C)C
 51 Fc1ccc(cc1)\C(=C\C(=O)OC)\C
 52 Clc1ccc(cc1)\C(=C\C(=O)OCC)\C
 53 [Na+].[O-]C(=O)\C=C/c1ccccc1
 54 O(CC)C(=O)\C=C\c1cc(ccc1)C(=O)C
 55 N(CCOc(=O)\C=C\c1ccccc1)(C)C
 56 O(C\C=C\c1ccccc1)C(=O)CC
 57 O(CC)C(=O)\C=C(/C)\c1ccccc1
 58 OC(=O)\C=C\c1cc(ccc1)C(=O)C
 59 OC(=O)\C=C\c1ccc(cc1)\C=C\C(=O)O
 60 O(C)C(=O)\C=C\c1c(ccc(c1)C)C
 61 OC(=O)\C=C(/C)\c1ccc(cc1)CC
 62 Fc1cc(ccc1)\C=C\C(=O)OC
 63 OC(=O)\C=C\c1c(cc(c(c1)C)C)C
 64 O(C)C(=O)\C=C\c1c(cc(c(c1)C)C)C
 65 Nc1ccc(cc1)\C=C\C(=O)OCC
 66 Fc1ccc(cc1)\C=C\C(=O)OC
 67 OC(=O)\C=C\c1cc(ccc1)C
 68 OC(=O)\C=C\c1ccc(cc1)C
 69 O(CCC)C(=O)\C=C\c1ccccc1
 70 O(CC(C)C)C(=O)\C=C\c1ccccc1
 71 OC(=O)\C=C\c1c(ccc(c1)C)C
 72 Fc1ccc(cc1)\C=C\C(=O)OCC
 73 O(C)C(=O)c1ccc(cc1)\C=C\C(=O)OC
 74 O(C\C=C\c1ccccc1)C(=O)C
 75 O(C)C(=O)\C=C/c1c2c(ccc1)cccc2
 76 Brc1ccc(cc1)\C=C\C(=O)OC
 77 O(C)C(=O)\C=C\c1ccc(cc1)CC
 78 O(C)c1ccc(cc1)\C=C\C(=O)OC
 79 O(C)C(=O)\C=C\c1c(cc(cc1)C)C
 80 OC(=O)\C=C\c1c(cccc1)CC
 81 OC(=O)\C=C\c1cc(c(cc1)C)C
 82 OC(=O)\C=C\c1c(cc(cc1)C)C
 83 Clc1ccc(cc1)\C=C\C(=O)OC
 84 Fc1c(cccc1)\C=C\C(=O)OC
 85 O(C)C(=O)\C=C\c1ccc(cc1)O

86 OC(=O)\C=C\c1c(cccc1)C
87 O(CC)C(=O)\C=C\c1ccccc1
88 OC(=O)\C=C\c1cc(cc(c1)C)C
89 OC(=O)\C=C\c1c(cc(cc1C)C)C
90 O(CC=C)C(=O)\C=C\c1ccccc1
91 O(C)C(=O)\C=C\c1ccc(cc1)C(=O)O
92 Clc1ccc(cc1)CC(=O)C(C)C
93 Clc1ccc(cc1)CCC(=O)C(C)C
94 Fc1ccc(cc1)C2C(C2)C(=O)C
95 Clc1ccc(cc1)[C@@H]2[C@H](C2)C(=O)O
96 Clc1cc(ccc1)CC(=O)CC(C)C
97 Clc1ccc(cc1)C2CCCC2=O
98 Clc1cc(ccc1)C2CC(=O)CC(=O)C2
99 Clc1ccc(cc1)C2CC(=O)CC(=O)C2
100 Clc1ccc(cc1)C(=O)C2CCC2
101 Clc1ccc(cc1)C2C(C2)C(=O)O
102 O1C(C2CC(C1)C(C=C2C)C)C(C)C
103 O1C(C2CC(C1)(CC=C2C)COC(=O)C)C(C)C
104 O(C1[C@@H]2C[C@H](C1)C=C2)C(=O)C
105 O(C(C1CCC(=CC1)C)(C)C)C(=O)C
106 O(C1C2CC(C1)C=C2)C(=O)C
107 O1C(C3C(C(C1)C(C=C3C)C)C)C2CCC(=CC2)C
108 O1C(C2C(C(C1)(C(C=C2C)C)CO)C)C
109 O(CC1=CCC2C(C1C2)(C)C)C(=O)C
110 O(C(C1CCC(=CC1)C)(C)C)C(=O)CC
111 O(C(C1CCC(=CC1)C)(C)C)C(=O)C(C)C
112 O(C(C(CC=C(C)C)CC=C(C)C)C)C(=O)C
113 O1C2(C3C(C(C1)C(C=C3C)C)C)CCCC2
114 O1C(C3CC(C1)CC=C3C)C2CCC(=CC2)C
115 O(C1CC(CC=C1C)C(=C)C)C(=O)C
116 O(C([C@@H]1CCC(C2=C(C1)C(CC2)C)C)(C)C)C(=O)C
117 O(C1C2C3C(C(C2)C1)CC=C3)C(=O)C
118 O(C(C1CCC(=CC1)C)(C)C)C(=O)CCC
119 CC(=O)OC[C@@]12[C@H](C)[C@@H](C(C)=CC1)[C@H](C)OC2
120 OC(CC=C)\C=C/C

Table S3. Final dataset of 17 molecules that are selected for the studies.

Sl No	Ligand numbering according to the main list	SMILES
1	11	<chem>[O]1C(CCC1)\C=C/C(=O)OCC\C(=O)C</chem>
2	13	<chem>[O]1C(CCC1)C=C(C(=O)OCC)C(=O)OCC</chem>
3	16	<chem>BrC1[O]C(CC1)\C=C\C(C)/C(=O)OCC</chem>
4	43	<chem>N(C1CC2C(CC1)OCCO2)C(=O)CC(=O)C</chem>
5	86	<chem>OC(=O)\C=C\C1C(CCCC1)C</chem>
6	92	<chem>ClC1CCC(CC1)CC(=O)C(C)C</chem>
7	93	<chem>ClC1CCC(CC1)CCC(=O)C(C)C</chem>
8	94	<chem>Fc1ccc(cc1)C2C(C2)C(=O)C</chem>
9	95	<chem>ClC1CCC(CC1)[C@@H]2[C@H](C2)C(=O)O</chem>
10	96	<chem>ClC1CC(CCC1)CC(=O)CC(C)C</chem>
11	97	<chem>ClC1CCC(CC1)C2CCCC2=O</chem>
12	98	<chem>ClC1CC(CCC1)C2CC(=O)CC(=O)C2</chem>
13	103	<chem>O1C(C2CC(C1)(CC=C2C)COC(=O)C)C(C)C</chem>
14	111	<chem>O(C(C1CCC(=CC1)C)(C)C)C(=O)C(C)C</chem>
15	117	<chem>O(C1C2C3C(C(C2)C1)CC=C3)C(=O)C</chem>
16	118	<chem>O(C(C1CCC(=CC1)C)(C)C)C(=O)CCC</chem>
17	119	<chem>CC(=O)OC[C@@]12[C@H](C)[C@@H](C(C)=CC1)[C@H](C)OC2</chem>

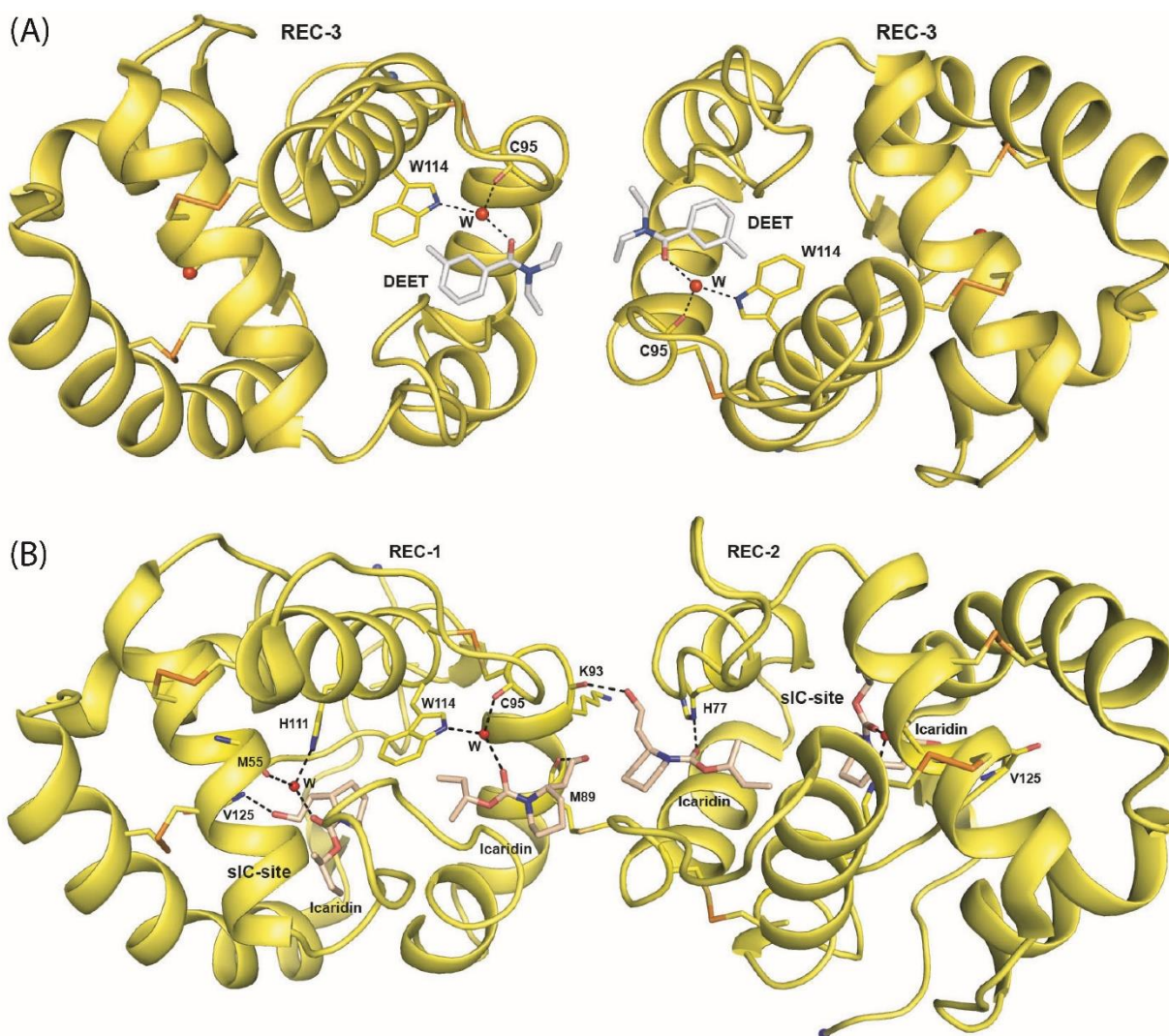


Figure S1. Overall structure of AgamOBP1-DEET & AgamOBP1-Icaridin complexes.

(A) DEET binds to a binding site located at the edge of a long hydrophobic tunnel running through each OBP1 monomer (**REC-3**) by exploiting numerous vdW interactions and one hydrogen bond to a conserved water molecule which in turn interacts with Thr114 and Cys95 residues. (B) Icaridin occupies two distinct binding sites on the OBP1 monomer, the DEET-site where it adopts two different orientations, and a second sIC-site located at the C-terminal region. At the DEET-binding site of monomer A (**REC-1**), it adopts a DEET-like binding mode and hydrogen bond network, while at monomer B (**REC-2**) is arranged in an anti-periplanar orientation and interacts through hydrogen bonds with His77 and Met89 of the other subunit. In the sIC-binding site, Icaridin adopts the same conformation in both protein monomers and participates in a direct hydrogen bond interaction with the main chain nitrogen of the C-terminal Val125 and in a water-mediated hydrogen bond to Met55 and His111. The detailed structural analysis of AgamOBP1-DEET and AgamOBP1-Icaridin complexes is presented by Tsitsanou et al. (2012) and Drakou et al. (2017), respectively.

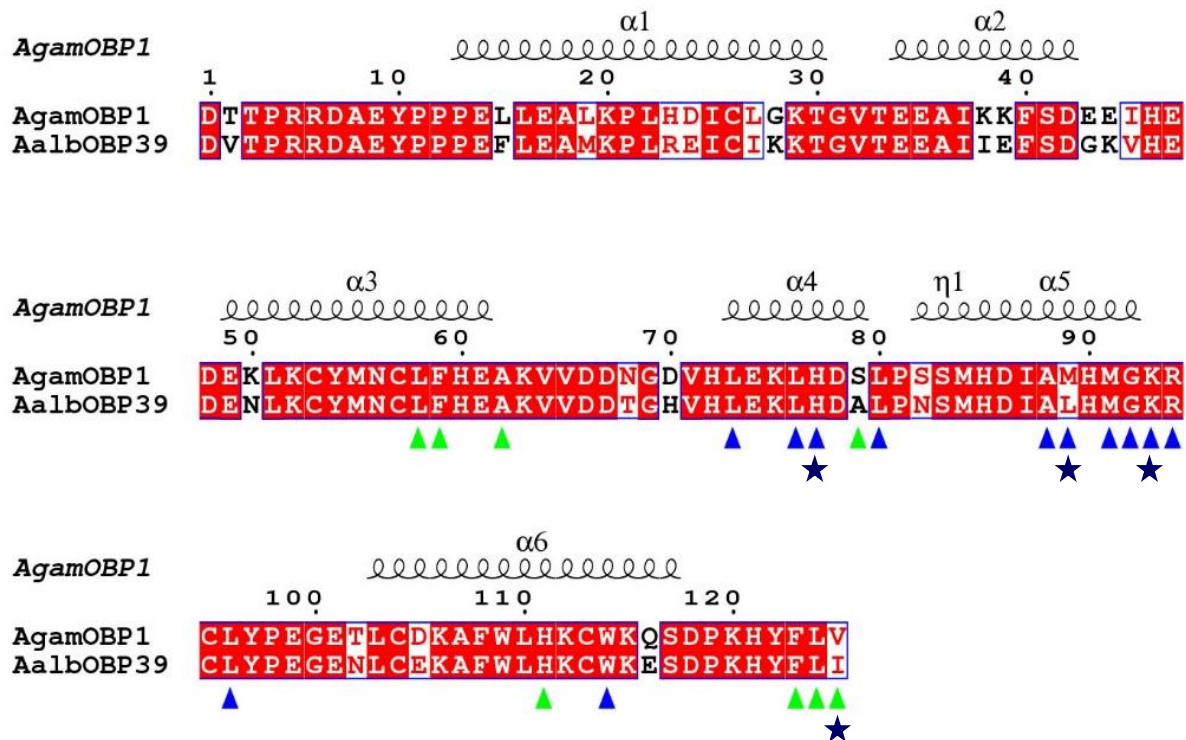


Figure S2. Sequence alignment of AgamOBP1 and AalbOBP39. Sequence alignment of the homologous AgamOBP1 (Uniprot: A0A1U7F4W2) and AalbOBP39 (Uniprot: M4W733) proteins from *Anopheles gambiae* and *Aedes albopictus* mosquitoes. Sequences were aligned using Clustal Omega online software and are represented by the single letter amino acid code. The residues of AgamOBP1 that are involved in hydrogen bonds and vdW interactions with DEET and Icaridin at the DEET-binding site are indicated by blue triangles. Those that interact with Icaridin in the secondary sIC-binding site are pointed by green triangles. Residues that form direct hydrogen bonds with the ligands are indicated by an asterisk.

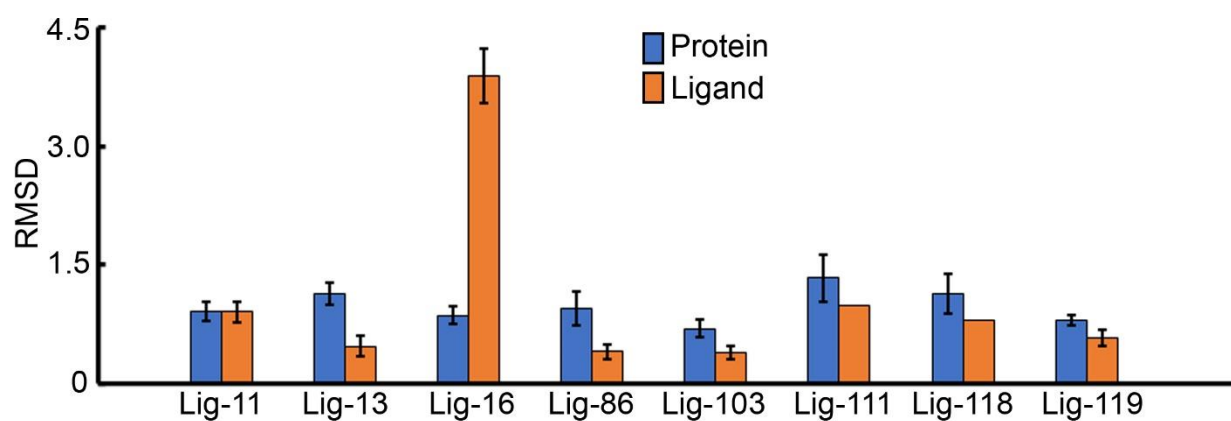


Figure S3. RMSD of the protein (marked as blue bars) and ligand (marked as orange bars) complexes for 500 ns simulation.

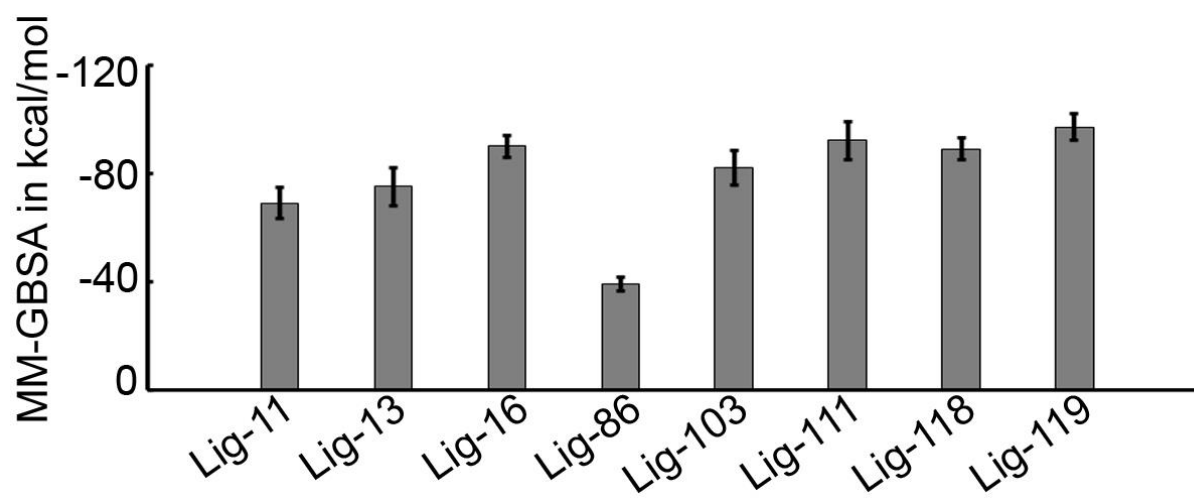


Figure S4. The MM-GBSA calculated for ligands from the trajectory of 500 ns.

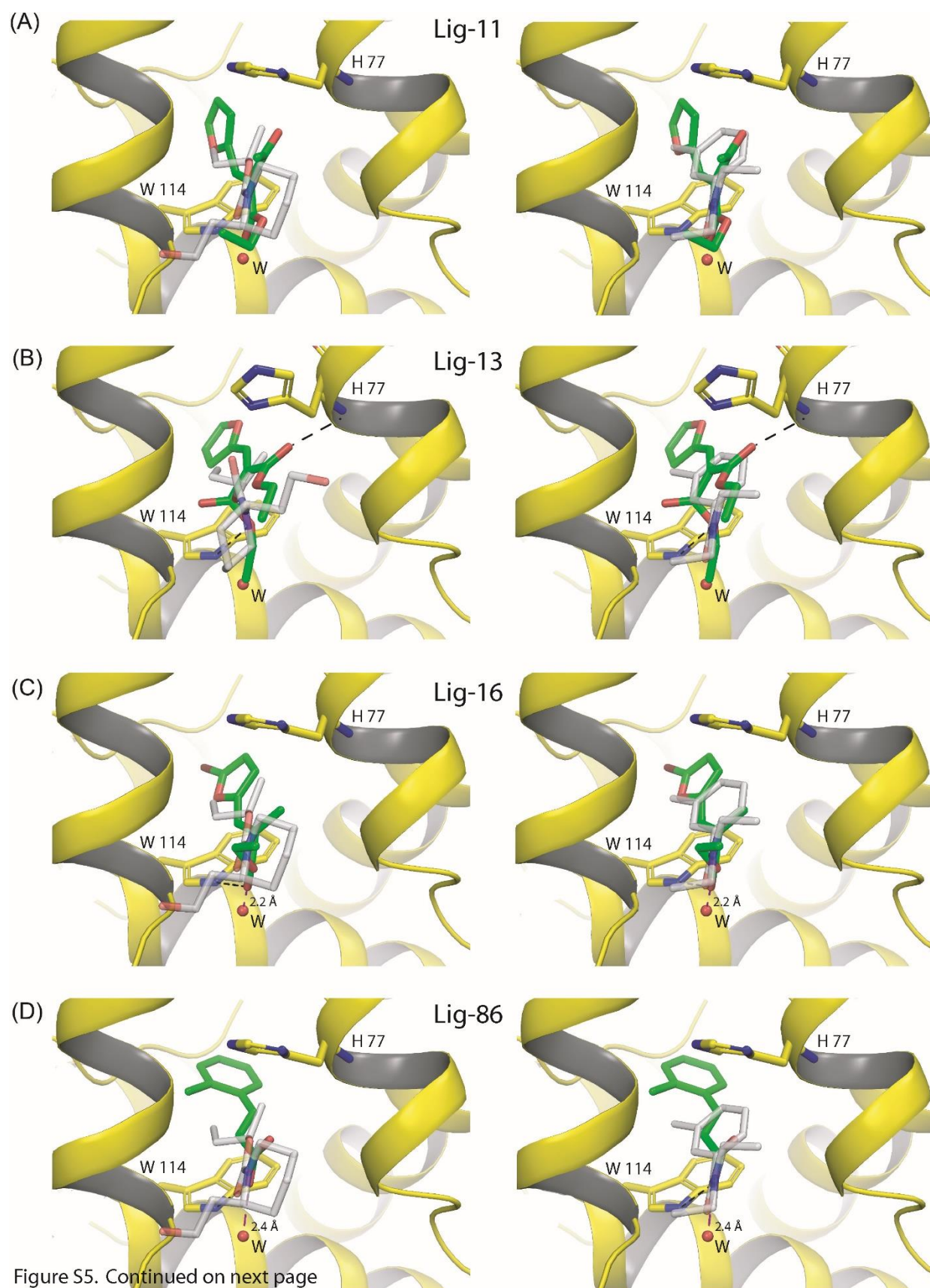


Figure S5. Continued on next page

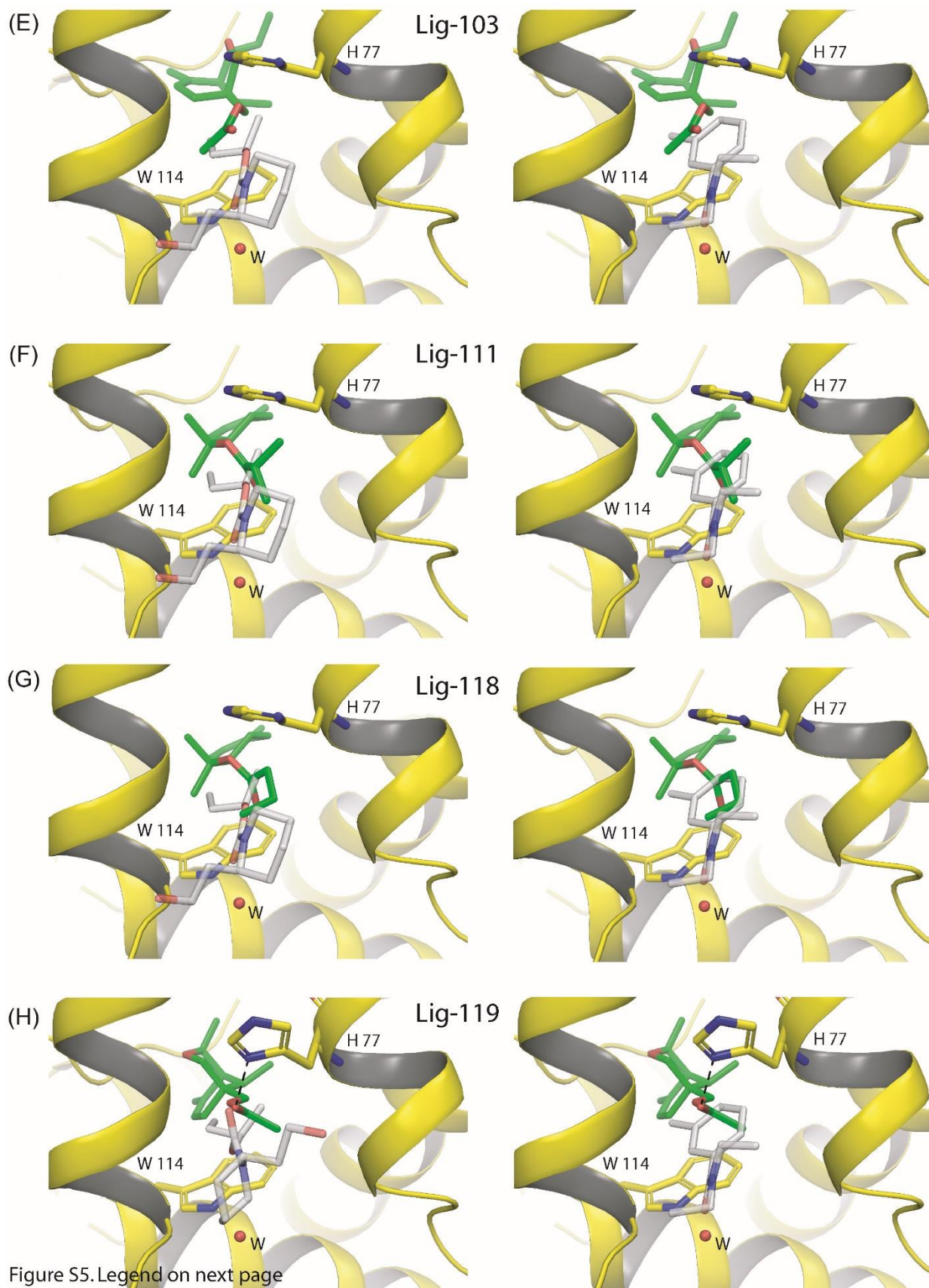


Figure S5. Structural comparison of ligands with DEET and Icaridin. Superimposition of docked ligand conformations (green) with the crystallographic conformations of (*left*) Icaridin (PDB: 5EL2; Rec-1 and 2) and (*right*) DEET (PDB: 3N7H; Rec-3) bound to AgamOBP1 (white). The binding modes of Lig-13 and Lig-119 which are hydrogen bonded to His77 are compared with the conformation of Icaridin interacting with His77 (Rec-2). All the other ligands are compared to the Icaridin molecule interacting with the structural water (Rec-1). Structural water molecule is shown as a red sphere. Hydrogen bonds between ligand atoms and His77 or Trp114 residues are presented as black dashed lines. The potential hydrogen bonds between the oxygen atoms of Lig-16 or Lig-86 and the water molecule are displayed as purple dashed lines.

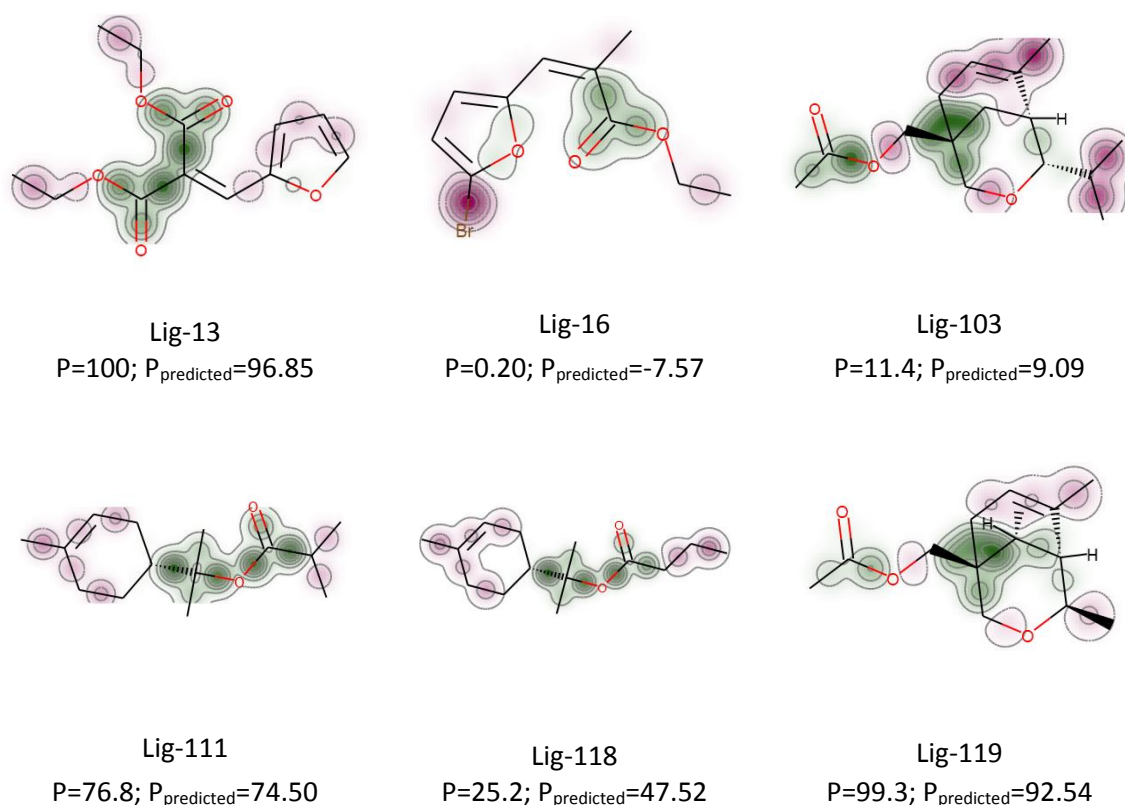


Figure S6. Contribution of six chemicals to the repellent activity. The figure shows the values of the observed repellent activity (P) and repellent activity predicted by the QSAR model (P_{predicted}). Green: Positive contribution; Red: Negative contribution.

REFERENCES

Drakou, C.E., Tsitsanou, K.E., Potamitis, C., Fessas, D., Zervou, M., Zographos, S.E., 2017. The crystal structure of the AgamOBP1*Icaridin complex reveals alternative binding modes and stereo-selective repellent recognition. *Cell Mol Life Sci* 74, 319-338.

Tsitsanou, K.E., Thireou, T., Drakou, C.E., Koussis, K., Keramioti, M.V., Leonidas, D.D., Eliopoulos, E., Iatrou, K., Zographos, S.E., 2012. Anopheles gambiae odorant binding protein crystal complex with the synthetic repellent DEET: implications for structure-based design of novel mosquito repellents. *Cell Mol Life Sci* 69, 283-297.