Computational Biology and Bioinformatics
2014; 2(3): 38-42
Published online June 20, 2014 (http://www.sciencepublishinggroup.com/j/cbb)
doi: 10.11648/j.cbb.20140203.12

In silico identification of potential inhibitors of dengue mosquito, *Aedes aegypti* chorion peroxidase

Edwin Plata Alcantara

National Institute of Molecular Biology and Biotechnology, University of the Philippines Los Banos, College, Laguna, Philippines

Email address:
epalcantara@uplb.edu.ph

To cite this article:
Edwin P. Alcantara. *In Silico* Identification of Potential Inhibitors of Dengue Mosquito, *Aedes Aegypti* Chorion Peroxidase. *Computational Biology and Bioinformatics.* Vol. 2, No. 3, 2014, pp. 38-42. doi: 10.11648/j.cbb.20140203.12

Abstract: The three dimensional structure of *Aedes aegypti* chorion peroxidase was computed by homology modeling. The ModWeb server provided the most accurate model with QMEAN score of 0.642. The protein model consists of 36.1% alpha-helices and 1% beta-strand. Ligand binding sites in *Aedes aegypti* chorion peroxidase were identified using SiteComp server. *In silico* docking of a subset of ZINC natural products database was focused on the predicted binding site. Three ligands were found to be potential inhibitors of *Ae. aegypti* chorion peroxidase.

Keywords: Homology Modeling, *Aedes Aegyti*, Chorion Peroxidase, Binding Site, *In Silico* Screening

1. Introduction

Dengue fever and dengue hemorrhagic fever (DF/DHF) are mosquito borne diseases of public health concerns in tropical and subtropical parts of the world [1], affecting millions of people annually [2]. The Philippines ranks second in incurring the dengue disease burden among countries in Southeast Asia [3]. In fact in 2010, the total number of cases and deaths attributed to dengue was highest in the Philippines [4]. The main insect vector of the disease is the mosquito, *Aedes aegypti* [5]. Currently, controlling this vector with insecticidal spraying remains an important option to minimize the incidence of dengue fever [6].

Natural products with mosquitocidal activity can be found from plants and microorganisms. The traditional approach to natural products discovery is through tedious screening of microbial and plant extracts followed by bioassay-guided identification and structure elucidation. However, the availability of public domain databases of protein sequences and experimentally determined structures of plant and microbial secondary metabolites together with advanced computational power offer cost-effective bioinformatics approach for discovery of new mosquitocidal compounds.

The chorion of *Aedes aegypti* eggs undergoes a hardening process following oviposition and individual chorion proteins become insoluble thereafter. The enzyme, chorion peroxidase is primarily responsible for the irreversible insolubilization of the chorion proteins after oviposition [7]. *Ae. aegypti* chorion peroxidase has not been investigated as potential target for development of natural products such as mosquitocides. The development of ovicidal compounds targeting chorion peroxidase would complement existing larvicidal and adulticidal compounds for control of *Ae. aegypti*. The objectives of this study are to construct a homology model of *Aedes aegyti* chorion peroxidase enzyme and to identify by computational method, potential inhibitors of chorion peroxidase.

2. Methodology

2.1. Homology Modeling

Homology modeling was carried out to predict the three dimensional (3D) structure of *Ae. aegypti* chorion peroxidase. The amino acid sequence of *Ae. aegypti* chorion peroxidase was downloaded from the NCBI website and uploaded separately to the Phyre2 [8], ModWeb [9], CPH [10] and RaptorX [11] homology modeling servers. Structure refinement of the predicted models was carried out using ModRefiner [12]. Quality of the predicted protein structures was evaluated by using online version of QMEAN and RAMPAGE. The binding sites in the protein was predicted with the SiteComp server [13]. The binding site predicted by the SiteComp server was validated with POOL [14] and DISCERN [15] protein functional site prediction servers.

2.2. In Silico Screening

Pharmacophore search in the ZINC compound database [16] for candidate inhibitors of *Ae. aegypti* chorion peroxidase (AaCP) was carried out using the on-line server, ZINCPharmer [17]. A well-defined pharmacophore model includes both hydrophobic volumes and hydrogen bond vectors. Heme bound to bovine lactoperoxidase (PDB I.D. 3q9k) was uploaded to the server to generate pharmacophoric features for screening the subset of natural products in the ZINC compound database. The software Autodock Vina [18] as implemented in PyRx (http://pyrx.sorceforge.net) was used to predict the binding pose and binding affinity of each ligand. Virtual screening was conducted with rigid receptor conformation. Docking of ligands was focused on the active site (size: x=25, y=25, z=25; center: 9.645, 2.54, 25.155). Computation was performed with a MacIntosh computer with quadcore Intel Core i5 CPU running at 2.7 GHz. Ligand with the most favorable binding energy (i.e. increasingly negative value) was considered a potential inhibitor of chorion peroxidase activity.

3. Results and Discussion

3.1. Model Quality Evaluation

Homology models of proteins have been previously used for discovery of ligands [19]. In this study, four homology modeling servers were utilized to predict the 3-D structure of *Ae. aegypti* chorion peroxidase. Each of the four servers automatically assigned a protein template for the uploaded amino acid sequence of *Ae. aegypti* chorion peroxidase. Homology modeling builds a three dimensional structure of the target protein (with no experimentally determined 3D structure) based on sequence identity to known protein structures [20, 21]. Therefore, sequence identity is good determinant for the quality of the model. In general, sequence of at least one related structure must have more than 30% identity [22]. As a general rule, those models with sequence identities between 25% to 50% can be used to assess target 'druggability" [23]. The most reliable model was obtained from the ModWeb server (Table 1). The server used bovine lactoperoxidase (PDB ID 3q9k) as template with 33% sequence identity with *Ae. aegypti* chorion peroxidase. The two other servers that produced good models were CPH and Phyre2. The low QMEAN score produced by the I-Tasser model might be due to low sequence identity with the template protein although model quality is not always directly related to the identity between template and target sequence [24]. In a previous benchmark study, it was demonstrated that Modeller performed better than other modeling programs [25].

Mod Web server is a Web interface to ModPipe [26] which in turn uses the Modeller program to calculate homology models of query protein sequences. MODELLER takes target-template alignment file as input and without user intervention it generates a 3D model. Initial step of model building is, identification of spatial restraints for example, distances and dihedral angles lying on the target sequence followed by alignment with template sequence. Interaction of many features of protein structure is analyzed statistically and used to derive spatial restraints on the target sequence [27]. The ModWeb model (Fig. 1) contained 36.1% alpha-helices, and 1% beta-strand. The stereochemical quality of the ModWeb-generated model was highest among the three models taking into consideration both the QMEAN score and percent sequence identity with the template protein (Table 1). The main chain conformation for 93.4% of all residues in the ModWeb model were within the most favoured regions, 5.2% in allowed regions and 1.4% in outlier regions as determined by Ramachandran plot analysis (Fig. 2).

Table 1. *Comparison of homology modeling server performance in prediction of Aedes aegypti chorion peroxidase three-dimentional structure.*

Server	Template	% Identity	QMEAN Score
Phyre2	2gjm	31.0	0.497
RaptorX	2o86	30.0	0.574
ModWeb	3q9k	33.0	0.642
CPH	2eha	30.9	0.604

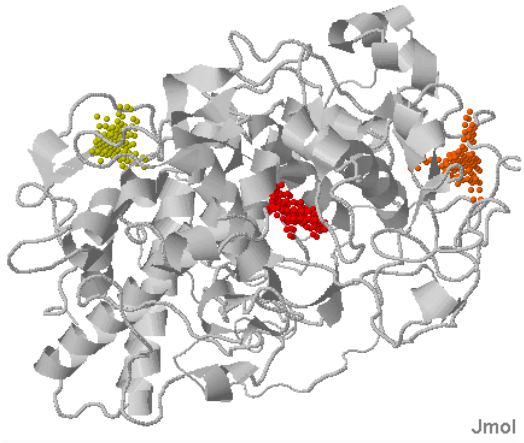


Figure 1. *ModWeb-predicted three dimensional structure of Aedes aegypti chorion peroxidase. Model structure was obtained from template 3q9kA. Modeled region was from residue 210 to 785. Length of target sequence is 790 residues. Predicted ligand binding site is shown as red sphere clusters. Each molecular interaction field (MIF) cluster represents a binding site. MIFs describe the spatial variation of the interaction energy between target molecule (i.e., chorion peroxidase) and a probe (i.e., methyl carbon) [11].*

3.2. Binding Site Identification

In silico analysis identified putative ligand binding (red sphere cluster) and decoy sites (orange and yellow clusters) in *Ae. aegypti* chorion peroxidase (Fig. 1). The putative ligand binding site corresponds to the heme binding site in 3q9k. The binding site in the ModWeb model is not identified in the other 3D structures generated by the other servers. The ligand binding site (total energy= -1404.83) contained 26 amino acid residues (Table 2). The two decoy sites were smaller in volume as compared to the putative ligand binding site. Statistically significant size difference exists between true ligand binding site and other sites on a particular protein surface [28]. The predicted binding site is

validated by the results of the POOL server analysis (Table 2). Fifteen (Gly300, Gln301, Ser304, His305, Thr308, Leu309, Arg447, Arg535, Pro536, Ala544, His547, Arg548, His551, Leu629, Arg633) out of 26 residues predicted by the SiteComp server were also identified as binding sites by the POOL server.

Table 2. *Results of POOL server analysis of the predicted Aedes aegypti chorion peroxidase binding site.*

Rank	POOL Score	Residue number
1	0.001666223979555	ARG:548
2	0.001088240067475	HIS:547
3	0.000595350051299	HIS:305
4	0.000279450032394	ARG:633
5	0.000153869987116	ASP:306
6	0.000151200016262	HIS:551
7	0.000132696004584	ASP:390
8	0.000107250001747	ASP:445
9	0.000085085004685	ASP:383
10	0.000083226004790	TYR:709
11	0.000058916004491	TYR:706
12	0.000058464003814	LYS:617
13	0.000054432006436	GLN:301
14	0.000053550003940	ARG:447
15	0.000051216004067	ARG:640
16	0.000049028003559	HIS:387
17	0.000038880003558	SER:304
18	0.000030096001865	TYR:501
19	0.000023760001568	THR:308
20	0.000018040000214	ARG:337
21	0.000017640000806	ARG:535
22	0.000017009999283	LEU:309
23	0.000016743999367	LYS:710
24	0.000013524000678	TYR:716
25	0.000013200000467	ALA:544
26	0.000011424001059	TYR:496
27	0.000010710001334	HIS:458
28	0.000008189999789	TYR:254
29	0.000007919999916	ASN:630
30	0.000007776000530	PHE:302
31	0.000007392000498	TYR:549
32	0.000006480000593	PRO:536
33	0.000006480000593	ARG:311
34	0.000006210000720	ARG:248
35	0.000005039999905	GLY:300
36	0.000004752000223	MET:537
37	0.000004576000720	TYR:395
38	0.000004049999916	LEU:629

Among the top ten amino acid residues listed in Table 2, five residues are also included in the top ten list of the DISCERN server. The five residues with corresponding DISCERN scores are as follows: Arg633=5.82, Arg548=4.70, His551=4.62, Asp306=4.25, and Asp445=4.21.

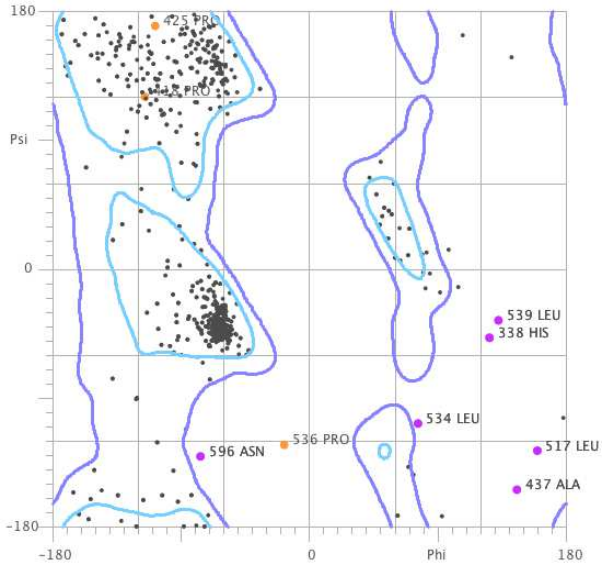


Figure 2. *Ramachandran plot of ModWeb-predicted three dimensional structure of Aedes aegypti chorion peroxidase. Outlier residues are indicated as magenta and orange dots.*

3.3. In Silico Screening

In silico screening is an alternative or complement to high throughput screening and permits screening of compounds which are not physically available in the laboratory. A preliminary pharmacophore filtering of a subset of ZINC natural products database yielded 243 compounds (data not shown). A summary of results of *in silico* docking of these pre-filtered compounds to *Ae. aegypti* chorion peroxidase are shown in Table 3. Three compounds with molecular weight ranging from 489.568 g/mol to 542.995 g/mol were predicted to have tight binding to *Ae. aegypti* chorion peroxidase as follows: ZINC 02158845 (-11.9 kcal/mol), ZINC 70701039 (-11.2 kcal/mol), and ZINC 12902647 (-11.1 kcal/mol) (Table 3). The amino acid residues interacting with the top scoring compound ZINC 02158845 are shown in Fig. 3. One residue, glutamine 450 is hydrogen bonded to O4 in ZINC 02158845. Nineteen residues (Ser304, His338, Phe339, Trp436, Arg447, Asn449, Pro536, Ala544, His547, Arg548, His551, Phe576, Asn577, Leu615, Phe616, Lys617, Glu618, Leu629, Arg633) in the binding site were in hydrophobic contact with the ligand. The interaction between these hydrophobic region of the binding site with the ligand are often observed to provide the driving force for binding [29]. The presence of proline in the binding site is unusual because this amino acid residue is generally unreactive and does not adopt many protein main-chain conformations [30]. Chorion peroxidase has not been utilized as target for development of mosquitocidal compounds. Enzyme inhibition and ovicidal assays will be required to validate the predicted ovicidal activity of these three compounds.

Table 3. *Result of binding site analysis of Aedes aegypti chorion peroxidase.*

Rank	Total Energy	Volume (Å)	Amino acid residues in the binding site
1	-1404.83	124	G300, Q301, S304, H305, T308, L309, H358, F339, R447, Q450, L451, R535, P536, A544, H547, R548, G550, H551, V554, I572, F576, L611, L615, L626, L629, R633
2	-1319.59	103	E255, W259,A260, P261, R262, H264, S265, V266, N269, L270, L271, P272, S273, A274, I277, F333, R337, F342, P343, E345, I631, Q632, R635, Y671, D676, D678
3	-1281.28	99	N293, L295, L561, M592, F593, F595, N596, I745, Q759, E760, A761, Q764, D768, N769, P771

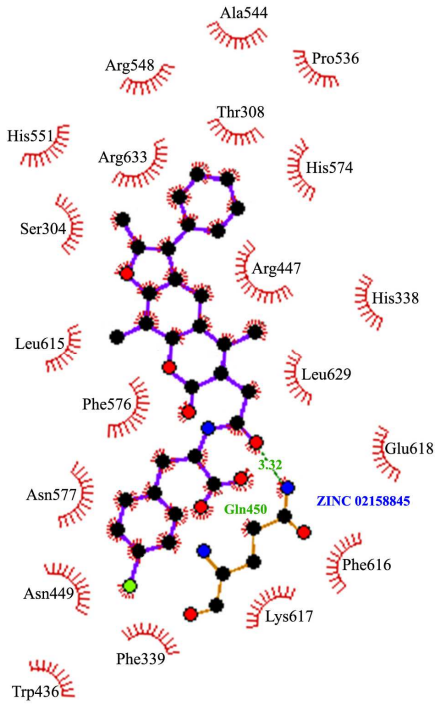
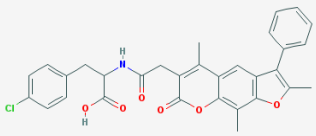
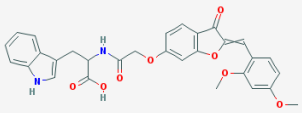
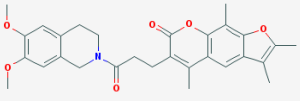


Figure 3. *2-D representation of amino acid residues in Aedes aegypti chorion peroxidase binding site interacting with ZINC 02158845. Hydrogen bonding is indicated by green dashed line between atoms involved. Hydrophobic contacts are represented by an arc with spokes radiating toward the ligand atoms they contact. Diagram was constructed using LigPlot V1.4.5 [31].*

Table 3. *Chemical properties of predicted Aedes aegypti chorion peroxidase inhibitors from subset of ZINC natural products database.*

Ligand ID	Molecular Weight (g/mol)	Molecular Formula	Chemical Structure
ZINC 02158845	542.995	C31H26ClNO6	
ZINC 70700093	541.536	C30H26N2O8	
ZINC 12902647	489.568	C29H31NO6	

4. Conclusion

Homology modeling is a very useful tool for predicting with good accuracy the 3D structure of *Ae. aegypti* chorion peroxidase. Experimental validation of the predicted anti-chorion peroxidase activity of the candidate ligands should be conducted.

References

[1] J. Ponlawat, G. Scott, and L.C. Harrington, "Insecticide susceptibility of *Aedes aegypti* and *Aedes albopictus* across Thailand," J. Med. Entomol. Vol. 42, pp.821–825, 2005.

[2] M. Jacobs, "Dengue: emergence as a global public health problem and prospects for control," Trans R Soc Trop Med Hyg. Vol. 94, pp. 7–8, 2000.

[3] D.S. Shepard, E.A. Undurraga, and Y.A. Halasa, "Economic and disease burden of dengue in Southeast Asia," PLoS Negl. Trop. Dis. Vol. 7, pp. e2055. 2013.

[4] Y. Arima and T. Matsui, "Epidemiologic update of dengue in the Western Pacific Region, 2010," West. Pacif. Surveil. Res. Jour. Vol. 2, pp. 4-8, 2011.

[5] M.S. Chang, E.M. Christophel, D. Gopinath, R. Abdur R, "Challenges and future perspective for dengue vector control in the Western Pacific Region", West. Pacif. Surveil. Res. Jour. Vol. 2, pp. 9-16, 2011.

[6] H. Khan, W. Akram, K. Shehzad, and E. Shaalan, "First report of field evolved resistance to agrochemicals in dengue mosquito, *Aedes albopictus* (Diptera: Culicidae), from Pakistan," Parasites & Vectors Vol. 4, pp. 146-157, 2011.

[7] J.S. Li and J. Li, "Major chorion proteins and their crosslinking during chorion hardening in *Aedes aegypti* mosquitoes," Ins. Biochem. Mol. Biol. Vol. 36, pp. 954-964, 2006.

[8] L.A. Kelly and M.J.E. Sternberg, "Protein structure prediction on the web: a case study using the Phyre server," Nat. Prot. Vol. 4, pp. 363-371, 2009.

[9] N. Eswar, B. John, N. Mirkovic, A. Fiser, V. Ilyin, et al., "Tools for comparative protein structure modeling and analysis," Nuc. Acids Res. Vol. 31, pp. 3375-3380, 2003.

[10] M. Nielsen, C. Lundegaard, O. Lund, T.N. Petersen, "CPHmodels-3.0- Remote homology modeling using structure guided sequence profiles," Nuc. Acids Res. Vol. 38, doi:10.1093/nar/gkq535, 2010.

[11] M. Kallberg, H. Wang, S. Wang, J. Peng, Z. Wang, et al., "Template-based protein structure modeling using the RaptorX web server," Nat. Prot. Vol. 7, pp. 1511-1522, 2012.

[12] D. Xu and Y. Zhang, "Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization," Biophy. J. Vol. 101, pp. 2525-2534, 2011.

[13] Y. Lin, "SiteComp: a server for ligand binding site analysis in protein structures," Bioinformatics Vol. 15, pp. 1172-73, 2012.

[14] S. Somarowthu and M. Ondrechen, "POOL server: machine learning application for functional site prediction in proteins," Bioinfor. Vol. 28, pp. 2078-2079, 2012.

[15] S. Sankararaman, F. Sha, J.F. Kirsch, M.I. Jordan, K. Sjolander, "Active site prediction using evolutionary and structural information," Bioinfor. Vol. 26, pp. 617-624, 2010.

[16] J.J. Irwin,and B.K. Shoichet, "ZINC- a free database of commercially available compounds for virtual screening," J. Chem. Inf. Model, Vol. 45, pp. 177–182, 2005.

[17] D. Koes D and C.J. Camacho, "ZINCPharmer: pharmacophore search of the ZINC database," Nuc. Acids Res. Vol. 40, pp. W409-W414, 2012.

[18] O. Trott and A.J. Olson, "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading," Jour. Comp. Chem. Vol. 31, pp. 455-461, 2010.

[19] M. Jacobson and A. Sali, "Comparative protein structure modeling and its applications to drug discovery," *Annu. Rep. Med. Chem., Vol. 39*, pp. 259–276, 2004.

[20] A. Sali, "100,000 protein structures for the biologist," Nat. Struct. Biol. Vol. 5, pp. 1029-1032, 1998.

[21] D. Vitkup, E. Melamud, J. Moult, and C. Sander, "Completeness in structural genomics," Nat. Struct. Biol. Vol. 8, pp. 559-566. 2001.

[22] S. Roberto and S. Andrej, "Comparative protein structure modeling: Introduction and practical examples with MODELLER," In: Protein Structure Prediction: Methods and Protocols. Ed: Webster, D. M., 97- 129. Humana Press. 2000.

[23] B. Wallner and A. Elofsson. "All are not equal: A benchmark of different homology modeling programs," Prot. Sci., vol. 14, pp. 1315-1327, February 2005.

[24] S. Bilal, S.B. Ali, S. Fazai, and A. Mir. "Generation of a 3D model for human cereblon using comparative modeling," Jour. Bioinf. Seq. Anal. Vol. 5, pp. 10-15, 2013.

[25] C.N. Cavasotto and S.S. Phatak, "Homology modeling in drug discovery: current trends and applications," Drug Disc.Today. Vol 14, pp. 676-683, 2009.

[26] U. Pieper, N. Eswar, B. Webb, E. Eramian, L. Kelly, et al., "MODBASE, a database of annotated comparative protein structure models, and associated resources," Nucleic Acids Res. Vol. 37, D347-D354. 2009.

[27] A. Sali and T.L. Blundell, "Comparative protein modeling by satisfaction of spatial restraints," Jour. Mol. Biol. Vol. 234, pp. 779-815.

[28] P.J. Hajduk, J.R. Huth, and S.W. Fesik. "Druggability indices for protein targets derived from NMR-based screening data," J. Med. Chem. Vol. 48, pp. 2518-2525,2005.

[29] M.D. Kelly, R.L. Mancera, "A new method for estimating the importance of hydrophobic groups in the binding site of a protein," Jour. Med. Chem. Vol. 48. pp. 1069-1078.

[30] M.J. Betts, R.B. Russell, "Amino acid properties and consequences of substitutions. In Bioinformatics for Geneticists (M.R. Barnes, I.C. Gray, eds). Wiley, 2003.

[31] A.C. Wallace, Laskowski R.A., and Thornton J.M. "LIG PLOT: a program to generate schematic diagrams of protein-ligand interactions," Protein Eng. Vol. 8, pp. 127-134, 1996.