IN SILICO SITE-DIRECTED MUTAGENESIS OF THE *Anopheles gambiae* ODORANT BINDING PROTEIN 20

JOSE ISAGANI B. JANAIRO[1,*],
PATRICIA ISABEL K. BRAVO[2], NINNA LOUISE G. MORANO[2],
DERRICK ETHELBHERT C. YU[2]

[1] *Biology Department, De La Salle University, 2401 Taft Avenue, Manila, 0922, Philippines (*jose.isagani.janairo@dlsu.ed)*
[2] *Chemistry Department, De La Salle University, 2401 Taft Avenue, Manila, 0922, Philippines*

Abstract: The *Anopheles gambiae* is a highly anthropophilic mosquito which is the leading vector for malaria. This disease has affected more than 500 million people worldwide. The *Anopheles gambiae* targets its hosts through the odors of the human skin and sweat where odorant molecules radiate. These odors elicit specific responses from the insect through the odorant – binding proteins (OBP). Recently, a specific type of OBP has been characterized which is known as the *Anopheles gambiae* odorant – binding protein 20 (AgamOBP20). This OBP is highly expressed in the female mosquito antennae during the peak of its host – seeking behavior and thus may play a role in olfactory perception. The binding site of the AgamOBP20 is composed primarily of hydrophobic residues wherein the importance of each residue is herein analysed to further understand the properties of *Agam*OBP20. This was carried out through computer – aided site – directed mutagenesis coupled with homology modelling and docking simulations wherein each residue in the binding site was changed to alanine and serine. Probable key amino acid residues were identified as LEU106, LEU107, and MET53 which are hypothesized to play a significant role in the protein – ligand interaction. These residues had the greatest impact in the binding free energy when mutated with alanine and serine. The presented results suggest that steric hindrance and hydrophobic interaction are crucial factors to consider on the manner in which the ligand binds with *Agam*OBP20. The molecular features and parameters obtained may be utilized for the development of new pesticides and repellents that are able to block the function of *Agam*OBP20 and may result to the disarray of the host – seeking behavior of the *Anopheles gambiae*.

Keywords: odorant binding proteins, *Anopheles gambiae*, site-directed mutagenesis

1. Introduction

Malaria is a disease acquired from the female *Anopheles* mosquitoes. It is caused by parasites that belong to the genus *Plasmodium.* Previous studies indicated that the parasite is introduced to the female mosquito vector through its accumulation of blood from infected humans. The parasite develops and reproduces inside the host where infection may spread once the mosquito bites more humans. The common symptoms of malaria include fever and a flu-like illness, which will lead to death if not treated. According to the Center for Disease Control (2015), about 3.2 billion people are at great risk of acquiring malaria. A total of 97 countries were calculated to have been experiencing malaria transmission within their domain in the year of 2015. There have been approximately 198 million cases of this disease

DOI 10.1515/nbec-2016-0016
© University of SS. Cyril and Methodius in Trnava

DE G DE GRUYTER OPEN

Bereitgestellt von Slovenská poľnohospodárska knižnica | Heruntergeladen 17.01.20 14:56 UTC

around the world and about 500,000 people have died due to malaria. The *Anopheles gambiae* mosquitoes are believed to able to locate their targets through olfactory sensing facilitated by odorant binding proteins (OBP) (ZHOU, 2010; PELOSI, 1994; PELOSI and MAIDA, 1990). The hypothesized role of OBPs in odor perception is that they are odorant molecule carriers. They are able to bind with odorant molecules and carry them across the mucus barrier to the olfactory receptor, triggering the mosquito's mechanism of olfaction (PELOSI, 1994; PELOSI and MAIDA, 1990; PELOSI and TIRINDELLI, 1989; PEVSNER *et al.*, 1988). OBPs are generally hydrophobic in nature and thus, aid hydrophobic compounds to traverse across the aqueous mucus layer. One of the newly discovered OBPs in the *A. gambiae* mosquito is the OBP20. A recent study using DNA microarray analysis on the *A. gambiae* under light/dark cycles showed that the gene expression of the AgamOBP20 is regulated by light/dark cycles, the period where its host seeking behavior is at its peak, suggesting that this OBP has a vital role in the olfactory responses (ZIEMBA *et al.*, 2013; RUND *et al.*, 2011). Evidently, using microarray-based survey on mRNA levels of the *A. gambiae*, showed that the AgamOBP20 is one of the OBPs that are expressed highest in the female antennae exhibiting around 3-6 fold elevated expression levels (BIESSMANN *et al.*, 2005). The crystal structure of the AgamOBP20 has a specific route for the arrival and the departure of the odorant molecules from humans into a certain binding pocket. The binding pocket in the AgamOBP20 is lined up with mostly hydrophobic residues which are: Met6, Met7, Gly10, Ile13, Met50, Met53, Ile70, Ile73, Met82, Leu86, Leu106, Leu107, Leu110, Phe117, Ile118, Phe119 and Pro120. The polar residues Thr55 and Ser66 are quite close on one edge of the protein and are strategically positioned so that they will be able to hydrogen bond with the polar groups of the ligands. These residues greatly contribute to the activity of the AgamOBP20 (ZIEMBA *et al.*, 2013). The manner of the conformation of the AgamOBP20 is associated with the binding free energy of the ligand. Currently, insufficient information is present on the odorant preference and the manner of which this OBP adapts to its odorants. However, a study on the docking simulations performed on certain carboxylic acids found in sweat and limburger cheese with AgamOBP20 showed that oleic acid could be a potential ligand since it exhibited the most favourable binding free energy value when docked with AgamOBP20. (JANAIRO *et al.*, 2015). Evidently, oleic acid is a component of limburger cheese, having an odor similar to the human skin and is one of the main attractant for the female *Anopheles gambiae*, further supporting the role of oleic acid as a potential ligand (KNOLS *et al.*, 1997).

Site – directed mutagenesis using *in silico* methods was employed by mutating each residue, within the binding site, with alanine and serine in order to identify the key amino acid residue that had the greatest effect on the protein – ligand interaction and to analyse the dominant molecular interactions between the *Agam*OBP20 and its potential ligand, oleic acid. Alanine and serine were used in order to determine the effect of polarity and steric hindrances. The binding behaviour of the mutated *Agam*OBP20 with the ligand was assessed through docking simulations. Results of the docking simulations were then analysed using the binding

Bereitgestellt von Slovenská poľnohospodárska knižnica | Heruntergeladen 17.01.20 14:56 UTC

free energy values that were calculated in order to determine the prevailing trend. Structure – based method was also employed in order to observe the changes in the conformation of the ligand and the surrounding residues. Results may be of use for the development of repellents and pesticides that are able to inhibit the function of the *Agam*OBP20.

2. Materials and Methods

2.1. Mutagenesis of the AgamOBP20

The sequence of the AgamOBP20 was obtained and used as a template for site-directed mutagenesis that was employed for each of the residues on the binding site. The sequence was obtained from the Research Collaboratory for Structural Bioinformatics Protein Databank (RCSB PDB), an online portal that feature free and a readily available storehouse of biomolecular 3D structures (BERMAN *et al.*, 2000). The sequence was obtained as a FASTA sequence using its protein ID 3VB1 with a resolution of 2.00 angstroms (ZIEMBA *et al.*, 2013). Each residue was replaced with alanine and serine according to their localization on the binding site. The mutated sequences for both alanine and serine were collected and were used to generate homolog model of the *Agam*OBP20.

2.2. Homology modelling

Homolog models of the AgamOBP20 mutants were produced using Swiss Model Server, a fully automated server for protein homology – modelling (BIASINI *et al.*, 2014; BORDOLI *et al.*, 2009; BORDOLI *et al.*, 2006). The templates selected were based on their identity match with the target sequence, global model quality estimation, and the quality mean. The global model quality estimation (GMQE) is a quality estimation that combines properties from the target – template alignment. The quality mean (QMEAN) is a composite scoring function that estimates the global and local model quality. Each residue is assigned a reliability score between 0 to 1 for both GMQE and QMEAN, which indicates the expected similarity with the target structure. The template with the highest value of the identity match, GMQE and QMEAN were chosen.

2.3. Ligand selection

The ligand selected for docking with the *Agam*OBP20 was oleic acid since it obtained the most favourable binding free energy value out of the major components in sweat and limburger cheese that were also docked with the *Agam*OBP20 (JANAIRO *et al.*, 2015). The ligand structure was optimized using Spartan ’14 (Wavefunction, Inc.) software at the DFT B3LYP level of theory.

Bereitgestellt von Slovenská poľnohospodárska knižnica | Heruntergeladen 17.01.20 14:56 UTC

2.4. Docking simulations

Docking simulation was then performed in order to closely observe the difference in the binding orientation of oleic acid for each of the mutated *Agam*OBP20 as well as the wild type, and to energetically determine the interactions involved. The homologs generated from the mutated sequences with alanine and serine were docked with oleic acid using ArgusLab (Planaria Software), a publicly available software that is reliable for virtual screening, docking simulations, molecular modelling, ligand pose construction and ligand pose selection (JANAIRO and JANAIRO 2012; ODA *et al.*, 2009). The simulation was confined only within 15 angstroms of the *x*, *y*, and *z* axes of the binding site with a 0.4 angstroms of grid resolution, having a regular precision and a maximum of 150 poses. Binding interaction was assessed in thermodynamic parameters using the binding free energy of the best ligand pose. The binding free energies obtained from the mutated sequence were compared with the binding free energy of the wild type. Structure-based observations were also employed to determine the changes in the conformation of the ligand and the surrounding residues.

3. Results and Discussion

The binding site of the AgamOBP20 is composed of hydrophobic residues namely M6, M7, G10, I13, M50, M53, I70, I73, M82, L86, L106, L107, L110, F117, I118, F119, and P120. Site – directed mutagenesis was conducted through homology modeling by replacing each residue with alanine and serine. Docking simulation was then employed for the mutated OBP20 and the wild type with oleic acid used as the ligand. Docking was performed in order to observe the difference in the binding of the ligand for each mutation in contrast to the binding of the ligand to the wild type. The key residues that are said to play an important role in the binding of the AgamOBP20 with its ligand were selected based on the greatest difference in the value of the calculated binding free energy of the alanine and serine mutations compared with the wild type. In order to maintain consistency, mutant OBP20 with its corresponding mutation will be denoted as mOBP20-mutation (i.e. mOBP20-LEU107A etc.)

The OBP20 with the mutated residues and the wild type are listed in Table 1 with their corresponding binding free energy after it has been docked with the potential ligand, oleic acid.

From the residues altered with alanine and serine, the following can be observed. The mOBP20-LEU107A and mOBP20 - LEU106S garnered the lowest value of binding free energy, indicating enhanced spontaneity. Taking into account the binding free energy of the wild type, which is -11.9752 kcal/mol, mOBP20-LEU107A and mOBP20-LEU106S resulted into a more favorable binding free energy.

The percentage difference of the binding free energy of the extremes with respect to the binding free energy of the wild type is shown in table 2. The alteration of leucine with both alanine and serine resulted into a binding free energy that is more favorable with the percentage difference of 12.68 % and 6.38 % from the wild type. Since both alanine and leucine are hydrophobic amino acids, it is difficult to conclude

Bereitgestellt von Slovenská poľnohospodárska knižnica | Heruntergeladen 17.01.20 14:56 UTC

that the enhanced spontaneity brought about by alanine substitution is absolutely due to hydrophobicity.

Table 1. Binding free energies of the mutated OBP20 and the wild type using oleic acid as the ligand.

Mutated OBP20 - Alanine Mutation	Binding Free Energy (kcal/mol)	Mutated OBP20 - Serine Mutation	Binding Free Energy (kcal/mol)	Wild Type	Binding Free Energy (kcal/mol)
L107A	-13.49	L106S	-12.74	OBP20	-11.98
P120A	-12.71	P120S	-12.55		
I13A	-12.65	L107S	-12.55		
M7A	-12.47	M50S	-12.39		
I70A	-12.20	I118S	-12.36		
M82A	-12.07	I73S	-12.19		
F117A	-12.04	G10S	-12.00		
F119A	-12.01	F117S	-11.97		
L110A	-12.01	M7S	-11.95		
G10A	-11.98	I13S	-11.94		
I73A	-11.98	F119S	-11.85		
M50A	-11.92	L86S	-11.83		
M6A	-11.73	M6S	-11.59		
L106A	-11.64	M82S	-11.40		
I118A	-11.59	L110S	-11.38		
L86A	-11.54	I70S	-10.82		
M53A	-11.14	M53S	-10.26		

This is further supported with the observations wherein serine mutations at L106, P120, L107, I118 and I73 resulted into a more favorable binding. This suggests that the ligand preferred a residue that is less bulky since both alanine and serine are less bulky than leucine. Steric interactions within the protein binding site have strong influence on the manner of which the ligand binds (SNYDER *et al.*, 2011; HLAVACEK *et al.*, 1999). However, the role of hydrophobicity in ligand binding cannot be discounted and still remains a crucial factor. This was observed when the values generated for alanine and serine mutations were compared.

Table 2. Percentage difference of the binding free energy values of the extremes with the wild type.

Mutated OBP20 Alanine Mutation	Percentage Difference (%)	Mutated OBP20 Serine Mutation	Percentage Difference (%)
L107A	12.68	L106S	6.38
M53A	-6.95	M53S	-14.32

It is evident that the binding free energy values generated from the mutated residues with alanine obtained lower values than that of the serine mutants indicating that the ligand preferred nonpolar residues to interact with. When M53 was altered with both alanine and serine, it resulted to a higher value of binding free energy.

Bereitgestellt von Slovenská poľnohospodárska knižnica | Heruntergeladen 17.01.20 14:56 UTC

This indicates that the affinity of the protein with the ligand did not become favorable. However, upon observing the values of other methionine residues such as M7, M82, M50 and M6, it was not consistent with the trend of M53 so it is safe to conclude that M53 garnered the highest value of binding free energy when replaced with alanine and serine due to its position with respect to the orientation of the ligand.

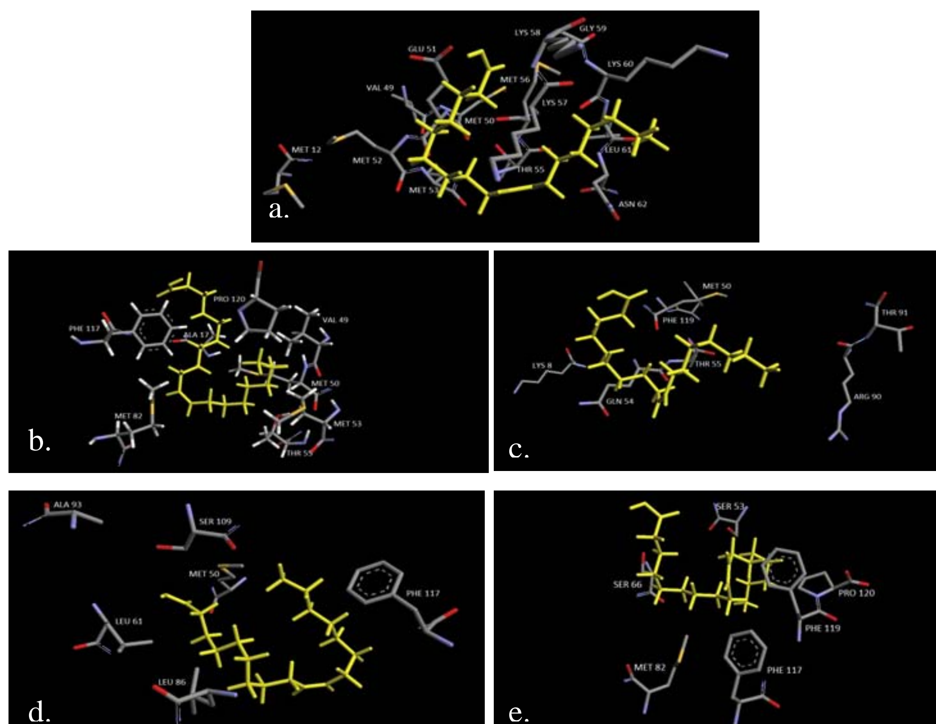

Fig. 1. Comparison of the binding of the ligand oleic acid (yellow) with the wild type (**a**) and the mutated *Agam*OBP20 (**b,c,d,e**) in ribbon model after docking. Fig. 1b shows the mutated OBP20-L107A after docking. Fig. 1c shows mutated OBP20-M53A after docking. Changes in the orientation of the residues as well as the ligand orientation in contrast to Fig. 1a is notable. Fig. 1d shows mOBP20-L106S after it has been docked with oleic acid. Fig. 1e shows mOBP20-M53S after docking. Notable variations in the ligand orientation are observed as well as with the specific residues the ligand interacts with. These conformational differences can be attributed to the mutations conducted.

Differences in ligand orientation among the mutants were thereafter assessed. The images of the mutated *Agam*OBP20 (L107A and L106S) were compared with the docked wild type (Fig. 1). Fig. 1a shows the wild type OBP20 (ribbon model) interacting with the oleic acid (yellow). The figure shows that most of the residues that are close to the ligand were the nonpolar residues such as V49, M50, and L61. However, there are also presence of polar residues such as N62 and T55. The OBP20 mutated with L107A is depicted in Fig. 1b. It can be deduced from Fig. 1b that the conformation of oleic acid is bent wherein the terminal atoms are observed to potentially express a hydrophobic interaction with F117 while the median portion of the oleic acid are shown to potentially express hydrophobic interaction

Bereitgestellt von Slovenská poľnohospodárska knižnica | Heruntergeladen 17.01.20 14:56 UTC

with nonpolar residues such as A107 and M50. It can also be depicted that the manner of which the ligand is bent may be due to the orientation of residues such as F117, bulky in nature, that are oriented towards the ligand. The OBP20 mutated with L106S is shown in Fig. 1d. It can be observed from Fig. 1d that the alkyl chain of the oleic acid is bent and is surrounded with nonpolar residues such as L61, L86, A93, M50 and F117. These residues are oriented towards the ligand that may have been the reason why the ligand bent itself in that manner. Presence of polar residues can also be noted such as S106.

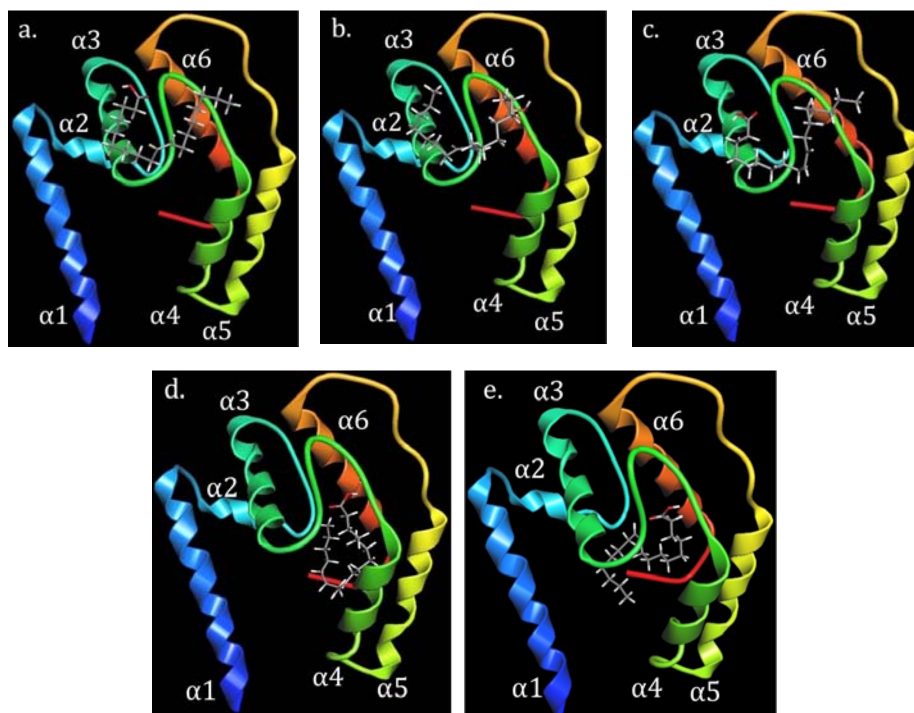

Fig. 2. Ribbon model of the wild type (**a**) and the mutated OBP20 with the extremes L107A (**b**), M53A (**c**), L106S (**d**), M53S in (**e**). Overall, variations in ligand orientation and localization within the binding pocket which can be attributed to the mutations are notable.

To observe the ligand orientation with respect to the conformation of the *Agam*OBP20, docked images of the wild type and the mutated OBP20 containing the extremes were collected (Fig. 2). The ribbon model of the wild type AgamOBP20 after docking with oleic acid is presented in Fig. 2a. It can be observed that the ligand is elongated in such a way that all parts of the chain are located on the outer most portion of the protein. It is evident that the carboxylic side is in close proximity with helices 2 and 3 where the upper middle portion of the alkyl chain is bent to such degree that it is relatively close to the lower portion of helix 3. Likewise, the terminal chain is bent in such a way that it is relatively close to the uppermost portion of helix 4.

Bereitgestellt von Slovenská poľnohospodárska knižnica | Heruntergeladen 17.01.20 14:56 UTC

The ribbon model of the OBP20 with the extremes L107A and M53A after it has been docked with oleic acid is presented in Fig. 2b and 2c, respectively. In Fig. 2b, it is clearly evident that there was a change in the conformation of oleic acid in contrast to that shown in Fig. 2c however, the ligand is still located on the outermost core of the protein. In here, the carboxylic side is now in close proximity with the upper portion of the helix 4. The alkyl chain is within the perimeter of helix 3. In Fig. 2c, where the OBP20 containing M53A was docked with oleic acid, the carboxylic side is now located near helix 3 and is elongated in such a way that the terminal side is near the proximity of helix 4. The ribbon model of the OBP20 with the extremes L106S and M53S is depicted in Fig. 2d and 2e, respectively. It is clearly evident that the ligand is now located onto the inner core of the protein. For OBP20 with the mutation L106S, the carboxylic side of the ligand is near the proximity of the helix 6. The ligand is bent to such degree that the middle portion of the alkyl chain is near the C terminal tail of the protein. The terminal tail of the ligand exists in the median core of the protein. For the OBP20 containing M53S, depicted in Fig. 2e, the ligand is also located within the inner portion of the protein where the carboxylic side is also near helix 6 however, the alkyl chain is more bent to such degree that the median portion of the alkyl chain is the located on the median core of the protein. The terminal portion of the alkyl chain exits near the C terminal tail of the protein.

4. Conclusions

In summary, the results of the calculations and modelling provided insights regarding the binding of oleic acid with *Agam*OBP20. From these calculations, key amino acids residues were identified which are hypothesized to play key roles in the binding of the ligand. These residues are L106, L107 and M53. These residues had the greatest impact in the binding free energy when these were mutated with alanine and serine. The significant change in the binding free energy of these mutants from the wild type is attributed to a decrease in steric hindrance within the binding site. These results suggest that both hydrophobic interaction and steric hindrance are crucial factors to consider in the binding of the ligand with this odorant – binding protein. The molecular features and parameters obtained may be a potential target for the new drugs and pesticides that are able to inhibit the function of *Agam*OBP20 and may result to the blocking of the olfactory mechanism of the *Anopheles gambiae*, the mosquito responsible for transmitting malaria in humans, leading to the disarray of its host seeking behaviour.

References

ARNOLD, K., BORDOLI, L., KOPP, J., SCHWEDE, T.: The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. Bioinformatics, 22(2), 2005, 195-201.

Bereitgestellt von Slovenská poľnohospodárska knižnica | Heruntergeladen 17.01.20 14:56 UTC

BERMAN, H.M., WESTBROOK, J., FENG, Z., GILLILAND, G., BHAT, T.N., WEISSIG, H., BOURNE, P.E.: The Protein Databank. Nuc. Acids Res., 28(1), 2000, 235-242.

BIASINI, M., BIENERT, S., WATERHOUSE, A., ARNOLD, K., STUDER, G., SCHMIDT, T., SCHWEDE, T.: SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. Nuc. Acids Res., 42(1), 2014, 252-258.

BIESSMANN, H., NGUYEN, Q.K., LE, D., WALTER, M.F.: Microarray-based survey of a subset of putative olfactory genes in the mosquito *Anopheles gambiae*. Insect Mol. Biol., 14, 2005, 575-589.

BORDOLI, L., KIEFER, F., ARNOLD, K., BENKERT, P., BATTEY, J., SCHWEDE, T.: Protein structure homology modelling using SWISS-MODEL Workspace. Nat. Protoc., 4, 2009, 1-13.

HLAVACEK, W.S., POSNER, R.G., PERELSON, A.S.: Steric effects on multivalent ligand-receptor binding: exclusion of ligand sites by bound cell surface receptors. Biophys. J., 76(6), 1999, 3031-3043.

JANAIRO, J.I.B., JANAIRO, G.C.: Homology modeling and comparative docking analysis of two naturally occurring pancreatic glucokinase mutants. Philipp. Sci. Lett., 5(1), 2012, 1-6.

JANAIRO, J.I.B., CARANDANG, J., VI, AMALIN, D.: Docking simulations and regression analysis on the binding of several carboxylic acids with the odorant - binding protein 20 of *Anopheles gambiae*. Rom. J. Biochem., 52(1), 2015, 61-65.

KNOLS, B.G.J., VAN LOON, J.J.A., CORK, A., ROBINSON, R.D., ADAM, W., MEIJERINK, J., DE JONG, R., TAKKEN, W.: Behavioural and electrophysiological responses of the female malaria mosquito *Anopheles gambiae* (Diptera: *Culicidae*) to Limburger cheese volatiles, Bull. Entomol. Res., 87, 1997, 151-159.

ODA, A., TAKAHASHI, O.: Validation of ArgusLab efficiencies for binding free energy calculations. Chem. Bio. Informat. J., 9, 2009, 52-61.

PELOSI, P., MAIDA, R.: Odorant binding proteins in vertebrates and insects: similarities and possible common function. Chem. Senses., 15, 1990, 205- 215.

PELOSI, P.: Odorant-Binding Proteins. Crit. Rev. Biochem.Mol. Biol., 29(3), 1994, 199-228.

PELOSI, P., TIRINDELLI, R.: Structure/activity studies and characterization of an odorant-binding protein. Receptor events and transduction in taste and olfaction. Chem. Senses, 1, 1989, 207-226.

PEVSNER, J., REED, R.R., FEINSTEIN, P.G., SNYDER, S.H.: Molecular cloning of odorant-binding protein: member of a ligand carrier family. Science, 241, 1988, 336-339.

RUND, S.S.C., HOU, T.Y., WARD, S.M., COLLINS, F.H., DUFFIELD, G.E.: Genome-wide profiling of diel and circadian gene expression in the malaria vector *Anopheles gambiae*. Proc. Natl. Acad. Sci. USA, 108, 2001, E421-E430.

SCHWEDE, T., KOPP, J., GUEX, N., PEITSCH, M.C.: SWISS-MODEL: an automated protein homologymodeling server. Nuc. Acids Res., 31(13), 2003, 3381-3385.

Bereitgestellt von Slovenská poľnohospodárska knižnica | Heruntergeladen 17.01.20 14:56 UTC

SNYDER, P.W., MECINOVIC, J., MOUSTAKAS, D.T., THOMAS, S.W., HARDER, M., MACK, E.T., LOCKETT, M.R., HÉROUX, A., SHERMAN, W., WHITESIDES, G.M.: Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase. Proc. Natl. Acad. Sci. USA, 108(44), 2011, 17889-17894.

ZHOU, J.: Odorant-binding proteins in insects. Vitam. Horm., 83, 2010, 241-272.

ZIEMBA, B.P., MURPHY, E.J., EDLIN, H.T., JONES, D.N.M.: A novel mechanism of ligand binding and release in the odorant binding protein 20 from the malaria mosquito *Anopheles gambiae*, Protein Sci., 22, 2013, 11-21.

Received *18 October 2016*
Accepted *7 December 2016*

Bereitgestellt von Slovenská poľnohospodárska knižnica | Heruntergeladen 17.01.20 14:56 UTC