# Recombinant Production of Odorant Binding Protein 1 from Aedes Aegypti (AaegOBP1) to Use as Recognition Element for Lung Cancer Biomarker Heneicosane

Asım Emir Çubukcuoğlu, Furkan Fıçıcı, Gizil Oğuz, Hande Ezgi Kepenek, Kaan Yapıcı, Ömer Faruk Güleç, Seçil Sude Öztabak, Yaren Özdemir

Istanbul Technical University

BioCrabs Team

GOGEC Competition Report

## ABSTRACT

Lung cancer has the highest mortality rate among cancers. First- and second-degree lung cancer can be hidden very well because there are no pain receptors in the lungs. Unless controlled by chance, lung cancer can often be detected when it metastasizes after the third stage. After the third and fourth stages, the recovery rates are quite low. Early diagnosis is very important here, as in other types of cancer. For this purpose, a system that can detect lung cancer from the breath was considered. The Volatile Organic Compounds (VOCs) in the breath were investigated and the molecule Heneicosane was detected to be the most specific. This molecule, unlike other VOCs, originated only from lung cancer cells. It has been found that the Heneicosane molecule works in nature as a pheromone of some species. Aedes aegypti mosquito was the selected study species in this context. In the pheromone pathway of this mosquito, odorant binding proteins (OBP), the protein that allows molecules to come together with olfactory receptors (ORs), were found. In the continuation of the study, OBP was studied in detail. It was primarily purified from mosquito antennae. It was synthesized recombinantly and purified again in usable form. E. coli BL21(DE3) strain and pET-22b plasmid vector were used for recombination. Purification was done by affinity chromatography using the His-tag method. Docking and modeling were studied in computer environment. In this research, AaegOBP1 was produced recombinantly in E. coli BL21 strain and purified with highest efficiency with the methods that are suggested. It was also concluded that, according to the results of CB-Dock visualization, heneicosane has the highest affinity to AaegOBP1 when it was compared with other pheromone and lung cancer Volatile organic compound biomarkers: Eicosane, octadecane, tridecane, undecane, and decane.

**KEYWORDS:** Lung cancer early diagnosis, volatile organic compounds (VOCs), Heneicosane, odorant binding protein (OBP), recombinant protein production and purification, His-tag, Modelling OBP1 and heneicosane affinity with CB-Dock

# 1. INTRODUCTION

Lung cancer, which causes death of 1.59 million people each year, is leading cancer-related deaths. 16% of lung cancer cases that can usually be detected at stages 3 and 4 are incurable or treatment is ineffective. As can be seen in Figure 1, lung cancer has the highest 5-year survival rate compared to colon, skin, and breast cancers. According to the National Lung Screening Study, low-dose computed tomography scans resulted in a 20% reduction in lung cancer death rates [1].

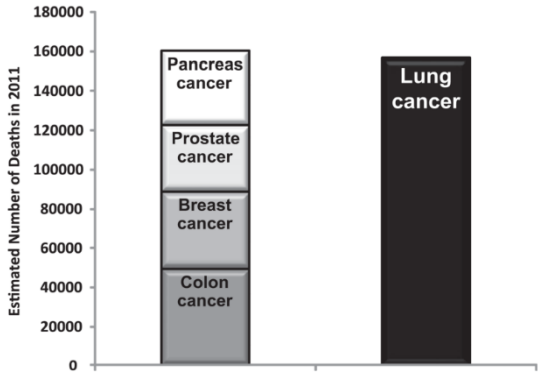


**Figure 1:** Estimated deaths from lung cancer compared with others combined.

There are 67 common volatile organic compounds (VOCs) at picomolar levels in the breath of people with lung cancer. In studies on lung cancer volatile organic compounds provided an accuracy of around 81% [2].

The most striking of the VOCs in the research was heneicosane alkane. As per the findings of the studies performed by Byun et al. and Yu et al., heneicosane is not detected in the samples collected from non-smokers and healthy individuals but is present in samples collected from people who suffer from lung cancer, it has been identified as a lung cancer identifier [3][4].

According to the studies on the larval cuticle of the *Aedes aegypti* mosquito species, heneicosane is one of the key compounds in attracting gravid mosquitoes to oviposit in treated substrates by acting as an oviposition pheromone to promote egg deposition [5]. These discoveries led to the classification of heneicosane as an *Aedes aegypti* oviposition pheromone, which is recognized by the antennal olfactory receptors of *Aedes aegypti* [6].

After the inhalation of odor molecules through the air, the olfactory receptors recognize these molecules, response to olfactory transport to brain as a signal by neurons, and a neural response occurs in the brain. The receptors could recognize odor molecule called olfactory receptors (ORs) [7][8]. It was known that odorant binding proteins (OBP) could bind to odor molecules and carry them to the receptors. On the other hand, OBP also protect the ORs because OBP could prevent reaching of too much volatile compounds to receptors [9].

Like most of volatile compounds, pheromones are also member of this group. For the insects, OBPs could ensure molecular interactions for pheromones, and they can also play role to transportation of pheromones to receptors [10]. Most of the

odorant binding proteins expressed in antenna cells of insects [11][12].

In this study, OBP protein obtained from the antennae of *Aedes aegypti* mosquito species was produced recombinantly. Recombinant proteins have critical place for biomedical industry, the pharmaceutical and treatment researches such as drug production [13]. Recombinant protein purification is essential for most structural and functional studies thus, many methods were suggested mostly based on expression of affinity tags allowing the facilitation of detection and purification. In His-tag purification, which is preferred in this project, recombinant protein is produced with six or nine histidine repeats which can be placed on either C- or N- terminus of the protein. Since, these repeated histidine residues have small size is uncharged in the condition of physiological pH, the application of His-tag has no effect on the protein folding and to the structure and function of the recombinant protein. Then the His-tagged recombinant protein can be purified by the immobilized metal ion affinity chromatography (IMAC). In IMAC method, the binding affinity of the Histidine tag to immobilized metal ions $Ni^{+2}$ or $Cu^{+2}$, is used for purification [14].

In this project, the focus is on the recombinant production of odorant binding protein in E. coli BL21(DE3) strain. Thus, this protein has been tried to purified with the His-tagged purification methods from *Aedes Aegypti*, and tried to be produced in the *E. coli* by the recombinant DNA methods.

Escherichia coli is studied intensively, especially for the last 60 years, and its genome was completely sequenced in 1997 [15][16]. This long history makes it the best studied and understood organism. Because E. coli is a well-known and simple bacterium, seen as a model organism, the "industrial workhorse", of in the fields of biotechnology and microbiology, which was also preferred in this study. E. coli can grow rapidly in simple cultures containing glucose and ammonium under laboratory conditions, with a doubling time of 20 minutes [15]. Advantage of E. coli in protein production by gene cloning is vector plasmids. Plasmids are small circular pieces of DNA that can replicate independently of the cell DNA. These parts, which provide information transfer between bacteria, carry antibacterial products or genes that provide antibiotic resistance. They can be easily modified in the laboratory and used in the production of recombinant proteins [17].

E. coli BL21(DE3) strain was preferred for the production of recombinant protein in this study, thanks to all these properties and its well-known knowledge.

***Molecular Docking***

Receptor-ligand docking is a structure-based drug design technique that has been

widely used to predict and visualize the interactions between a receptor and a ligand, for it significantly increases efficiency, shortens time and lowers research costs [18]. Docking might be performed either with or without information on the protein's binding site. When a protein's binding pocket is known, docking is carried out under the restrictions of a receptor binding site, and the algorithm places the ligand inside the binding pocket while finding the models that are structurally closest to the reference structure, that is, the most stable protein-ligand complexes. To discover the most stable composition of the complex when the binding site of the receptor is unknown, the algorithm docks the ligand with the entire surface of the protein, a process known as blind docking [19].

***Molecular Dynamics Simulations***

Computing methods called molecular dynamics (MD) simulations are computational tools which are used to study the thermal averages of molecular structures, conformational changes in molecules that are thermally accessible, and the dynamic characteristics of biomolecules in solution over a range of timescales [20] As force fields which consist of customized formulas and variables that specify a molecule's potential energy surface are used in MD simulations, their accuracy in obtaining quantitative calculations of molecular interactions and behavior under various environmental conditions is enhanced.

A number of evolving trajectories are averaged in molecular dynamics to illustrate the likely sequence of changes that would occur in a typical evolution of the molecule, or time-stepping is utilized to compute numerous snapshots that are generally used for generating a probability distribution.

***Project Value***

While cancer causes the death of approximately 10 million people in the world, lung cancer is the most common type. In 2020, while the diagnosis of lung cancer was made to approximately 2.21 million people in the world, and caused the death of 1.80 million people [21]. Early diagnosis plays a very important role in the treatment of lung cancer. In human breath there are lots of volatile organic compounds such as alkanes, alcohols, aldehydes and ketones. Previous studies show that the data is obtained from breath and cell culture by using GC-MS (gas chromatography-mass spectrometry) in different cancer types. One molecule, 21 carbon straight chain alkane heneicosane, becomes prominent that only seen in lung cancer patients, so that it is lung cancer biomarker. Heneicosane is also a pheromone in *Aedes aegypti*, the yellow

fever mosquito, and can be detected by the odorant binding proteins in the antenna of *Aedes aegypti*. In this project, the interaction between heneicosane and the odorant binding protein 1 (OBP1) are examined under molecular docking. To conclude, this project aims to change the binding site of these interaction and mutate the protein to bind more specifically to heneicosane and more strongly in order to detect the biomarker easier than other VOCs [23].

***Specific Aims***

**Aim 1:** Detection of Lung Cancer

*Sub Aim 1:* Developing a method for early detection for lung cancer

*Sub Aim 2:* Detecting lung cancer from breath with a biomarker

**Aim 2:** Presenting an alternative method for early detection of lung cancer by detection of volatile organic molecules from breath

*Sub Aim 1:* Detection of the biomarker heneicosane, which is only found specifically in lung cancer cells and can be found in breath

**Aim 3:** Amplification of the odorant binding protein of Aedes aegypti capable of detecting the molecule

*Sub Aim 1:* Isolation of cell membrane protein from antenna

Sub Aim 2: Genetically cloning the protein in E coli bacteria

**Aim 4:** Understanding the binding mechanism between odorant binding protein 1 (AaegOBP1) of *Aedes aegypti* and heneicosane

*Sub Aim 1:* Identification of key AaegOBP1 residues for heneicosane binding by assessing the binding affinities for mutations of common residues which were detected for the complexes of AaegOBP1 with octadecane, tridecane, undecane and decane.

***Human Centered Design***

GC-MS, is a big and expensive device, also not so accessible for the patients. The best-known technique for early diagnosis is lung tomography which can emit radiation and also not so accessible. Patients may use medication for screening. These old methods need an alternative approach, so that a more effective, user-friendly and cheaper concept can be used. There are lots of biosensors that detect alcohols, but not specific to lung cancer. There is no any alkane sensor, just because it is tough to detect straight chain alkanes using biosensors, because alkanes do not have functional groups. According to this aim, after a mutation specific to heneicosane, a complex of odorant receptors and odorant binding proteins can be adapted to a biosensor, which provides a more effective, accessible, mobile, sustainable, sensitive and safety way for early diagnosis [25].

***Originality***

In previous studies, volatile organic compounds include heneicosane are not

detected by using odorant binding proteins and odorant receptors for the purpose of early diagnosis of lung cancer. There are only a few studies that research lipid bilayer sensor systems [26][27]. This research gives perspective to the other researches that, AaegOBP1 can be used as a recognition element for detecting heneicosane from human's exhaled breath due to diagnose lung cancer in early stages.

***Biosafety and Biosecurity***

In this project, the *E. coli* BL21 strain is used, which is not pathogenic. Therefore, there is no need a BSL 2 or BSL 3 labs. However, to avoid the contamination the experiments must be done in a cabinet. On the other hand, instead of TRIzol the RNA isolation kit can be favorable, because of the toxicity of TRIzol. The Temed which is used by the preparation of SDS-PAGE Gel is also toxic and must be make ready for the experiment in the cabinet, additionally, experiment must be made in the cabinet.

# 2. MATERIALS AND METHODS

## 2.1 MATERIALS

In this research, for experimental design, 1000 *Aedes Aegypti* antennas, Leibovitz's medium and Grace's medium supplemented with lactalbumine hydrolysate and yeastolate, L-cysteine-activated papain protease, incubator, Pasteur pipette, Falcon Petri dishes, RNeasy Mini Kit – Qiagen, TRIS/HCl buffer which had 20 mM TRIS and pH 7 value, centrifuge tubes, centrifuge, SMART RACE cDNA Amplification Kit, forward primer (KpnI recognition sequence -*5'-GCGGGGTACCCGACGTTACTCCGC GGCGTG-3'*) and reverse primer (BamHI recognition sequence - *5'-GCGCGGATCC TTAAATCAGGAAGTAATGC-3'*), MMLV reverse transcriptase, Q5 high-fidelity DNA polymerase kit (NEB), thermal cycler, agarose powder, TBI buffer, RedSafe dye, agarose gel electrophoresis tank, 6X loading dye, DNA ladder, scalpel, vortex, PCR Clean-up Gel Extraction Kit (MN), pET-22b(+) vector, 10X NEB rCutSmart buffer (1X buffer ingredients: 50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100µg/mL BSA) which has a pH 7.8-8 value, *MscI*, *BamHI*, *KpnI* restriction enzymes, T4 DNA polymerase, dNTPs, T4 Ligase enzyme, *E. coli* BL21(DE3) strain, LB broth, 100 µg/ml ampicillin containing agar plates, plasmid-free and plasmid containing *E. coli* BL21(DE3), water bath, orbital shaker, glass spreader, ampicillin containing LB medium, 2L flask, 1mM stock of IPTG, dry ice/ethanol bath, ice/water bath, Immobilized Metal Affinity Chromatography (IMAC) with Ni-NTA resins, dd$H_2O$, washing buffer, elution buffer, Centriprep-10 filter, High-Performance Liquid Chromatography (HPLC), Hydroxyalkoxypropyl-dextran type VI resin as column material for HPLC, methanol, pH 4.5 citric acid buffer, SDS loading buffer, resolving buffer, with 0.1 % Coomassie

Brilliant Blue R-250 and CB- Dock, Protein Data Bank, PubChem, NAMD, VMD as software were needed materials.

## 2.2 METHODS

Firstly, from the antenna cells were obtained as taking 1000 antennas from female *Aedes aegypti* in 1 mL of 3 parts Leibovitz's L15 medium and 2 parts of Grace's medium, which was supplemented with lactalbumine hydrolysate and yeastolate (3+2 medium) and after that, the antennas were taken from that solution. Then, the disruption of antenna is occurred with the incubation of these antennas in L-cysteine-activated papain protease which has 1 mg/mL at 30°C for 75 minutes and then, the medium is placed into 1-1.5 mL of 3+2 medium. After that, trituration is done with Pasteur pipette gently. The enzyme activity is stopped by rinsing the antennas in 3+2 medium with 3 times before the trituration step. After that, the cell suspension which is obtained from previous steps, placed on uncoated Falcon petri dish and waited for 45 minutes for settling the cells on the petri dish. Cultured cells were incubated at 22°C for 2 to 3 weeks in humid conditions in incubator. Then, after 2 to 3 weeks, grown antenna cells are obtained and they can be isolated for further applications.

In this research it is suggested that the usage of RNeasy Mini Kit – Qiagen for RNA purification. At the beginning, the grown antenna cells were harvested with 300 µL of TRIS/HCl buffer which has pH 7 with 20 mM TRIS and were centrifuged at 500 x g for 4 minutes. Then, supernatant was discarded from the centrifuge tube and after that, the cell pellet was loosened with flicking the tube carefully. After that, for purification of RNA, the procedure was followed as it was indicated in RNeasy Mini Kit – Qiagen.

For the cDNA synthesis from the isolated and purified *Aedes Aegypti*, SMART RACE cDNA Amplification Kit was used and the cDNA synthesis procedure was followed as it was indicated in the SMART RACE cDNA Amplification Kit. For the synthesis of insert cDNA, forward primer was designed as KpnI recognition sequence - *5′-GCGGGGTACCCGACGTTACTCCGCGGCGTG-3′* and reverse primer was designed as BamHI recognition sequence - *5′-GCGCGGATCCTTAAATCAGGAAGTAATGC-3′* as it was already demonstrated in the research of Leite et al. [28]. In this research, Moloney murine leukemia virus' reverse transcriptase enzyme was used for cDNA synthesis according to the SMART RACE cDNA Amplification Kit. The cDNA synthesis reaction was incubated at 42°C for 1.5 hour in air incubator.

Then, Q5 high-fidelity DNA polymerase kit from the NEB for PCR reaction and procedure is suggested in this research. PCR reaction was set up by adding, 3 µL of cDNA template, 30.5 µL nuclease free water, 10 µL of 5X Q5 reaction buffer, 2.5 µL of each forward and reverse primers, 1 µL of dNTPs and 0.5 µL of Q5 DNA polymerase enzyme respectively according to the Q5 high-fidelity DNA polymerase kit (NEB). The reaction was incubated in thermal cycler which was adjusted to 98°C for 35-40 seconds; then, to 50-72°C for 20-30 seconds; and at the final, to 72°C for 2 minutes and 20-30 seconds. After the, PCR reaction, the reaction solution was waited at 4°C.

After PCR reaction, with the addition of 1 g agarose powder to 100 mL of TBI buffer following the melting step in microwave, 1% agarose gel was made. Then, 1 µL RedSafe was added to the agarose gel and mixed. After that, agarose gel which was melted, was placed into casting tray for turning into solid form and after this step, solid agarose gel was placed onto electrophoresis tank. Then TBI buffer was added again to the electrophoresis tank until the height of the TBI buffer was reached to 2 or 3 mm over the agarose gel. Then 5 µL of PCR product was mixed with 1 µL 6X loading dye and after the addition of 2.5 µL DNA ladder, the dyed PCR product was added to the wells of agarose gel electrophoresis. Then the agarose gel electrophoresis was run at 100 volts for 40 minutes. After that, the bands were analyzed under the UV light.

Then, the PCR product which is located to the place correspond to 401 bp in the agarose gel, was taken and purified again due to obtain insert DNA sequence since the AaegOBP1 DNA sequence has 401 bp size [29]. With the usage of scalpel, DNA fragment was cut from the agarose gel and taken in a clean tube. Then, to the clean tube, NTI buffer was added with 1:2 ratio. Incubation was done at 50°C for 5 to 10 minutes and then, vortex was applied to the sample until the agarose gel was dissolved. Then, the PCR product was purified with the usage of PCR Clean-up Gel Extraction Kit (MN).

After the purification of PCR product, restriction was applied to both PCR product and the vector which was selected as pET-22b(+) in this research. For the restriction of pET-22b(+) plasmid DNA, the reaction was set up by adding nuclease free water, 10X NEB rCutSmart buffer, plasmid DNA, *MscI* and *BamHI* restriction enzymes.

For the restriction of PCR product, the reaction was set up with the addition of nuclease free water, 10X NEB rCutSmart buffer, PCR product, *KpnI* and *BamHI*

restriction enzymes respectively. Then, this restricted insert DNA was blunted with the usage of T4 DNA polymerase with dNTP. These restricted insert DNA and plasmid DNA was purified with the usage of PCR Clean-up Gel Extraction Kit (MN) again separately. Then again purified plasmid DNA was loaded to the agarose gel electrophoresis. The wells of agarose gel electrophoresis were combined to 1 well in each row and purified plasmid DNA was mixed with 6X loading dye before the addition of purified plasmid DNA samples to the wells. DNA ladder was also added to the agarose gel electrophoresis. Agarose gel electrophoresis was run at 100 volts for 40 minutes. This step was applied for checking the vector's size after the purification. Since pET-22b(+) vector has, 5493 bp, the purified vector sample which was located around 5493 bp was taken and isolated from agarose gel with the same procedure which was explained above and purified again with PCR Clean-up Gel Extraction Kit (MN). Then, insert DNA was ligated to the vector with T4 Ligase enzyme and ligation reaction was incubated 16°C overnight and after the incubation, the solution was waited at 65°C for 10 minutes for inactivation of ligase enzyme.

Then, 2.5 µl of the prepared recombinant plasmid was inserted into 50 µl competent *E. coli* BL21(DE3) strain cells by transformation. The transformation process was completed by incubation in an orbital shaker. LB broth was preferred as the medium in the stages. For the antibiotic resistance test, 100 µg/ml agar containing ampicillin antibiotic, and 100 µl each plasmid-free and plasmid containing *E. coli* BL21(DE3) cells, as control groups, were used. 50 µl of frozen competent *E. coli* BL21(DE3) strain sample was taken and left to thaw. Since the cells were taken from -80°C, they were allowed to thaw slowly by placing them on ice again, as they could be damaged if thawed suddenly at room temperature. Immediately after thawing, 2.5 µl of recombinant plasmid was added to the tube. The mixture was left on ice for a further 30 seconds. Then, the tube was incubated for 45 minutes in a hot water bath at 42 ºC. At the end of the incubation, the tube was again placed on ice to provide heat shock. 250 µl of LB medium was added to the sample and incubated for 1 more hour at 37 ºC at 250 rpm in an orbital shaker. At this point, it is assumed that recombinant cells are obtained at the end of the transformation process. However, this assumption should be confirmed by a control experiment. This confirmation was provided by the antibiotic gene found in the plasmid. 3 plates were prepared with 100 µg/ml agar containing ampicillin antibiotic. To the first, 100 µl of plasmid-free *E. coli* BL21(DE3) cells were added and cultivated by spreading with a

glass spreader, as a negative control group. In the second, 100 µl of cells with plasmid were cultivated as a positive control group. On the third plate, 100 µl cultivated cells were inoculated in the cells prepared by transformation. After drying, the plates were incubated at 37 ºC for 16 hours.

After the growing, single colony onto the plate was collected by inoculation loop, and it was transferred into tube includes LB medium ampicillin medium. Tube was placed into orbital shaker at 200 rpm and incubated at 37 °C overnight to obtain starter culture.

After the incubation, starter culture was transferred into 2 L flask include 1 L of LB Broth ampicillin medium and it was incubated at 37 °C until the cells reach to mid-exponential phase; thus, optical density of culture at 600 nm ($OD_{600}$) reaches to 0.6 [30]. Then, in order to induce gene expression 1 ml of 1mM stock of IPTG was added into culture and incubated for 3 hours [31]. After that, cells transferred into microcentrifuge tubes and they were centrifuged at 9000 rpm for 30 minutes to precipitate cell and supernatant part was removed.

Free-thaw method was chosen for the cell lysis. According to literature, this method ensures the highest efficiency for purification periplasmic proteins. Freeze-thaw cycle was performed as submerging of tube include pelleted cells in dry ice/ethanol bath for 2 minutes to freeze and submerging in 0 °C ice/water bath for 8 minutes to thaw three times and after the final thaw pellet part was resuspended in 20 mM pH 8.0 Tris-HCl buffer. Then, sample was centrifuged at 16.000xg for 10 minutes. The supernatant part could be used in purification steps [32]. In the Immobilized Metal Affinity Chromatography (IMAC) part, Ni-NTA resin was placed into gravity flow column and washed with $ddH_2O$. Then, resin was equilibrated with lysis buffer by flow through column. Ni-NTA resin and clear lysate were mixed in column (closed lid) and transferred to microcentrifuge tube, and incubated on a shaker with ice for 30 minutes. The sample was centrifuged at 1000 rpm for 1 minute, and then pH 8.0 washing buffer was added on tube and sample was centrifuged at 1000 rpm for 1 minute, supernatant was removed and addition of washing buffer on pellet and centrifugation part repeated one more. Then, same step was repeated with elution buffer 1, elution buffer 2 and elution buffer 3 instead of washing buffer respectively. The flow through after each washing and elution step was collected in different tubes, collected sample after third elution will be used.

**Table 1:** The ingredients and order of flow of mobile phases were used in His-Tag Purification.

| | Mobile Phases | pH | Ingredients |
|---|---|---|---|
| **His-tag** | Lysis Buffer | 8.0 | 50 mM $NaH_2PO_4$, 300 mM NaCl, 10 mM imidazole |
| | Washing Buffer | | 50 mM $NaH_2PO_4$, 300 mM NaCl, 20 mM imidazole |
| | Elution Buffer 1 | | 50 mM $NaH_2PO_4$, 300 mM NaCl, 100 mM imidazole |
| | Elution Buffer 2 | | 50 mM $NaH_2PO_4$, 300 mM NaCl, 100 mM imidazole |
| | Elution Buffer 3 | | 50 mM $NaH_2PO_4$, 300 mM NaCl, 100 mM imidazole |

Sample was filtered with Centriprep-10. Desalting procedure was performed, 5-ml HiTrap desalting columns were used and water was a mobile phase. In order to obtain purest protein, multi-step chromatography was preferred.

After the His-Tag purification, in order to obtain more pure protein also High-Performance Liquid Chromatography was used. Hydroxyalkoxypropyl-dextran type VI resin was used as a column material, it was mixed with HPLC grade methanol and placed into glass column. Then, resin was washed with methanol, and equilibrated with 50 mM pH 4.5 citric acid buffer. Protein sample was dissolved in 50 mM pH 4.5 citric acid buffer in a centrifuge tube, and equilibrated resin was added on the sample. The mixture was incubated for 1 hour by stirrer at room temperature. After the incubation, resin & sample mixture transferred to glass column and elution was performed with 50 mM pH 4.5 citric acid buffer.

In order to analyze purity of protein, SDS-Page method could be used. Samples was denatured in SDS loading buffer as a 1 sample / 4 buffer ratio, for 15 minutes at 95 °C. %12 resolving gel is suitable for 15 kda protein subunit. Running could be performed at 120 mW. Staining procedure was performed at room temperature for 3 hours with 0.1 % Coomassie Brilliant Blue R-250. Destaining was performed with %40 methanol, %10 acetic acid solution in water. On the other hand, Liquid Chromatography-Electrospray Ionization/Mass Spectroscopy could be an alternative for high resolution assay to determine purity of protein.

To inspect the interactions between odorant binding proteins (OBPs) of Aedes aegypti and the volatile organic compounds (VOCs), heneicosane, eicosane, tridecane, undecane, decane, and octadecane, the blind docking method was used for there are known gaps in the literature regarding the binding mechanism between the VOCs and the OBPs of Aedes aegypti.

The CB-Dock web server was used to perform blind docking and visualize the

results of these processes. The crystal structure of the odorant binding protein 1 (AaegOBP1) of Aedes aegypti was retrieved from Protein Data Bank with the identification code 3K1E at 1.85 Å resolution and separated from the coordinates of the ligand, then blind docked with the structural data files of heneicosane, eicosane, tridecane, undecane, decane, and octadecane, retrieved from PubChem, which contains information about chemical compounds [33] [34] [18].

As the docking results of lung cancer non-specific VOCs and AaegOBP1 obtained from the CB-Dock server were further examined for similar binding activities, TRP114, LEU76, PHE15, and LEU80 residues were discovered to be common residues to bind to AaegOBP1. As heneicosane was intended to be used as a biomarker it was aimed to prevent lung cancer non-specific VOCs from binding to AaegOBP1. Thus, W114A, L76A, F15A, and L80A mutations were suggested to decrease the binding affinity between lung cancer non-specific VOCs and AaegOBP1. W114, L76, F15, and L80 residues of the FASTA sequence for AaegOBP1 were replaced with A114, A76, A15, A80, and the His-tag motif was eliminated. The artificial intelligence (AI) tool Alphafold 2 was used to determine the protein structure of the generated mutations [35] [36]. Using the CB-Dock server, each mutant was docked with heneicosane, eicosane, tridecane, undecane, decane, and octadecane [34][18].

To assess the binding affinity for each one of the complexes these VOCs form with AaegOBP1 under different environmental conditions, Molecular Dynamics simulations were utilized. For each mutation, the top 5 docking models were selected to further continue the Molecular Dynamics steps were selected to further continue the Molecular Dynamics steps.

The MD simulations involved two steps for minimization and equilibrium and were carried out using Nanoscale Molecular Dynamics (NAMD) at pH 7 and pH 7.5, 300 K, 0.15 mol/L KCl concentration, and non-bond interactions cutoff set to 12 for 100 ns [37] [38]. Following the minimization run, equilibrium MD simulations for mutant and VOC docking models were performed.

## 3. RESULTS

Eicosane, octadecane, tridecane, undecane, and decane molecules are also considered as volatile organic compounds and can be used as a biomarker for lung cancer when exhaled breath sample was taken from humans. These molecules also serve as pheromones in *Aedes Aegypti* and can bind to OBP1 with different affinities.

Upon inspection of the top-ranked docking models for VOCs and 3K1E, heneicosane

was seen to have the best Autodock Vina score of -7.5, followed by Autodock Vina scores of -7.4, -7.1, -6.4, -6 and -5.6 belonging to the docking models of eicosane, octadecane, tridecane, undecane, and decane.

Therefore, based on the Autodock Vina docking scores for these docking models, heneicosane demonstrated a greater binding affinity than the alternatives.

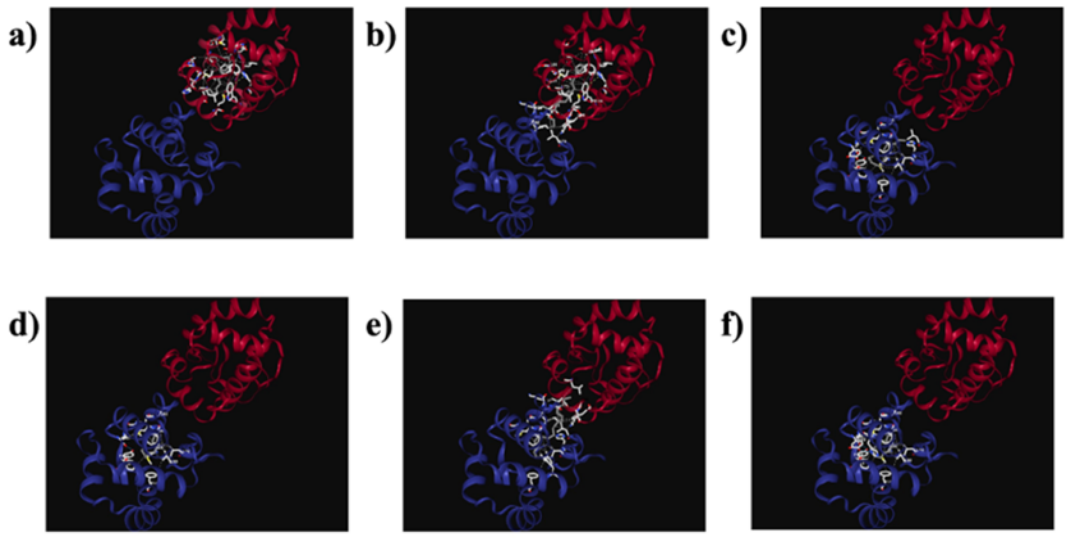


**Figure 1**: **(a)** CB-Dock visualization of the top-ranked model for blind docking of heneicosane to odorant binding protein 1 of Aedes aegypti (AaegOBP1). **(b)** CB-Dock visualization of the top-ranked model for blind docking of eicosane to AaegOBP1. **(c)** CB-Dock visualization of the top-ranked model for blind docking of tridecane to AaegOBP1. Heneicosane and eicosane were observed to bind mainly with Chain A, as tridecane interacted with Chain B of the receptor [18]. **(d)** CB-Dock visualization of the top-ranked model for blind docking of decane to AaegOBP1. **(e)** CB-Dock visualization of the top-ranked model for blind docking of octadecane to AaegOBP1. **(f)** CB-Dock visualization of the top-ranked model for blind docking of undecane to AaegOBP1. Decane and undecane were observed to bind strictly with Chain B, as octadecane interacted with both Chain A and Chain B of the receptor. TRP11, LEU76, PHE15, and LEU80 residues were observed to be common residues for the binding of octadecane, tridecane, undecane, and decane to AaegOBP1 [18].

## REFERENCES


1. Dela Cruz, C. S., Tanoue, L. T., & Matthay, R. A. (2011). Lung Cancer: Epidemiology, Etiology, and Prevention. *Clinics in Chest Medicine*, *32*(4), 605–644. https://doi.org/10.1016/j.ccm.2011.09.001

2. Phillips, M., Gleeson, K., Hughes, J. M. B., Greenberg, J., Cataneo, R. N., Baker, L., & McVay, W. P. (1999). Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study. The Lancet, 353(9168), 1930–1933. https://doi.org/10.1016/s01406736(98)07552-7

3. Byun, H. G., Yu, J. B., Huh, J. S., & Lim, J. O. (2014). Exhaled Breath Analysis System based on Electronic Nose Techniques Applicable to Lung Diseases. *Hanyang Medical Reviews*, *34*(3), 125. https://doi.org/10.7599/hmr.2014.34.3.125

4. Yu, J., Wang, D., Wang, L., Wang, P., Hu, Y., Ying, K., Pardo, M., & Sberveglieri, G. (2009). Detection of Lung Cancer with Volatile Organic Biomarkers in Exhaled Breath and Lung Cancer Cells. *AIP Conference Proceedings*. https://doi.org/10.1063/1.3156506

5. Seenivasagan, T., Sharma, K. R., Sekhar, K., Ganesan, K., Prakash, S., & Vijayaraghavan, R. (2008). Electroantennogram, flight orientation, and oviposition responses of Aedes aegypti to the oviposition pheromone n-heneicosane. *Parasitology Research*, *104*(4), 827–833. https://doi.org/10.1007/s00436-008-1263-2

6. Gonzalez, P. V., González Audino, P. A., & Masuh, H. M. (2014). Electrophysiological and behavioural response of Aedes albopictuston-heinecosane, an ovipositional pheromone of Aedes aegypti. Entomologia Experimentalis Et Applicata, 151(3), 191–197. https://doi.org/10.1111/eea.12184

7. Archunan G. (2018). Odorant Binding Proteins: a key player in the sense of smell. *Bioinformation*, *14*(1), 36–37.

8. Pelosi P. (1994). Odorant-binding proteins. *Critical reviews in biochemistry and molecular biology*, *29*(3), 199–228.

9. Başaran, M. M., & BACANLI, M. (2020). Drug Induced Taste and Smell Disorders. *Turkish Journal of Family*

*Medicine and Primary Care*, *14*(1), 153-162.

10. Zhou, J.-J., Robertson, G., He, X., Dufour, S., Hooper, A. M., Pickett, J. A., … Field, L. M. (2009). Characterization of Bombyx mori Odorant-binding Proteins Reveals that a General Odorant-binding Protein Discriminates Between Sex Pheromone Components. *Journal of Molecular Biology, 389*(3), 529–545.

11. Liu, Z., Liang, X.-F., Xu, L., Keesey, I. W., Lei, Z.-R., Smagghe, G., & Wang, J.-J. (2020). An Antennae-Specific Odorant-Binding Protein Is Involved in Bactrocera dorsalis Olfaction. *Frontiers in Ecology and Evolution, 8*.

12. Calvello, M., Brandazza, A., Navarrini, A., Dani, F. R., Turillazzi, S., Felicioli, A., & Pelosi, P. (2005). Expression of odorant-binding proteins and chemosensory proteins in some Hymenoptera. *Insect Biochemistry and Molecular Biology, 35*(4), 297–307.

13. Pham, P. V. (2018). *Medical Biotechnology. Omics Technologies and Bio-Engineering, 449–469.* doi:10.1016/b978-0-12-804659-3.00019-1

14. Spriestersbach, A., Kubicek, J., Schäfer, F., Block, H., & Maertens, B. (2015). *Purification of His-Tagged Proteins. Methods in Enzymology, 1–15.* doi: 10.1016/bs.mie.2014.11.003

15. Asefa, M. *et al.* (2020) "Escherichia coli as Model Organism and its Application in Biotechnology: A Review," *Global Journal of Biotechnology & Biochemistry*, 15. Available at: https://doi.org/10.5829/idosi.gjbb.2020.15.01.141135.

16. Idalia, V.-M. N., & Bernardo, F. (2017). *Escherichia coli as a Model Organism and Its Application in Biotechnology. Escherichia Coli - Recent Advances on Physiology, Pathogenesis and Biotechnological Applications.* doi:10.5772/67306

17. Cronan, J. E. (2014). *Escherichia coli as an Experimental Organism. eLS.* doi: 10.1002/9780470015902.a0002026.pub2

18. Liu, Y., Grimm, M., Dai, W. T., Hou, M. C., Xiao, Z. X., & Cao, Y. (2019). CB-Dock: a web server for cavity detection-guided protein–ligand blind docking. *Acta Pharmacologica Sinica*, *41*(1), 138–144. https://doi.org/10.1038/s41401-019-0228-6

19. Jofily, P., Pascutti, P. G., & Torres, P. H. M. (2021). Improving Blind Docking in DOCK6 through an Automated Preliminary Fragment Probing Strategy. *Molecules*, *26*(5), 1224. https://doi.org/10.3390/molecules26051224

20. Hansson, T., Oostenbrink, C., & van Gunsteren, W. (2002). Molecular dynamics simulations. *Current Opinion in Structural Biology*, *12*(2), 190–196. https://doi.org/10.1016/s0959-440x(02)00308-1

21. World Health Organization. (2022). Cancer. https://www.who.int/news-room/factsheets/detail/cancer

22. Thriumani, R., Zakaria, A., Hashim, Y. Z. H. Y., Jeffree, A. I., Helmy, K. M., Kamarudin, L. M., ... & Persaud, K. C. (2018). A study on volatile organic compounds emitted by in-vitro lung cancer cultured cells using gas sensor array and SPME-GCMS. *BMC cancer*, *18*(1), 1-17.

23. Seenivasagan, T., Sharma, K. R., Sekhar, K., Ganesan, K., Prakash, S., & Vijayaraghavan, R. (2009). Electroantennogram, flight orientation, and oviposition responses of Aedes aegypti to the oviposition pheromone n-heneicosane. *Parasitology Research*, *104*(4), 827-833.

24. Yu, J., Wang, D., Wang, L., Wang, P., Hu, Y., & Ying, K. (2009, May). Detection of lung cancer with volatile organic biomarkers in exhaled breath and lung cancer cells. In AIP Conference Proceedings (Vol. 1137, No. 1, pp. 198-201). American Institute of Physics.

25. Bajtarevic, A., Ager, C., Pienz, M., Klieber, M., Schwarz, K., Ligor, M., ... & Amann, A. (2009). Noninvasive detection of lung cancer by analysis of exhaled breath. *BMC cancer*, *9*(1), 1-16.

26. Kolesov, D. V., Ivanova, V. O., Sokolinskaya, E. L., Kost, L. A., Balaban, P. M., Lukyanov, K.A., ... & Bogdanov, A. M. (2021). Impacts of OrX and cAMP-insensitive Orco to the insect olfactory heteromer activity. Molecular Biology Reports, 48(5), 4549-4561.

27. Misawa, N., Fujii, S., Kamiya, K., Osaki, T., Takaku, T., Takahashi, Y., & Takeuchi, S. (2019). Construction of a biohybrid odorant sensor using biological olfactory receptors embedded into bilayer lipid membrane on a chip. ACS sensors, 4(3), 711-716.

28. Leite, N. R., Krogh, R., Xu, W., Ishida, Y., Iulek, J., Leal, W. S., & Oliva, G. (2009). Structure of an odorant-binding protein from the mosquito Aedes aegypti suggests a binding pocket

covered by a pH-sensitive “Lid”. *PLoS One*, *4*(11), e8006.

29. Sengul, M. S., & Tu, Z. (2010). Expression analysis and knockdown of two antennal odorant-binding protein genes in Aedes aegypti. *Journal of Insect Science*, *10*(1).
30. Fazaeli, A., Golestani, A., Lakzaei, M., Rasi Varaei, S. S., & Aminian, M. (2019). *Expression optimization, purification, and functional characterization of cholesterol oxidase from Chromobacterium sp. DS1. PLOS ONE, 14(2), e0212217.*
31. Zhang, Z., Kuipers, G., Niemiec, Ł., Baumgarten, T., Slotboom, D. J., de Gier, J.-W., & Hjelm, A. (2015). *High-level production of membrane proteins in E. coli BL21(DE3) by omitting the inducer IPTG. Microbial Cell Factories, 14(1).*
32. Johnson, B. H., & Hecht, M. H. (1994). Recombinant Proteins Can Be Isolated from E. coli Cells by Repeated Cycles of Freezing and Thawing. Nature Biotechnology, 12(12), 1357–1360.
33. H.M. Berman, K. Henrick, H. Nakamura (2003) Announcing the worldwide Protein Data Bank Nature Structural Biology 10 (12): 980.
34. Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B. A., Thiessen, P. A., Yu, B., Zaslavsky, L., Zhang, J., & Bolton, E. E. (2023). PubChem 2023 update. Nucleic Acids Res., 51(D1), D1373–D1380. https://doi.org/10.1093/nar/gkac956
35. Jumper, J., Evans, R., Pritzel, A. et al. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021).
36. Mirdita, M., Ovchinnikov, S., Steinegger, M. ColabFold - Making protein folding accessible to all bioRxiv 2021.08.15.456425.
37. Humphrey, W., Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics", J. Molec. Graphics, 1996, vol. 14, pp. 33-38.
38. James C. Phillips, David J. Hardy, Julio D. C. Maia, John E. Stone, Joao V. Ribeiro, Rafael C. Bernardi, Ronak Buch, Giacomo Fiorin, Jerome Henin, Wei Jiang, Ryan McGreevy, Marcelo C. R. Melo, Brian K. Radak, Robert D. Skeel, Abhishek Singharoy, Yi Wang, Benoit Roux, Aleksei Aksimentiev, Zaida Luthey-Schulten, Laxmikant V. Kale, Klaus Schulten, Christophe Chipot, and Emad Tajkhorshid. Scalable molecular dynamics on CPU and GPU architectures with NAMD. Journal of Chemical Physics, 153:044130, 2020. doi:10.1063/5.0014475