Interaction of N-3-oxododecanoyl homoserine lactone with transcriptional regulator LasR of *Pseudomonas aeruginosa*: Insights from molecular docking and dynamics simulations

Hovakim Grabski*[1], Lernik Hunanyan[1], Susanna Tiratsuyan[1, 2], Hrachik Vardapetyan[1]

[1]Bioinformatics Group, Russian-Armenian University, Yerevan, Armenia

[2]Faculty of Biology, Yerevan State University, Yerevan, Armenia

* to whom correspondence should be addressed.

Email hovakim.grabski@rau.am

Phone: +374 98 73 02 94

E-mail addresses: hovakim.grabski@rau.am, lernik.hunanyan@rau.am, susanna.tiratsuyan@rau.am, hrachik.vardapetyan@rau.am

ABSTRACT

In 2017 World Health Organization announced the list of the most dangerous superbugs and among them is *Pseudomonas aeruginosa*, which is an antibiotic resistant opportunistic human pathogen. . The central problem is that it affects patients suffering from AIDS, cystic fibrosis, cancer, burn victims etc. *P. aeruginosa* creates and inhabits surface-associated biofilms. Biofilms increase resistance to antibiotics and host immune responses, because of that current treatments are not effective. It is imperative to find new antibacterial treatment strategies against *P. aeruginosa*, but detailed molecular properties of the LasR protein are not clearly known to date. In the present study we tried to analyze the molecular properties of the LasR protein as well as the mode of its interactions with autoinducer (AI) the N-3-oxododecanoyl homoserine lactone (3-O-C12-HSL). We performed docking and molecular dynamics (MD) simulations of the LasR protein of *P. aeruginosa* with the 3-O-C12-HSL ligand. We assessed the conformational changes of the interaction and analyzed the molecular details of the binding of the 3-O-C12-HSL with LasR. A new interaction site of the 3-O-C12-HSL with the beta turns in the short linker region (SLR) of LasR protein was found. We have also performed LasR monomer protein docking and found a new form of dimerization. This study may offer new insights for future experimental studies to detect the interaction of the autoinducer with the beta turns in the short linker region (SLR) of LasR protein and a new interaction site for drug design.

Keywords: AMR, LasR, 3-O-C12-HSL, autoinducer, SLR, homology, docking, dynamics

Abbreviations: PDB, Protein data bank; MD, Molecular Dynamics; PCA, Principal Component Analysis; PC, Principal Component; 3-O-C12-HSL, N-3-oxododecanoyl homoserine lactone; AI, autoinducer; SLR, Short Linker Region; BLAST, Basic local alignment search tool; DBI, David-Bouldin Index; psF, pseudo-F statistic.

INTRODUCTION

Pseudomonas aeruginosa (*P. aeruginosa*) is a Gram-negative, monoflagellated, obligatory aerobic bacteria (1, 2). It can be found in many diverse environments such as soil, plants, hospitals, etc. (1, 3). *P. aeruginosa* is an opportunistic human pathogen, because it rarely infects healthy persons. The central problem is that it affects patients suffering from AIDS, cystic fibrosis, cancer and burn victims (2, 4). Most of the deaths caused by cystic fibrosis are due to this pathogen (1, 5). The pathogenicity of *P. aeruginosa* occurs due to the synthesis of virulence factors such as proteases, hemolysins, exotoxin A, production of antibiotic pyocyanin, Hydrogen Cyanide (HCN), secretion systems of Types 1 (T1SS), 2 (T2SS), 3 (T3SS), 5 (T5SS), and 6 (T6SS) (6), ramnolipids and biofilm formation by this organism (7). Biofilm formation is characteristic to nearly all bacteria where cell-to-cell communication occurs as the population density increases in the human body during pathology. This system is called quorum sensing (QS) that uses hormone-like molecules called autoinducers (AI) that are accumulated in the extracellular matrix. When a threshold is reached, the AIs bind to its cognate receptor and then a response regulator modulates gene expression of QS genes virulence, which includes adaptation, colonization, antibiotic resistance, plasmid conjugation, etc. Among *P. aeruginosa* the AI type 1 system is represented by LuxI/LuxR-typed proteins. AI-1 diffuses freely through the bacterial membrane and binds to the transcriptional activator LuxR. The latter has two LuxI/LuxR-type systems, the first LasI that produces 3-O-C12-HSL and the second, RhlI that synthesizes C4-HSL, both AHLs regulate virulence and biofilm formation. Thus, in *P. aeruginosa* the operon called hcnABC is responsible of HCN biosynthesis through enzyme HCN synthase (8). Exposure to HCN can lead to neuronal necrosis through the inhibition of cytochrome c oxidase, the terminal component of the aerobic respiratory chain (9, 10). Three transcriptional regulators that perform in a cluster (8) control the transcription of hcnABC genes through LasR, ANR and RhlR (11).

Nonetheless it was proposed that LasR was the crucial activator of hcnABC genes through mutagenesis experiments (11). The transcriptional activator protein LasR regulates the target gene expression by recognizing a conserved DNA sequence termed as lux box (12, 13). LasR has two domains, 1) ligand binding domain at N-terminus (LBD)

and 2) DNA binding domain at C-terminus (DBD) (14). In DBD LasR has a DNA binding Helix-Turn-Helix (HTH) motif. (15). Binding of AI 3-O-C12HSL stabilizes LasR and promotes its dimerization, thereby contributing to the resulting LasR-AI homodimer complex to contact the promoter of the target DNA and activate gene transcription (14).

During colonization and invasion, both the pathogen and the host are exposed to molecules released by each other like bacterial AIs or hosts stress hormones and cytokines. The mechanisms and receptors that might be involved in cross-talk between microbial pathogens and their hosts are not well described to date. LuxR homologues studies have demonstrated that they are homodimers, and consist of two domains. These two functional domains are joined by a short linker region (16, 17).

There remains a need for understanding of the LasR monomer, because till date there is no molecular detail information about LasR monomer behaviour. Hence, we analyzed the molecular details of the interactions of the 3-O-C12-HSL with LasR protein. So far this is the first report that shows that the 3-O-C12-HSL can interact as well with the beta turns in the short linker region (SLR) of LasR. This study may be utilized for the development of new therapeutics against *P. aeruginosa* by targeting both the LBD as well as the beta turns in the SLR of LasR and inhibit activation of genes.

MATERIAL AND METHODS

Analysis of LasR protein sequence

The methodology was based on a previous study (1). The Amino acid sequence of LasR protein of *P. aeruginosa* was obtained from UniprotKB (Uniprot ID P25084). The crystal structure of the LBD of LasR protein (amino acid residues 7 to 169) from *P. aeruginosa* was acquired from the Protein Data Bank (PDB) (18) (PDB ID: 3IX3). However, the entire structure of LasR protein was needed in order to have a better understanding of LasR protein properties.

For this reason the amino acid sequence was used to search in the PDB databank using BLAST (19) to identify suitable templates for homology modeling. The crystal structure

of QscR, bound to 3-O-C12-HSL was found to be the best fit (PDB code 3SZT) (16). The full list of similar structures obtained from BLAST search are shown in Table 1.

Table 1 List of similar structures to LasR obtained from BLAST search (1).

PDB ID	Sequence identity (%)	Query coverage (%)	E-value	Structure
3SZT	30	97	2×10^{-26}	Quorum Sensing Control Repressor, QscR, Bound to N-3-oxo-dodecanoyl-L-Homoserine Lactone
1FSE	37	28	2×10^{-4}	Crystal structure of Bascillus subtilis regulatory protein gene
4Y15	30	76	3×10^{-16}	Crystal structure of Sdia in complex with 3-oxo-c6-homoserine Lactone
3ULQ	32	25	0.09	Crystal structure of the Anti Activator Rapf complexed with the response regulator coma DNA binding domain

Reconstruction of LasR monomer

HHPred web server (20) server was used for the homology modeling of the full LasR protein. The templates used for the homology modeling of LasR are 3SZT, 3IX3, 1FSE and 3ULQ. The aforementioned templates were used to reconstruct the final model of LasR. The final reconstructed model of LasR protein was verified using PROCHECK (21) (Figure S1, S2 of the Supporting Information), Verify3D (22) and Ramachandran

Plots. More than 96.65% of the amino acid residues in the final reconstructed model had a 3D–2D score > 0.2 as indicated by Verify3D (Figure S3 of the Supporting Information) for a good computational model (22). Ramachandran Plot (Figure S1 of the Supporting Information) revealed that no amino acid residues were found in the disallowed regions. The PROSA value of this model was -6.69 which suggests that the quality of the homology model is good (23), although the obtained model is slightly different (1), it is still viable.

Constructing 3-O-C12-HSL model

The 3D model of 3-O-C12-HSL (Figure 1) was acquired from PubChem (24) to study the interactions of the ligand with LasR.

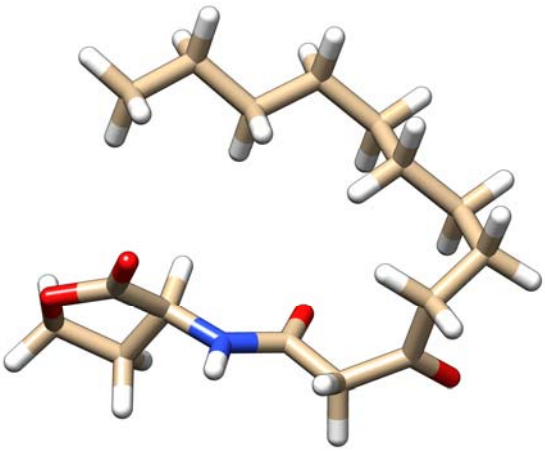

Figure 1: The stick representation of the 3O-C12-HSL AI ligand.

The ligand parameters were calculated for the General Amber Force Field (25) by using the acpype tool (26) with AM1-BCC partial charges (27).

Molecular Dynamics Simulations of LasR protein systems

MD simulations of all systems were conducted with the GROMACS suite, version 5.1.2 (28), utilizing the Amber ff99SB-ILDN force field (29).

In all cases, Short-range non-bonded interactions were cut off at 1.4 nm. Particle Mesh Ewald (PME) (30, 31) was used for the calculation of long-range electrostatics.

Set 1. *LasR monomer simulation.* In order to generate the starting structure of LasR monomer before docking, it was placed in a dodecahedron box of TIP3P water (32), to which 100 mM NaCl was added, including neutralizing counter-ions. Following two steepest descents minimization, LasR monomer was equilibrated in two stages. The first stage involved simulating for 200 ps under a constant volume (NVT) ensemble. The second stage involved simulating for 200 ps under a constant-pressure (NPT) for maintaining pressure isotropically at 1.0 bar. Production MD simulation was conducted for 100 ns in the absence of any restraints. Temperature was sustained at 300 K using V-rescale (33) algorithm. For isotropical regulation of the pressure the Parrinello-Rahman barostat (34) was used.

Set 2. Molecular dynamics simulations using docked structures. It has been shown that docking has its limitations (35). For this reason after finishing molecular docking simulations we chose the structures of LasR bound to 3-O-C12-HSL to perform molecular dynamics (MD) simulations. A time step of 2 fs was used during heating and the production run and coordinates were recorded every 1 ps. Three simulations of 300 ns were performed. Further details about the simulation protocol can be found above.

LasR–3-O-C12-HSL ligand blind docking experiments

To build the LasR–ligand complex, the ligand 3-O-C12-HSL was docked with LasR monomer using Autodock Vina (36). However, AutoDock Vina currently uses several simplifications that affect the obtained results. The most notable simplification as the creators' note (36) is the use of a rigid receptor. Vina provides a parameter called 'Exhaustiveness' to change the amount of computational effort used during a docking experiment (36, 37). But the default exhaustiveness value is 8 and the creators claim that it should increase the time linearly and decrease the probability of not finding the minimum exponentially (36, 37), hence increasing this value leads to an exponential increase in the probability of finding the energy minima.

The whole protein conformational space was searched, using grid box dimensions 60×62× 48 A˚. Following exhaustiveness values were tested in this study: 8, 16, 32, 64, 128, 256, 512, 1024, 2048 and 4096.

Principal component (PC) (38) and cluster analysis using K-means algorithm (39) was performed (Figure 2 and Figure S4 and Figure S5 of the Supporting Information). The results demonstrate that the number of interaction sites don't change in the interval using exhaustiveness from 1024 to 4096. Exhaustiveness value of 1024 was chosen as it provides good results, good speed and thorough sampling of the docked configurations.

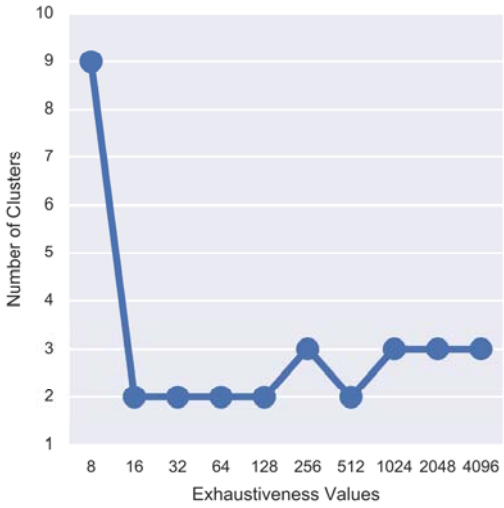

Figure 2: Exhaustiveness determination by number of clusters.

Exhaustiveness value was increased to a value of 1024 and maximum number of binding modes to generate set to 20. After that 100 independent docking calculations were carried out with random initial seeds.

Analysis of docking conformations and trajectories

We performed the analysis of docking conformations and trajectories by a combination of Autodock Vina (36), Gromacs (28) in addition to custom python scripts, which uses pandas (40), scikit-learn (41) and MDTraj (42).

Here is the list of program packages we used for analysis:

- The MDTraj python library (42) was used for the trajectory analysis.
- Plot visualization was done using matplotlib (43) and Seaborn package (44).

- Figures and videos were prepared with PyMOL (45), VMD (46) and UCSF Chimera (47).

The analysis protocol approach was similar to the works of Wolf et al. (48) and Zamora et al. (49) and the following techniques were used for the analysis:

- Principal component analysis (PCA) is an unsupervised statistical technique that is often used to make data easy to explore and visualize. PCA tries to explain the maximum amount of variance with the fewest number of principal components (38). The process of applying PCA to a protein trajectory is called Essential Dynamics (ED) (49-51). PC analysis, performed on Cartesian coordinates has become an important tool for the examination of conformational changes.

- Cluster analysis is another unsupervised technique that tries to identify structures within the data. It is a data exploration tool for dividing a multivariate dataset into groups. Clustering algorithms can be grouped into partitional and hierarchical clustering approaches (52).

NMR calculations

Finally, the trajectory of LasR monomer simulation was used for the calculation of chemical shifts. Sparta+ (53) and ShiftX2 (54) were used to predict the chemical shifts of backbone atoms of the protein with the help of wrapper functions of MDTraj (42).

Protein-Protein Docking

When docking homology models, it is best if there is an experimental evidence to suggest the general interaction site (within ~10 Å). Representative structures from molecular dynamics simulations were used for protein-protein docking using Cluspro (55). From the experimental X-ray data it was found that ‘Trp152’, ‘Lys155’ and ‘Asp156’ from H10 play an important role during dimerization.

The distances between ‘Trp152’ from chain A and ‘Asp156’ from chain B was 0.276 nm, ‘Asp156 from chain A and ‘Lys155’ from chain B was 0.279nm. These residues were used as attraction constraints.

Entire flowchart

The whole methodology is presented as a flowchart for a better comprehension:

- LasR protein sequence taken from UniprotKB (sequence ID P25084) and performed blastp of LasR to identify suitable templates for reconstruction.
- HH-pred web server was used for the reconstruction of LasR monomer structure and validated using Verify 3D, Procheck, Prosa.
- The 3D model of 3-O-C12-HSL acquired from pubchem web server.
- Docking of 3-O-C12-HSL ligand with the LasR monomer performed using Autodock Vina.
- PCA and cluster analysis performed on docking conformations.
- Extraction of centroid conformations from cluster analysis.
- Ligand parameters generated using Acpype interface in the framework of the AMBER force field.
- Using centroid conformations as starting points for molecular dynamics simulations using Gromacs.
- Analysis of molecular dynamics trajectory files using MDTraj.
- Performed protein-protein docking using ClusPro.

RESULTS

Compactization of LasR without 3-O-C12-HSL

We performed a simulation run of 100 ns using a standard MD protocol for the assessment of conformational changes of LasR monomer (Movie 1 of the Supporting

Information). The overall stability of the molecule was assessed using the root-mean-square-deviation (RMSD) of the protein atoms. RMSD was calculated with reference to the initial frame through time of MD run (Figure S6 of the Supporting Information). Another suitable measurement for the stability of the LasR monomer structure is radius of gyration (Figure S7 of the Supporting Information).

During the examination of MD trajectories, Principal Component Analysis (PCA) (37) is usually used for the visualization of the motions of the system. Generally to capture more than 70% of the variance in the trajectory data the first handful of components are sufficient (49). PC analysis can uncover the fundamental movements contained in an MD trajectory, however it does not group the snapshots into different clusters (52). This can be accomplished by the clusterization of the PC data.

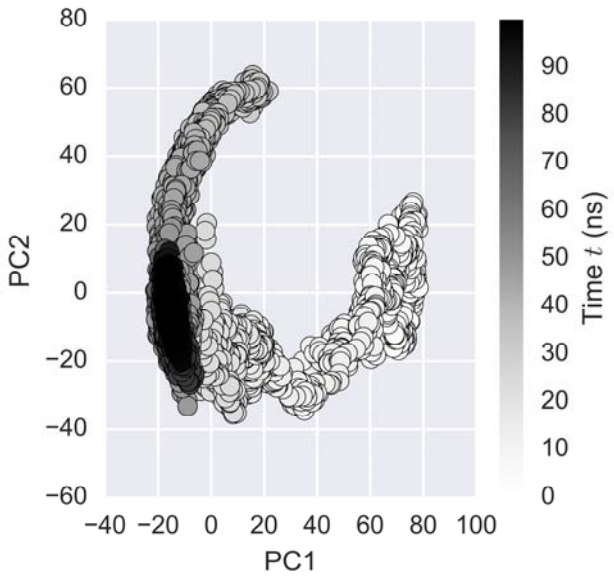

Figure 3: LasR monomer simulation trajectory projected onto its first two principal components. The black scale indicates the time progression from 0 ns (white) to 100 ns (black).

For the selection of initial starting structure of the LasR for docking study we performed cluster analysis on the LasR monomer run for the selection of a starting point. By identifying a distinct representative structure of the most populated cluster, this will allow to perform blind docking on the whole structure. It should be also noted that cluster

analysis allows to evaluate the frequent conformations of LasR. For the clustering analysis, we chose the Agglomerative Clustering algorithm with Ward linkage (56) as the most appropriate as they are deterministic, allowing for reproducibility of resulting clusters, thus minimizing the amount of bias.

We performed several rounds of Agglomerative clustering with Ward linkage (details can be found in the methods section). The accuracy of the clustering was assessed by the help of the Davies–Bouldin Index (DBI) (57), Silhouette Score (58), Dunn Index (59) and the pseudo-F statistic (pSF or Calinski Harabasz) (60) metrics (Figure S8 of the Supporting Information). An optimal number of clusters were chosen, simultaneously accounting for a low DBI, high Silhouette, high Dunn Index and high pSF values.

The distribution of clusters over simulation is visualized in Figure 4 and the four cases are: cluster 4 (black) at the beginning of the simulation (after equilibration), cluster 2 (green) in the middle, cluster 1 (light green) and cluster 3 (dark blue) at its end.

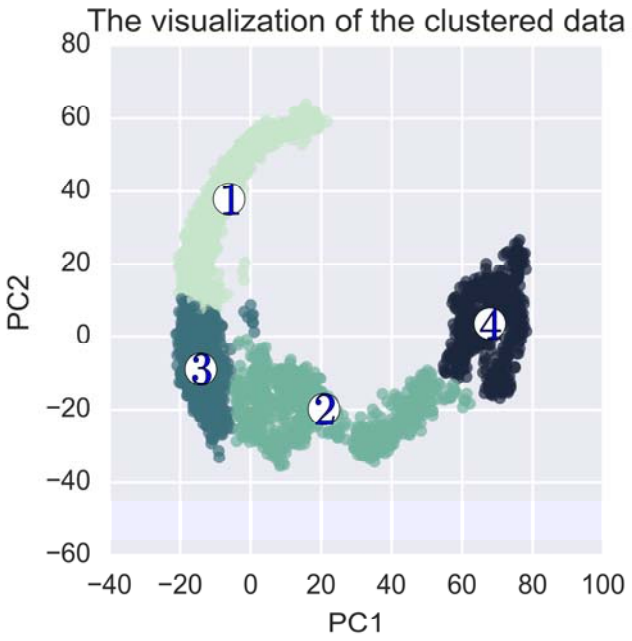

Figure 4: Clustering results of ward-linkage algorithm formed by first two PCs. The entire MD trajectory data was used for the clusterization.

The clusterization defined by the first two PCs (Figure 4) provides a coherent picture and it also supported by a good DBI, Dunn, Silhouette Score and psF value (Figure S8 of the

Supporting Information). It also shows clearly that the simulation has converged (Figure 5).

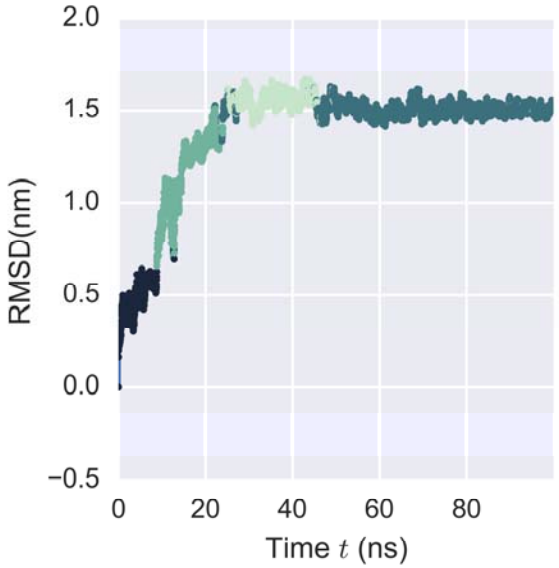

Figure 5: Color-coded RMSD of the simulation obtained from Ward-linkage cluster analysis. The colors represent the clusters that are demonstrated in Figure 4.

For the validation of the quality of molecular dynamic simulation, theoretical NMR shifts were calculated using Sparta+ (53) and ShiftX2 (54). For the NMR shifts calculation, snapshots from cluster 3 (Figure 4 and Figure 5) were used.

In Figure 6 you can see that there is a strong linear relationship between experimental and simulated ShiftX2 NMR shift values (Figure S9 of the Supporting Information), thus this demonstrates the quality of MD simulation.

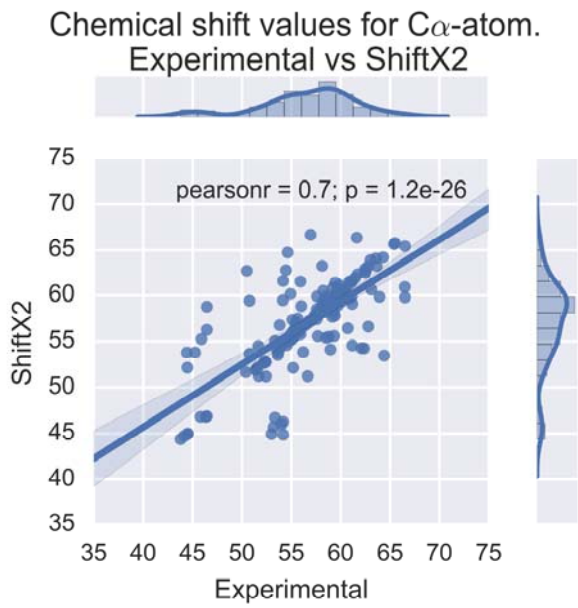

Figure 6: Chemical shifts of Cα- atoms. Experimental vs ShiftX2 predictions.

In Figure 7 you can see that there is a strong linear relationship as well between experimental and simulated Sparta+ NMR shift values (Figure S9 of the Supporting Information).

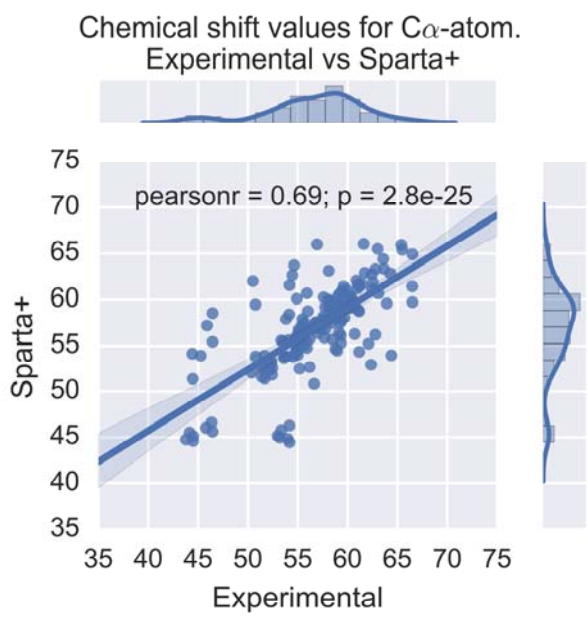

Figure 7: Chemical shifts of Cα- atoms. Experimental vs Sparta+ predictions.

After that for docking study the representative structure (Figure 8c) was extracted from cluster 3 (Figure 4 and Figure 5).

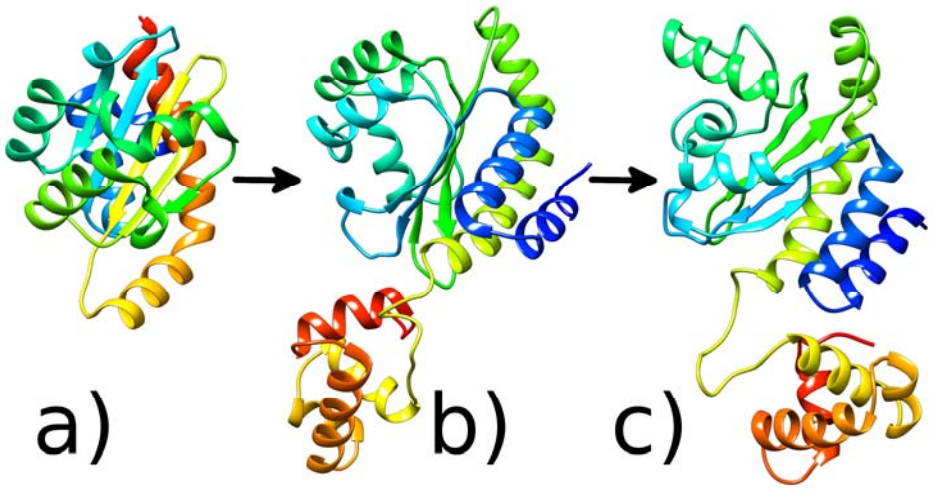

Figure 8: 3D LasR protein structures: a) Crystal structure of the N-terminal LBD of LasR protein b) Homology Modeling c) Representative Structure after 100ns MD run. Images generated with UCSF Chimera (47).

The representative structure of LasR system and its differences from the homology modeled structure are highlighted in Figure 8c. Secondary structure analysis of the representative structure was performed as well using PDBSum (61) (Figure S10 of the Supporting Information).

To our opinion, this is the most complete study to include the dynamics of the full-length LasR molecule (residues 1 to 239). It has also been shown that the dynamics of the complete C-terminal region of LasR modulate N-terminal region, as will be discussed later.

Docking analysis of with LasR monomer

In this study, molecular docking approach was used to inspect the possible binding modes of 3-O-C12-HSL in LasR monomer. PCA and cluster analysis were performed on docking data (Figure 8c). The results show three binding sites, cluster 2 corresponds to

experimental data, while cluster 1 and cluster 3 do not (Figure 9). These results clearly support the findings of Bottomley et al. (14), where it was also demonstrated that 3-O-C12-HSL binds to LBD.

We performed several rounds of K-Means clustering (details are available in the section of methods). The accuracy of the cluster analysis was evaluated using the DBI (57), Dunn Index (58), Silhouette score (59) and the pSF (60) metrics (Figure S10 of the Supporting Information). An optimal number of clusters were chosen for docking results, simultaneously accounting for a low DBI, high Dunn, high Silhoeutte and high pSF values.

We generated 2000 docked poses and performed representative structure extraction for use in MD simulations of the LasR 3-O-C12-HSL binding sites. The resulting representative structures from each cluster are shown in Figure 10. These cluster representative structures were produced by finding the centroid conformations.

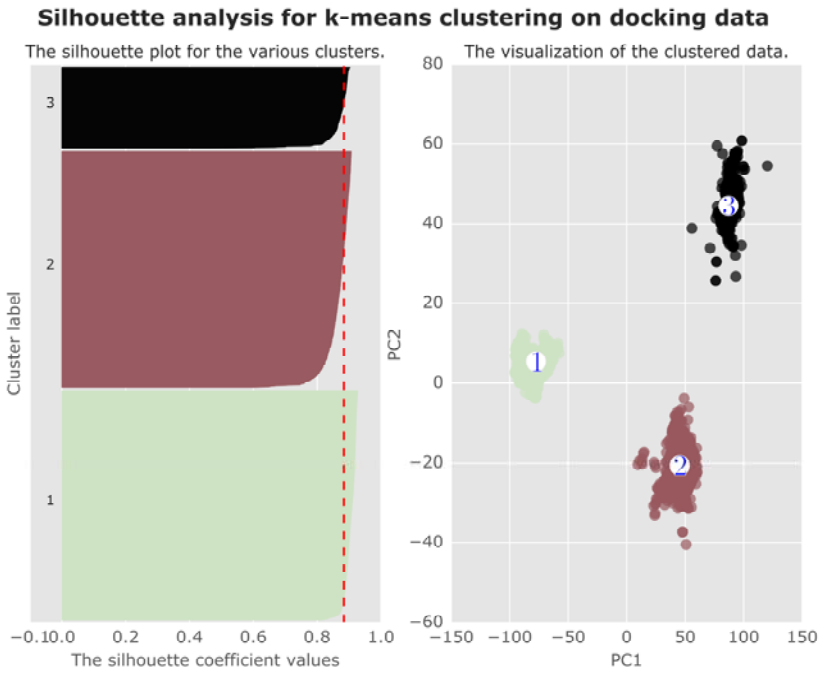

Figure 9: Silhouette plot and clustering results on the docking data using k-means algorithm formed by first two PCs.

Representative structures from each cluster were extracted. The binding energy for representative structure of cluster 1 is -5.4 kcal/mol, as for the mean binding affinity for the whole cluster is -5.257 ±0.233 kcal/mol (Figure S12 of the Supporting Information). Cluster 1 contains 839 docked poses from 2000, about 41.95%. For cluster 2 the binding affinity for the representative structure is -5.1 kcal/mol and for the whole cluster -5.593 ±0.386 kcal/mol (Figure S13 of the Supporting Information) and this one corresponds to the experimental binding site data (14). Cluster 2 contains 864 docked poses from 2000, about 43.2%. For cluster 3, the representative structure features the highest binding affinity -5.7 kcal/mol and for the whole cluster -5.264 ±0.27 kcal/mol (Figure S14 of the Supporting Information). Cluster 3 contains 297 docked poses from 2000, about 14.85%, which is a rather unstudied area.

There is a huge amount of literature that suggests that molecular docking are not appropriate for the prediction of binding affinity or binding poses of protein-ligand complexes, however they can still provide important information (35, 62).

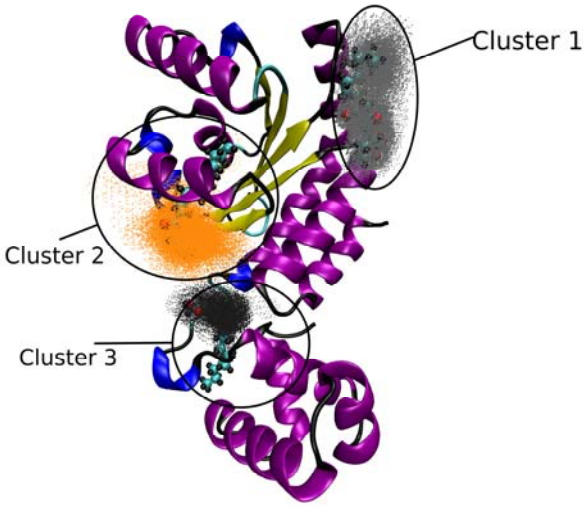

Figure 10: 3D visualization of the analyzed docking data with their representative structures and clusters.

Binding Modes of 3-O-C12-HSL

We performed three 300 ns simulations using standard MD protocol. Overall, 900ns of simulation data was used for the analysis of 3-O-C12-HSL interaction with LasR monomer. The representative structures were taken from docking results (Figure 10) and used as starting points for MD simulation with LasR. Simulations were conducted for sufficient time to allow the positions of 3-O-C12-HSL to reach equilibrium in LasR molecule.

The overall stability of the molecule was assessed using the mass-weighted root-mean-square-deviation (RMSD) of the backbone atoms. RMSD was calculated with reference to the initial snapshot for the different independent MD runs. Figure 11 shows that Simulation 2 and 3 experience a substantial RMSD deviation from the initial starting point. Simulation 2 corresponds to cluster 2 in docking simulations, while Simulation 3 corresponds to cluster 3 (Figure 10). Simulation 1 which corresponds to cluster 1, the molecule of 3-O-C12-HSL did not fixate and reach equilibrium, so further research was not performed (Figure 11) (Movie 2 of the Supporting Information).

Simulation 2 (Movie 3 of the Supporting Information) shows that after 230ns, the structure becomes stable. While Simulation 3 (Movie 4 of the Supporting Information) changes its conformation in 100ns and becomes stable till 300ns.

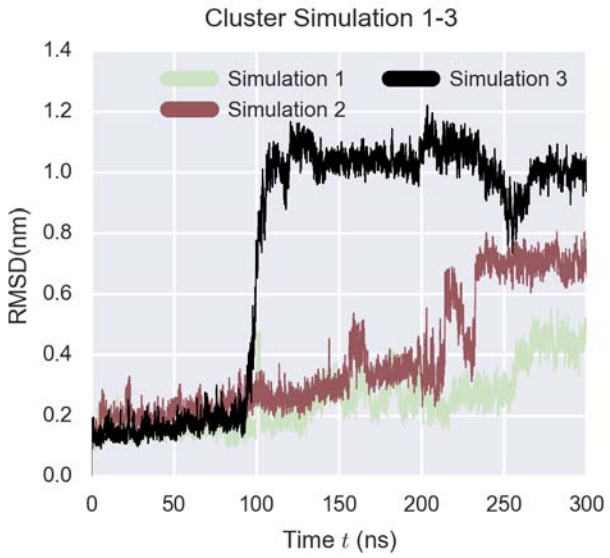

Figure 11: RMSD evolution during 300 ns of simulation for 3 independent runs using representative structures of 3-O-C12-HSL docking poses as starting points.

The root-mean-square fluctuation (RMSF) was used for the assessment of the flexibility of the LasR monomer. In Figure 12 the average per-residue RMSF for each cluster simulation runs is plotted.

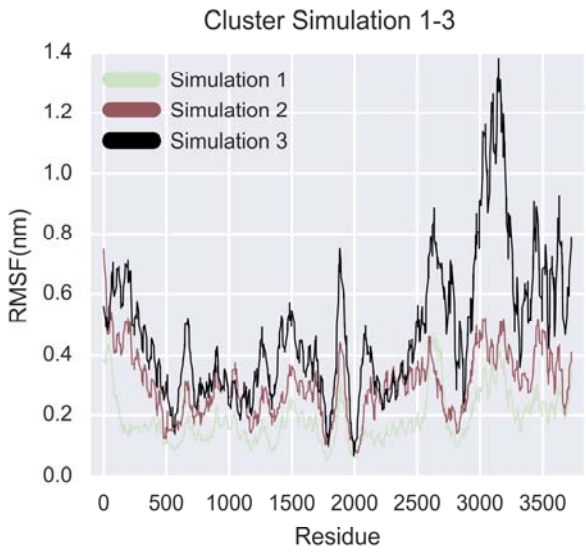

Figure 12: RMSF comparison for each individual run.

From Figure 12 it visible that the residues from 165 to 176, which correspond to beta turns in the SLR of LasR, are of high mobility (Figure S10 of the Supporting Information). Simulations 2 and 3 show that 3-O-C12-HSL has two binding modes, one with LBD, which corresponds to experimental data and simulation 3 with the beta turns region (Figure S10 of the Supporting Information) in the SLR of LasR.

PCA and Cluster analysis were performed on simulation 2 (Figure S15 of the Supporting Information) as well as hydrogen bond analysis based on cutoffs distance and angle according to the criterion of Wernet Nilson (63) using MDTraj (42). Over the course of cluster 1, 3-O-C12-HSL with LasR established an average of 0.655±0.651 of hydrogen bonds, while in cluster 2 the average is 0.042±0.202 (Figure S16 of the Supporting Information). Over the course of simulation 2, 3-O-C12-HSL establishes a large number of hydrophobic contacts with amino acid side chains of the LBD of LasR protein (Figure

S13 of the Supporting Information), a phenomenon that is not unexpected, given the large hydrophobic surface area of the LasR LBD and the low solubility of 3-O-C12-HSL in water. In simulation 2, 3-O-C12-HSL has hydrophobic interactions mainly with amino acids from H8, H10 and A strand (Figure S17 from supplementary material). RMSD analysis of the conformations between LasR monomer and LasR bound to 3-O-C12-HSL in LBD is equal to 7.027 Ångström (Figure 13).

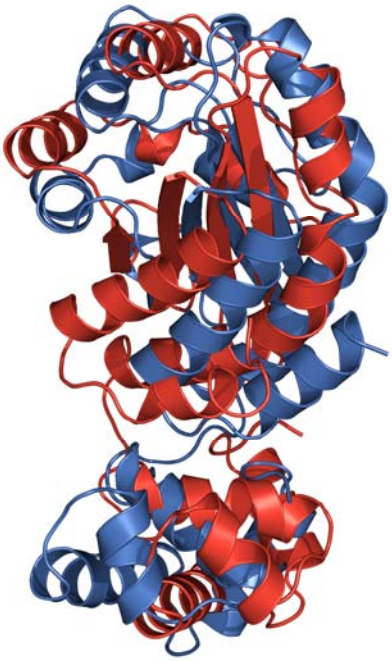

Figure 13: LasR monomer structures blue color) LasR monomer without 3-O-C12-HSL, red color) LasR monomer with 3-O-C12-HSL bound to LBD.

Residues that participate in hydrophobic interactions are shown in Figure 14 and Figure S17 of the Supporting Information.

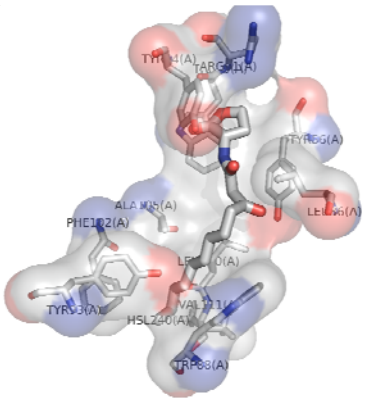

Figure 14: Insertion of 3-O-C12-HSL in LBD of LasR, from centroid snapshot of cluster 2 in simulation 2 (Figure S15 of the Supporting Information).

The second binding mode involves the interaction of 3-O-C12-HSL with the beta turns in the region of 170-175 amino acids.

PCA, cluster and hydrogen bond analysis were also performed on simulation 3 (Figure S14 of the Supporting Information). Over the course of cluster 2, 3-O-C12-HSL with LasR established an average of 1.506±0.742 hydrogen bonds, for cluster 3 the average 0.228±0.492 (Figure S13 of the Supporting Information), while in cluster 1 the average is 0.652±0.654. Over the course of simulation 3, 3-O-C12-HSL establishes hydrogen bonds and large number of hydrophobic contacts with amino acid side chains in the beta turns in the SLR of LasR protein (Figure S15 of the Supporting Information). In simulation 3, 3-O-C12-HSL forms hydrogen bonds mainly with ‘Lys182’ and ‘Leu177’ of the beta turns in the SLR of LasR (Figure S15 of the Supporting Information). RMSD analysis of the conformations between LasR monomer and LasR bound to 3-O-C12-HSL in DBD is equal to 1.677 Ångström (Figure 15).

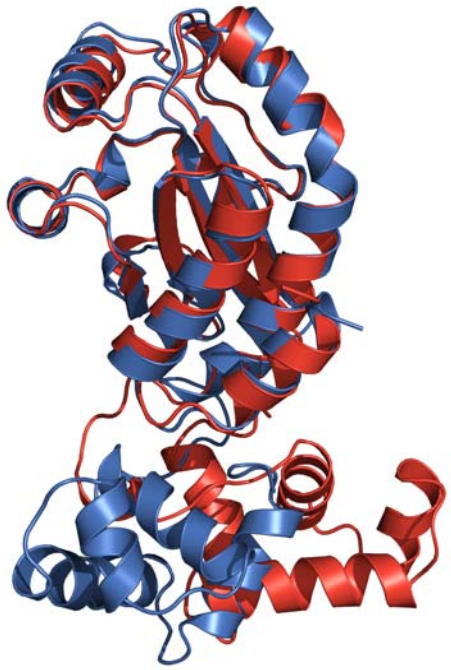

Figure 15: LasR monomer structures. blue color) LasR monomer without 3-O-C12-HSL, red color) LasR monomer with 3-O-C12-HSL bound to beta turns in the SLR of LasR.

Residues that participate in hydrogen and hydrophobic interactions are shown in Figure 16.

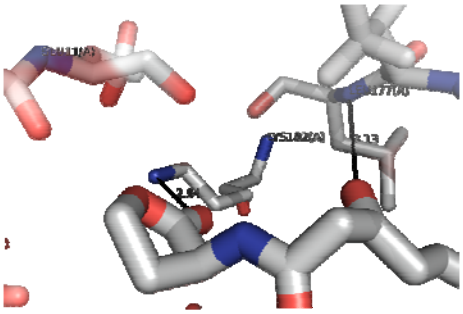

Figure 16: Putative hydrogen bonds of 3-O-C12-HSL with Lys182 and Leu177 from centroid snapshot of cluster 1 of simulation 3 (Figure S18 of the Supporting Information).

Complete data sets for MD simulations are available as Supporting Information (Figures S15-S20).

Protein Docking of MD models

It is known that to LasR binds to the corresponding promoter DNA region as a dimer (1). Thus, the monomeric LasR–3-O-C12-HSL complexes were subjected to dimerization. The protein docking experiments using structures from MD runs were performed with ClusPro 2.0 (55, 64–66) because of its success in the CAPRI (Critical Assessment of Predicted Interaction). Each centroid conformation was extracted from simulations 2 and 3 and was used for docking. For the selection of the model we used the approach as recommended by the authors of ClusPro (55), suggesting to find the most populated clusters.

From simulation 2 the top model contains 122 members and the scores for the docking model were -1440.7 for the center and -1517.9 for the lowest energy, suggesting a favorable binding mode (Figure 17).

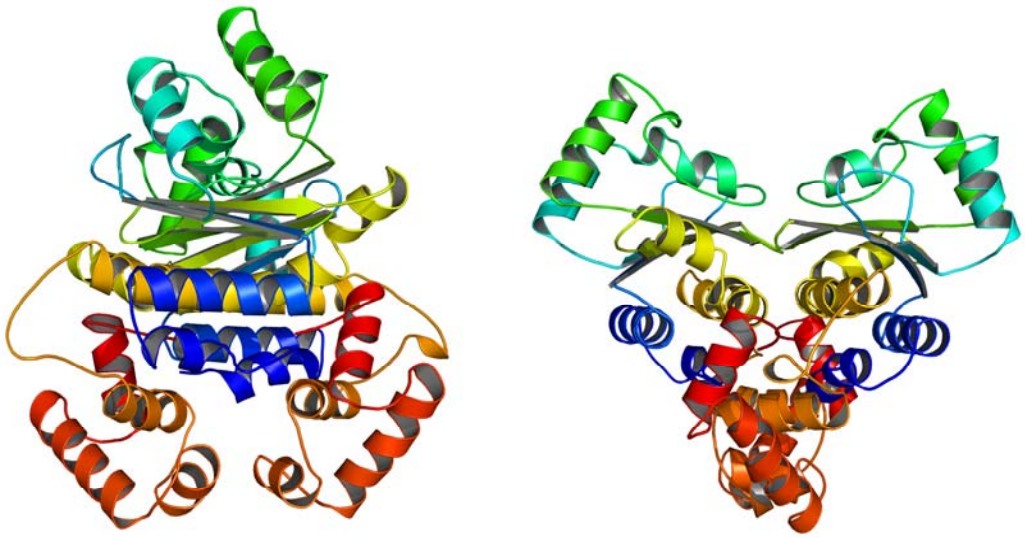

Figure 17: LasR Protein-Protein docking result based on centroid conformation from cluster 2 of simulation 2 (3-O-C12-HSL with LBD). Left and front views.

For simulation 3 docking the top model contains 80 Members and the scores for the docking model were -951.4 for the Center and -1332.0 and for the lowest energy, thus suggesting a favorable binding mode (Figure 18).

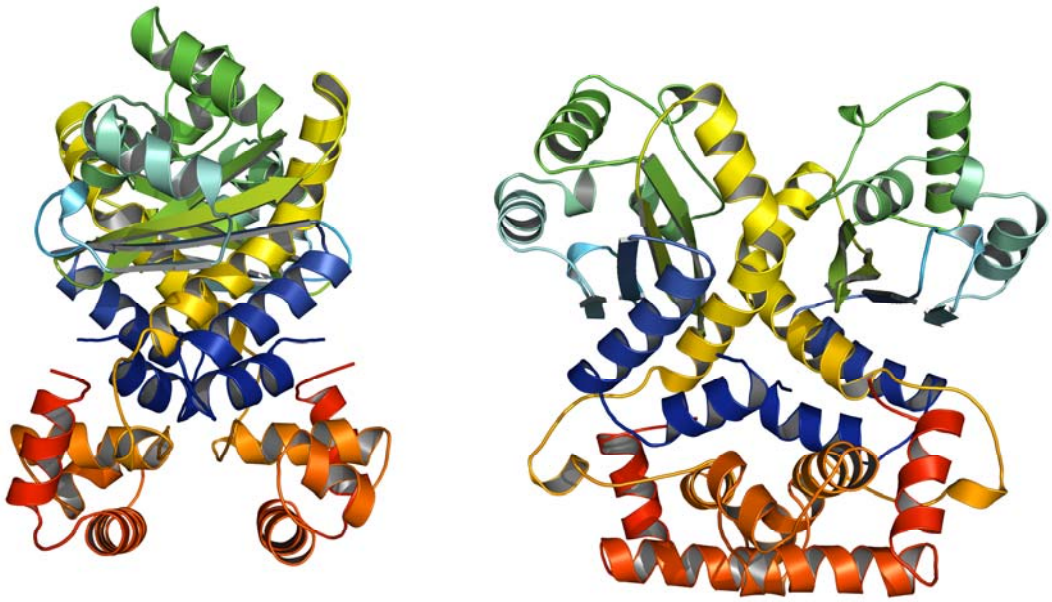

Figure 18: LasR Protein-Protein docking result based on centroid conformation from cluster 3 of simulation 3 (3-O-C12-HSL with beta turns in the SLR). Left and front views.

The interaction of 3-O-C12-HSL with the beta turns in the region of 170-175 creates a favorable conformation for dimerization. From our analysis, it is observed that 3-O-C12-HSL has two binding modes to LasR protein. The 3-O-C12-HSL, AI ligand (Figure 1) having a long hydrophobic tail is capable of binding both to LBD and to the beta turns in the SLR of LasR protein. The binding of 3-O-C12-HSL provokes conformational transitions. The two binding conformations are capable of dimerization.

CONCLUSION

From the simulations it can be safely concluded that the AI ligand 3-O-C12-HSL, can bind to LBD and to the beta turns in the SLR of LasR protein. The interaction with the beta turn is a novel site. Both conformational transitions favor dimerization, so this raises the question what would be the roles of different binding sites, it could be a two-step activation or resistance mechanism. This part needs further research. This study may reveal new insights of the interactions of the autoinducer 3-O-C12-HSL ligand with LasR protein. Results from this study may be used for future drug development endeavors.

ACKNOWLEDGMENTS

We thank Yerevan Physics Institute for providing time on the cluster for the molecular dynamics simulations.

Funding: This work was supported by grant NIR 25/15 of the Russian-Armenian University.

REFERENCES

1. Chowdhury, N., and Bagchi, A. (2016) Molecular insight into the activity of LasR protein from Pseudomonas aeruginosa in the regulation of virulence gene expression by this organism. Gene, 580(1), 80-87.

2. Botzenhart, K., and Döring, G. (1993) Ecology and epidemiology of Pseudomonas aeruginosa. In Pseudomonas aeruginosa as an Opportunistic Pathogen (pp. 1-18). Springer US.

3. Lederberg, J. (2000) Infectious history. Science, 288(5464), 287-293.

4. Irvin, R. T. (1993) Attachment and colonization of Pseudomonas aeruginosa: role of the surface structures. In Pseudomonas aeruginosa as an Opportunistic Pathogen (pp. 19-42). Springer US.

5. Govan, J. R., and Deretic, V. (1996) Microbial pathogenesis in cystic fibrosis: mucoid Pseudomonas aeruginosa and Burkholderia cepacia. Microbiological reviews, 60(3), 539-574.

6. Filloux, A. (2011) Protein secretion systems in Pseudomonas aeruginosa: an essay on diversity, evolution, and function. Frontiers in microbiology, 2.

7. Van Delden, C., and Iglewski, B. H. (1998) Cell-to-cell signaling and Pseudomonas aeruginosa infections. Emerging infectious diseases, 4(4), 551.

8. Laville, J., Blumer, C., Von Schroetter, C., Gaia, V., Défago, G., Keel, C., and Haas, D. (1998) Characterization of the hcnABC gene cluster encoding hydrogen cyanide synthase and anaerobic regulation by ANR in the strictly aerobic biocontrol agent Pseudomonas fluorescens CHA0. Journal of Bacteriology, 180(12), 3187-3196.

9. Solomonson, L. P. (1981) Cyanide as a metabolic inhibitor. Cyanide in biology/edited by B. Vennesland.

10. Niquet, J., Baldwin, R. A., Allen, S. G., Fujikawa, D. G., and Wasterlain, C. G. (2003) Hypoxic neuronal necrosis: protein synthesis-independent activation of a cell death program. Proceedings of the National Academy of Sciences, 100(5), 2825-2830.

11. Pessi, G., and Haas, D. (2000) Transcriptional Control of the Hydrogen Cyanide Biosynthetic Genes hcnABC by the Anaerobic Regulator ANR and the Quorum-Sensing Regulators LasR and RhlR inPseudomonas aeruginosa. Journal of Bacteriology, 182(24), 6940-6949.

12. Fuqua, C., and Greenberg, E. P. (1998) Self perception in bacteria: quorum sensing with acylated homoserine lactones. Current opinion in microbiology, 1(2), 183-189.

13. Egland, K. A., and Greenberg, E. P. (1999) Quorum sensing in Vibrio fischeri: elements of the luxI promoter. Molecular microbiology, 31(4), 1197-1204.

14. Bottomley, M. J., Muraglia, E., Bazzo, R., and Carfì, A. (2007) Molecular insights into quorum sensing in the human pathogen Pseudomonas aeruginosa from the structure of the virulence regulator LasR bound to its autoinducer. Journal of Biological Chemistry, 282(18), 13592-13600.

15. Fukushima, J., Ishiwata, T., Kurata, M., You, Z., and Okuda, K. (1994) Intracellular receptor-type transcription factor, LasR, contains a highly conserved amphipathic region which precedes the putative helix—turn—helix DNA binding motif. Nucleic acids research, 22(18), 3706-3707.

16. Lintz, M. J., Oinuma, K. I., Wysoczynski, C. L., Greenberg, E. P., and Churchill, M. E. (2011) Crystal structure of QscR, a Pseudomonas aeruginosa quorum sensing signal receptor. Proceedings of the National Academy of Sciences, 108(38), 15763-15768.

17. Costas, C., López-Puente, V., Bodelón, G., González-Bello, C., Pérez-Juste, J., Pastoriza-Santos, I., and Liz-Marzán, L. M. (2015) Using surface enhanced Raman scattering to analyze the interactions of protein receptors with bacterial quorum sensing modulators. ACS nano, 9(5), 5567.

18. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., ... and Bourne, P. E. (2000) The protein data bank. Nucleic acids research, 28(1), 235-242.

19. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. Journal of molecular biology, 215(3), 403-410.

20. Söding, J., Biegert, A., and Lupas, A. N. (2005) The HHpred interactive server for protein homology detection and structure prediction. Nucleic acids research, 33(suppl 2), W244-W248.

21. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. Journal of applied crystallography, 26(2), 283-291.

22. Eisenberg, D., Lüthy, R., and Bowie, J. U. (1997) VERIFY3D: Assessment of protein models with three-dimensional profiles. Methods in enzymology, 277, 396-404.

23. Wiederstein, M., and Sippl, M. J. (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic acids research, 35(suppl 2), W407-W410.

24. Bolton, E. E., Wang, Y., Thiessen, P. A., and Bryant, S. H. (2008) PubChem: integrated platform of small molecules and biological activities. Annual reports in computational chemistry, 4, 217-241.

25. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. Journal of computational chemistry, 25(9), 1157-1174.

26. da Silva, A. W. S., and Vranken, W. F. (2012) ACPYPE-Antechamber python parser interface. BMC research notes, 5(1), 367.

27. Jakalian, A., Jack, D. B., and Bayly, C. I. (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. Journal of computational chemistry, 23(16), 1623-1641.

28. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX, 1, 19-25.

29. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins: Structure, Function, and Bioinformatics, 78(8), 1950-1958.

30. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. The Journal of chemical physics, 98(12), 10089-10092.

31. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. The Journal of chemical physics, 103(19), 8577-8593.

32. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. The Journal of chemical physics, 79(2), 926-935.

33. Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. The Journal of chemical physics, 126(1), 014101.

34. Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. Journal of Applied physics, 52(12), 7182-7190.

35. Chen, Y. C. (2015) Beware of docking! Trends in pharmacological sciences, 36(2), 78-95.

36. Trott, O., and Olson, A. J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry, 31(2), 455-461.

37. Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., & Olson, A. J. (2016). Computational protein-ligand docking and virtual drug screening with the AutoDock suite. Nature protocols, 11(5), 905-919.

38. Hotelling, H. (1933) Analysis of a complex of statistical variables into principal components. Journal of educational psychology, 24(6), 417.

39. Hartigan, J. A., and Wong, M. A. (1979) Algorithm AS 136: A k-means clustering algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(1), 100-108.

40. McKinney, W. (2011) pandas: a foundational Python library for data analysis and statistics. Python for High Performance and Scientific Computing, 1-9.

41. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., and Vanderplas, J. (2011) Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(Oct), 2825-2830.

42. McGibbon, R. T., Beauchamp, K. A., Harrigan, M. P., Klein, C., Swails, J. M., Hernández, C. X., and Pande, V. S. (2015) MDTraj: A modern open library for the analysis of molecular dynamics trajectories. Biophysical journal, 109(8), 1528-1532.

43. Hunter, J. D. (2007) Matplotlib: A 2D graphics environment. Computing In Science and Engineering, 9(3), 90-95.

44. Waskom, M., Botvinnik, O., Hobson, P., Cole, J. B., Halchenko, Y., Hoyer, S., and Miles, A. (2014) seaborn: v0. 5.0 (November 2014).

45. DeLano, W. L. (2002) The PyMOL molecular graphics system.

46. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. Journal of molecular graphics, 14(1), 33-38.

47. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. Journal of computational chemistry, 25(13), 1605-1612.

48. Wolf, A., & Kirschner, K. N. (2013). Principal component and clustering analysis on molecular dynamics data of the ribosomal L11· 23S subdomain. Journal of molecular modeling, 19(2), 539-549.

49. Zamora, J. E., Papadaki, M., Messer, A. E., Marston, S. B., & Gould, I. R. (2016). Troponin structure: its modulation by Ca 2+ and phosphorylation studied by molecular dynamics simulations. Physical Chemistry Chemical Physics, 18(30), 20691-20707.

50. Hayward, S., and De Groot, B. L. (2008) Normal modes and essential dynamics. Molecular Modeling of Proteins, 89-106.

51. Amadei, A., Linssen, A., and Berendsen, H. J. (1993) Essential dynamics of proteins. Proteins: Structure, Function, and Bioinformatics, 17(4), 412-425.

52. Tan, P. N., Steinbach, M., and Kumar, V. (2006) Introduction to data mining, Addison Wesley Publishers.

53. Shen, Y., and Bax, A. (2010) SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. Journal of biomolecular NMR, 48(1), 13-22.

54. Han, B., Liu, Y., Ginzinger, S. W., and Wishart, D. S. (2011) SHIFTX2: significantly improved protein chemical shift prediction. Journal of biomolecular NMR, 50(1), 43.

55. Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., and Vajda, S. (2017) The ClusPro web server for protein-protein docking. Nature Protocols, 12(2), 255-278.

56. Johnson, S. C. (1967) Hierarchical clustering schemes. Psychometrika, 32(3), 241-254.

57. Davies, D. L., and Bouldin, D. W. (1979) A cluster separation measure. IEEE transactions on pattern analysis and machine intelligence, (2), 224-227.

58. Rousseeuw, P. J. (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 20, 53-65.

59. Dunn, J. C. (1973) A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters.

60. Caliński, T., and Harabasz, J. (1974) A dendrite method for cluster analysis. Communications in Statistics-theory and Methods, 3(1), 1-27.

61. Laskowski, R. A., Hutchinson, E. G., Michie, A. D., Wallace, A. C., Jones, M. L., and Thornton, J. M. (1997) PDBsum: a Web-based database of summaries and analyses of all PDB structures. Trends in biochemical sciences, 22(12), 488-490.

62. Warren, G. L., Andrews, C. W., Capelli, A. M., Clarke, B., LaLonde, J., Lambert, M. H., and Tedesco, G. (2006) A critical assessment of docking programs and scoring functions. Journal of medicinal chemistry, 49(20), 5912-5931.

63. Wernet, P., Nordlund, D., Bergmann, U., Cavalleri, M., Odelius, M., Ogasawara, H., and Pettersson, L. G. M. (2004) The structure of the first coordination shell in liquid water. Science, 304(5673), 995-999.

64. Comeau, S. R., Gatchell, D. W., Vajda, S., and Camacho, C. J. (2004) ClusPro: a fully automated algorithm for protein–protein docking. Nucleic acids research, 32(suppl 2), W96-W99.

65. Comeau, S. R., Gatchell, D. W., Vajda, S., and Camacho, C. J. (2004) ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics, 20(1), 45-50.

66. Kozakov, D., Brenke, R., Comeau, S. R., & Vajda, S. (2006) PIPER: an FFT-based protein docking program with pairwise potentials. Proteins: Structure, Function, and Bioinformatics, 65(2), 392-406.