Interaction of N-3-Oxododecanoyl Homoserine Lactone with LasR protein 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

∗ Corresponding author.

E-mail addresses: hovakim.grabski@rau.am, lernik.hunanyn@rau.am, hrach.vardapetyan@rau.am

Abstract:

In 2017 World Health Organization announced the list of the most dangerous superbugs and among them is *Pseudomonas aeroginosa*, which is an antibiotic resistant and opportunistic human pathogen. This organism attacks human patients suffering from such diseases as AIDS, cancer, cystic fibrosis, etc. Current therapies lack efficacy, partly because *P.aeruginosa* creates and inhabits surface-associated biofilms conferring increased resistance to antibiotics and host immune responses. There is an urgent requirement for novel antibacterial treatments to counter *P. aeruginosa*. However, till date the detailed molecular properties of the LasR protein are not clearly known. Therefore, in this work, an attempt has been made to analyze molecular properties of the LasR protein as well as the mode of its interactions with the N-3-oxododecanoyl homoserine lactone (3-O-C12-HSL) autoinducer(AI). We analyzed the three-dimensional structure of the LasR protein of *P. aeruginosa* as well as performed docking and molecular dynamics (MD) simulations of the protein with the AI 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 turn in the vicinity of DNA binding domain of the protein was found. We have also performed LasR monomer protein docking and found a new form of dimerization. This study may, therefore, offer new insights for future experiments to determine the interaction of the autoinducer with the beta turn in the vicinity of DNA binding domain of LasR protein and a new interaction site for drug design.

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; BLAST, Basic local alignment search tool; DBI, David-Bouldin Index; psF, pseudo-F statistic.

1. Introduction

Pseudomonas aeruginosa (*P. aeruginosa*) is a Gram-negative, monoflagellated, obligatory aerobic, Bacillus of the family Pseudomonadaceae and was first introduced at the end of the nineteenth century (1, 2). It has been found in environments such as soil, water, in human bodies, in animals, sewage, plants, and hospitals (3). *P. aeruginosa* is an opportunistic human pathogen, because it rarely infects healthy persons. The central problem is that it infects patients who are immunologically compromised, like those suffering from AIDS, cystic fibrosis, cancer or severe wound from burning (2, 4). This organism is a major issue and it is known to cause about 40–60% deaths in hospitals. About 90% of cystic fibrosis deaths are considered to be due to the secondary infection by this pathogen (5). The pathogenicity of *P. aeruginosa* occurs due to the synthesis of virulence factors like proteases, hemolysins, exotoxin A, production of antibiotic pyocyanin, Hydrogen Cyanide (HCN), Type II secretion system, ramnolipids and biofilm formation by this organism (6). 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 HCN biosynthesis is catalyzed by the membrane-bound enzyme HCN synthase encoded by the operon called hcnABC (7). HCN can potentially inhibit the cytochrome c oxidase, the terminal component of the aerobic respiratory chain, and HCN exposure can lead to neuronal necrosis (8, 9). The transcription of hcnABC genes is regulated by three transcriptional regulators, ANR, LasR and RhlR (10). These transcriptional regulators perform in a cluster (7). However, by several mutagenesis experiments, it was proposed that LasR was the most important activator of hcnABC though the entire aforementioned transcriptional regulators (ANR, LasR and RhlR) were essential for the optimal expression of the hcnABC genes (10). The transcriptional activator protein LasR belongs to the family of LuxR type transcriptional regulators. It regulates the target gene expression by recognizing a conserved DNA sequence termed as lux box (11, 12). LasR has two domains, 1) the N-terminal autoinducer binding domain(LBD) and 2) the C-terminal DNA binding domain(DBD) (13). In the C-terminal domain, LasR has a DNA binding Helix-Turn-Helix (HTH) motif. (14). The activity of the LasR protein is regulated by the autoinducer ligand 3-O-C12-HSL which when binds to the LasR protein helps the protein to dimerize.

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. The activity of the LasR protein is regulated by the AI ligand N-3-Oxododecanoyl Homoserine Lactone (3-O-C12-HSL) which when binds to the LasR protein helps the protein to dimerize. There remains a need for understanding of the LasR monomer, because till date there is no molecular detail information about LasR monomer behaviour. Therefore, in this work we tried to analyze the mechanism of 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 viccinity of DNA binding domain. This study may therefore be used to develop new therapeutics to prevent the infections caused by *P. aeruginosa* by targeting both the LBD as well as the beta turns in the viccinity of DBD and inhibit activation of genes.

2. Material and methods

2.1 Sequence analysis of LasR

The methodology was based on a previous study(1). The Amino acid sequence of LasR protein of *P. aeruginosa* was obtained from UniprotKB (Sequence ID P25084). The

crystal structure of the N-terminal LBD of LasR protein (amino acid residues 7 to 169) from *P. aeruginosa* was available in the Brookhaven Protein Data Bank (PDB) (PDB ID: 3IX3, chain — A). However, in order to have a better understanding of LasR protein properties, the structure of the entire LasR protein was needed.

Therefore, the amino acid sequence was used to search in the PDB databank to find suitable template(s) for homology modeling using BLAST (15). From this BLAST search, the best match was found to be the crystal structure of QscR, bound to 3-O-C12-HSL (PDB code 3SZT, chain — A) (16) with 30% sequence identity, 97% query coverage and E-value of $1 \times 10{-25}$. However, the other similar structures which were obtained from BLAST search results were as follows:

1. 1FSE, chain—A with 37% sequence identity, 28% query coverage and E-value of $2 \times 10{-4}$. It is the crystal structure of Bascillus subtilis regulatory protein gene.

2. 4Y15, chain—A with 30% sequence identity, 75% query coverage and E-value of $3 \times 10{-16}$. It is the crystal structure of Sdia,an Escherichia coli homologue of the quorum-sensing regulator, in complex with 3-oxo-c6-homoserine Lactone.

3. 3ULQ, chain—B with 32% sequence identity, 25% query coverage and E-value of 0.09. It is the crystal structure of the Anti Activator Rapf complexed with the response regulator coma DNA binding domain.

2.2 Modeling of LasR monomer

The protein LasR is a 239 residue long protein. The homology modeling of LasR protein was done using the HHPred web server (17). The HHPred web server uses a multi-template modeling approach and the templates, used for the homology modeling of LasR, are 3SZT_A, 3IX3_A, 1FSE_A and 3ULQ_B. These structures were also identified by BLAST(16) to be the suitable templates for homology modeling of LasR protein. The final model of LasR protein spanning the amino acid residues from 1 to 6 and 170 to 239 was built using the aforementioned templates. The root mean squared deviation (RMSD) of the built model was checked by superposing the coordinates of the modeled protein on to the template 3SZT, chain — A and it was found to be 1.512 Å. The stereo-chemical fitness of the modeled LasR protein was checked by PROCHECK (18) (Figure S1,S2 of the Supporting Information), Verify3D (19) and Ramachandran Plots. The structure was found to be stereo-chemically fit.More than 96.65% of the amino acid residues in the final modeled structure had a 3D–2D score >0.2 as specified by Verify3D(Figure S3 of the Supporting Information) for a good computational model (Eisenberg et al., 1997). No amino acid residues were found to be in the disallowed regions of the Ramachandran Plot (Figure S1 of the Supporting Information). The PROSA value of this model was -6.69 which represents a good model quality for homology models (20), although the obtained model is slightly different (1), it is still viable.

2.3. Building the model of the ligand

The LasR protein is known to exert its function when it gets attached to the ligand, 3-O-C12-HSL which is the AI (Figure 1). Thus to study the mechanism of the ligand

binding interactions with LasR, the three dimensional model of the ligand was generated. The model was obtained from PubChem web-server (21).

The ligand parameters were calculated for the General Amber Force Field (22) by using the acpype tool (23) with AM1-BCC partial charges (24).
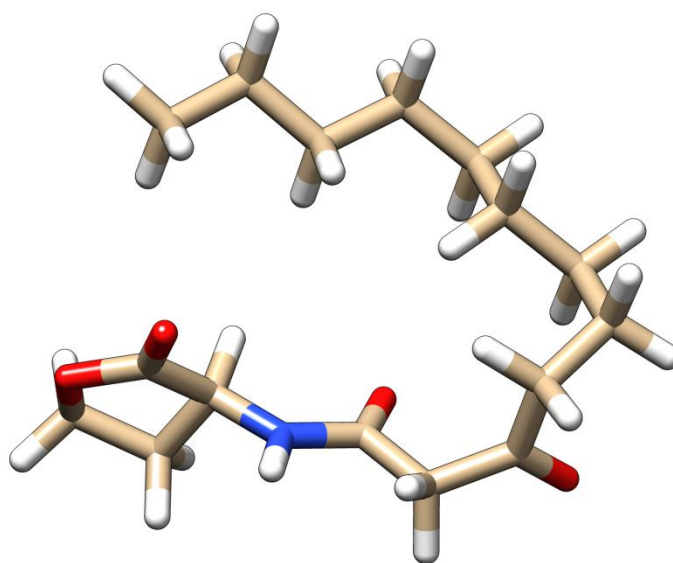

Figure 1: The stick representation of the 3O-C12-HSL AI ligand.

2.4. Molecular Dynamics Simulations of LasR protein systems.

MD simulations of all systems were conducted with the GROMACS suite, version 5.1.2 (25), utilizing the Amber ff99SB-ILDN force field (26).

In all cases, Short-range nonbonded interactions were cut off at 1.4 nm, with long-range electrostatics calculated using the particle mesh Ewald (PME) algorithm.(27, 28) Dispersion correction was applied to energy and pressure terms to account for truncation of van der Waals terms. Periodic boundary conditions were applied in all directions.

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 (29),to which 100 mM NaCl was added, including neutralizing counterions. Following two steepest descents minimization, LasR monomer was equilibrated in two steps, with position restraints applied to peptide heavy atoms throughout. The first phase involved simulating for 200 ps under a constant volume (NVT) ensemble. Protein and nonprotein atoms were coupled to separate temperature coupling baths, and temperature was maintained at 300 K using the Berendsen weak coupling method(31.).Following NVT equilibration, 200 ps of constant- pressure (NPT) equilibration were performed, also using weak coupling(31) to maintain pressure isotropically at 1.0 bar. Production MD simulations were conducted for 100 ns in the absence of any restraints. For this data

collection period, the V-rescale(32) was used to maintain temperature, and the Parrinello-Rahman barostat(33,34) was used to isotropically regulate pressure. This combination of thermostat and barostat ensures that a true NPT ensemble is sampled.(32)

Set 2. Molecular dynamics simulations using docked structures. It was 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; coordinates were recorded every 1 ps. A 300 ns MD trajectory for each system was obtained under constant pressure at 300 K. Further details concerning the model building and simulation protocol can be found above.

2.5 LasR–3-O-C12-HSL ligand blind docking simulations

To build the LasR–ligand complex, the autoinducer ligand 3-O-C12-HSL was docked with LasR monomer using Autodock Vina(36), based on a simple scoring function and rapid gradient-optimization conformational search. However, it should be noted that AutoDock Vina currently uses several simplifications that affect the obtained results. The most significant simplification 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. Note that the default exhaustiveness value of the original Vina is 8 and the authors claim that it "should increase the time linearly and decrease the probability of not finding the minimum exponentially" (36), 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(PCA) (37) and cluster analysis using K-means algorithm (38) 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 and good speed.

After the exhaustiveness was chosen, 100 independent docking calculations were carried out with different initial seeds within a box of dimension that was used before. All parameters were set to default, except the exhaustiveness value, which we increased to a value of 1024 to allow thorough sampling of the docked configurations and maximum number of binding modes to generate set to 20.
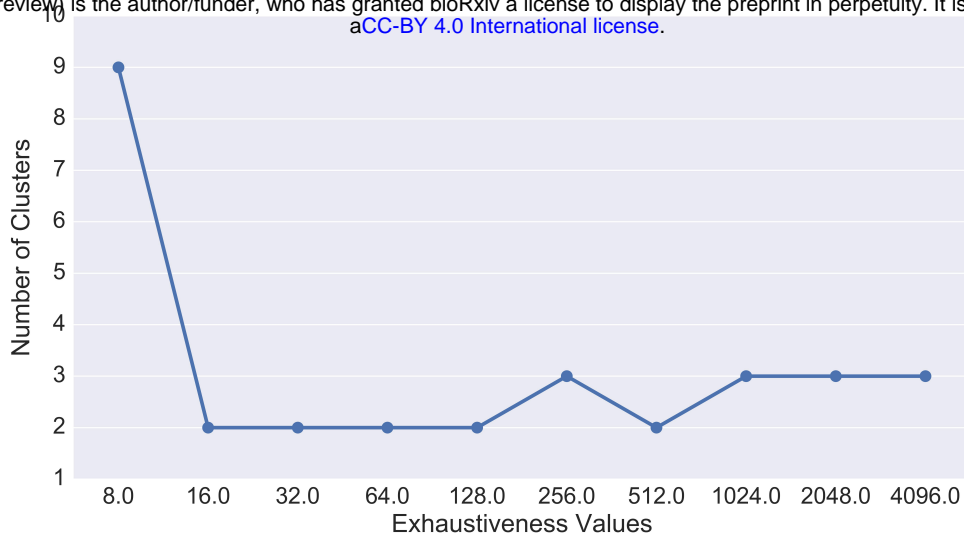
Figure 2: Exhaustiveness determination by number of clusters.

2.6 Analysis of docking conformations and trajectories

We performed the analysis of docking conformations and trajectories by a combination of Autodock Vina(36), Gromacs(25) in addition to custom python scripts, which uses pandas(39), scikit-learn(40) and MDTraj(41).

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

- The MDTraj python library(41) was used for the trajectory analysis.

- Plot visualization was done using matplotlib(42) and Seaborn package(43).

- Figures and videos were prepared with PyMOL(44), VMD(45) and UCSF Chimera(46).

We used the following techniques for the analysis:

- Principal component analysis is an unsupervised statistical technique used to emphasize variation and bring out strong patterns in a dataset. It's often used to make data easy to explore and visualize.. Developed about 100 years ago by Pearson and Hotelling (47), it remains a popular dimension reduction technique (47). The PCs describe concerted atomic displacements and can highlight major conformational changes between the structures. These motions are essential for protein function (48). The dynamics in this low-dimensional subspace – spanned by the first few PCs – was termed "essential dynamics" (49). PC analysis, performed on Cartesian coordinates has proven to be a valuable tool for studying conformational changes.

- Cluster analysis is another unsupervised technique for grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar to each other than to those in other groups (clusters). by minimizing intra-cluster and maximizing inter-cluster differences. Therefore, most clustering algorithms require a measure of similarity, or "distance", of objects. Clustering algorithms can be divided into partitional and hierarchical clustering approaches (50).

2.6. NMR calculations

Finally, the trajectory of LasR monomer simulation was used for the calculation of chemical shifts. Sparta+ (51) and ShiftX2 (52) were used to predict the chemical shifts of backbone atoms of the protein. In the prediction process, waters were removed from each trajectory using mdtraj's wrapper functions (41).

2.8 Protein-Protein Docking

In general, computational docking approaches strive to sample all possible interactions between two proteins to discern the biologically-relevant interaction. 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 (53). 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. In its basic operation, ClusPro outputs the structures at the centers of the 10 most populated clusters

2.7 Entire flowchart

The entire methodology is presented as a flowchart for the better understanding of the reader:

- Sequence of LasR taken from UniprotKB(sequence ID P25084). Protein BLAST of LasR was performed to find out suitable templates for homology modeling.

- Homology modeling done using HH-pred web server. The modeled structure was validated using Verify 3D, Procheck, Prosa.

- The structure of 3-O-C12-HSL obtained from pubchem web server.

- The AI 3-O-C12-HSL ligand docked with the LasR monomer 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.

3. Results

3.1 Compactization of LasR without 3-O-C12-HSL

We performed a simulation run of 100 ns using a standard MD protocol for the assesemnt of conformational changes of LasR monomer. To assess the overall stability of the molecule, the root-mean-square-deviation (RMSD) of the protein atoms was calculated with reference to the initial frame through time MD runs (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).

When analyzing MD simulations, Principal Component Analysis (PCA) (47) is commonly applied to visualize the motions that the system is exploring through the simulation. Each of these motions, or Principal Components (PCs), is responsible for an amount of variance in the original data that is quantified by its associated eigenvalue.

Thus, usually the first handful of PCs are enough to capture more than 70% of the variance in the original trajectory (the exact amount depends on the original trajectory data set). Though PC analysis reveals the main motions contained in an MD trajectory, it does not partition the snapshots into distinct conformational classes (50). This can be achieved by clustering the PC data.
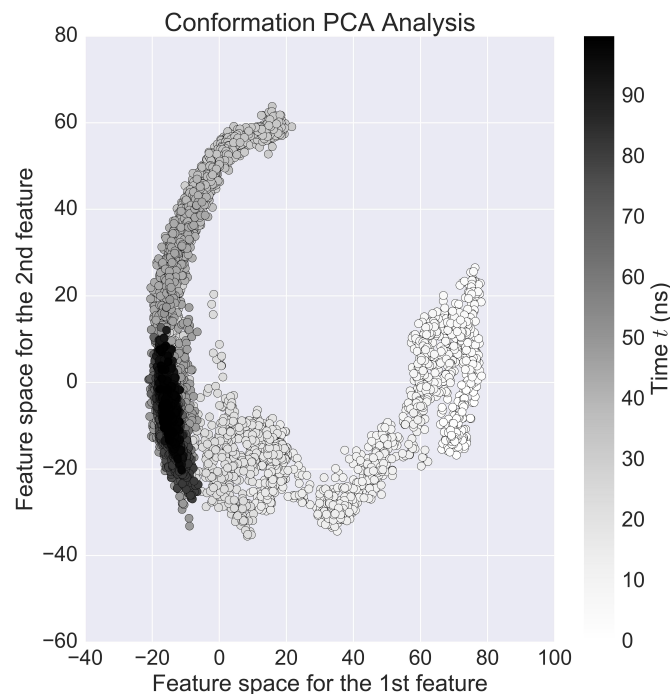

Figure 3: LasR monomer simulation trajectory projected onto its first two principal components. The black scale indicates the time progression from 0 ns (white to 100 ns (black).

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. Additionally, this cluster analysis permits assess what are the recurrent conformations that LasR adopts throughout the MD runs. For the clustering analysis, we chose the Agglomerative Clustering algorithm with Ward linkage (54) 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 using different parameters (details can be found in the Computational methods section). The accuracy of the different clustering rounds was assessed by the help of the Davies–Bouldin Index(DBI) (55), Silhouette Score (56), Dunn Index (57) and the pseudo-F statistic(pSF or Calinski Harabasz) (58) metrics. These four metrics were plotted vs. the cluster count that was achieved using different number of cluster quantity parameters(Figure S8 of the Supporting Information). An optimal number of clusters were chosen for each region, simultaneously accounting for a low DBI, High Silhouette, High Dunn Index and high pSF values.

The distribution of clusters over simulation is divisive (Figure 4). Clearly separated in all four cases are: cluster 4 (black) at the beginning of the simulation (after equilibration),cluster 2 (green) in the middle, cluster 1 (lightgreen) and cluster 3 (darkblue) at its end.
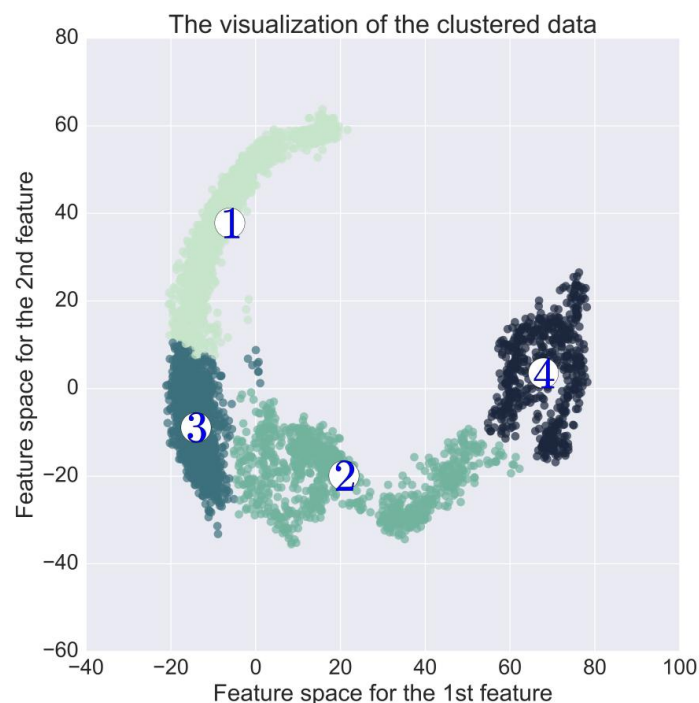

Figure 4: Clustering results of ward-linkage algorithm on 2d plane formed by first two PCs. Clustering was performed on the entire MD trajectory data.

In this study, the clustering in the subspace defined by the first two PCs (Figure 4) provides the most coherent picture. Moreover, this clustering is supported by a good DBI, Dunn,Silhoutte Score and psF value (Figure S8 of the Supporting Information). It also shows clearly that the simulation has converged (Figure 5).
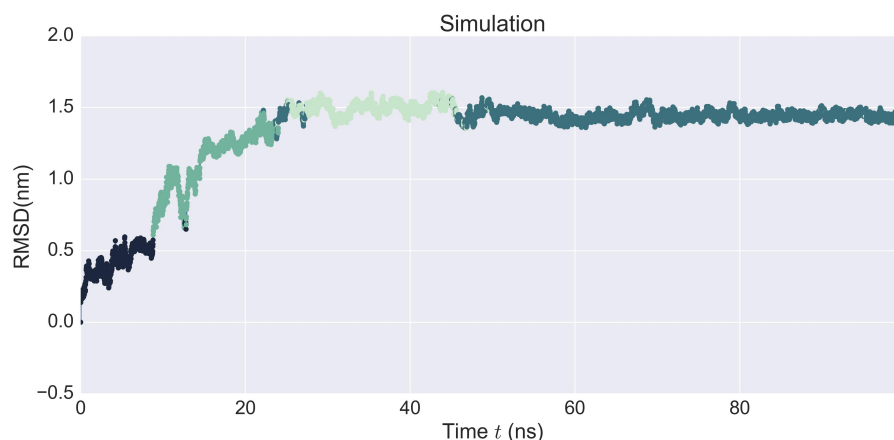

Figure 5: RMSD over simulation time and color-coded by clusters obtained from Ward-linkage algorithm for the the complete data .The colors correspond to the clusters as defined in Figure 4.

For the validation of the quality of molecular dynamic simulation, theoretical NMR shifts were calculated using Sparta+ (51) and ShiftX2 (52). For the NMR shifts calculation, snapshots from cluster 3(Figure 4 and Figure 5) were used.

In Figure 6 and Figure 7 you can see that there is a strong linear relationship between experimental and simulated NMR shift values (Figure S9 of the Supporting Information), thus this demonstrates the quality of MD simulation.
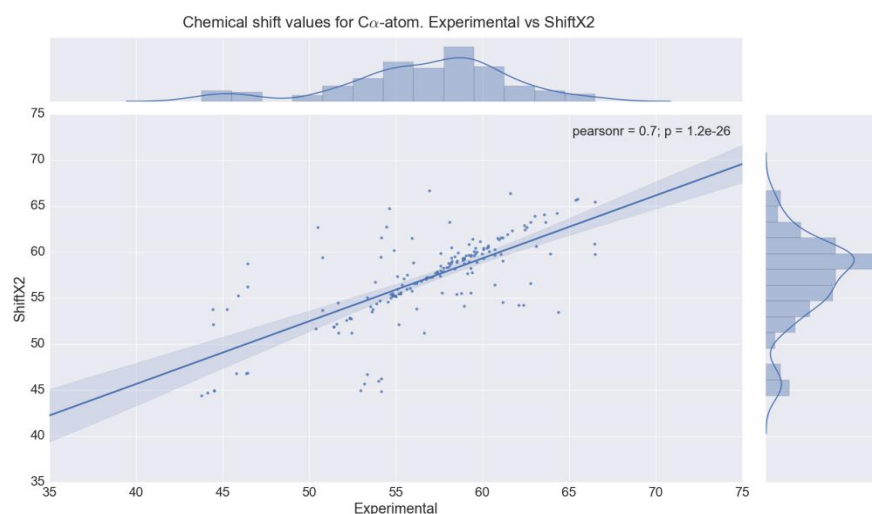

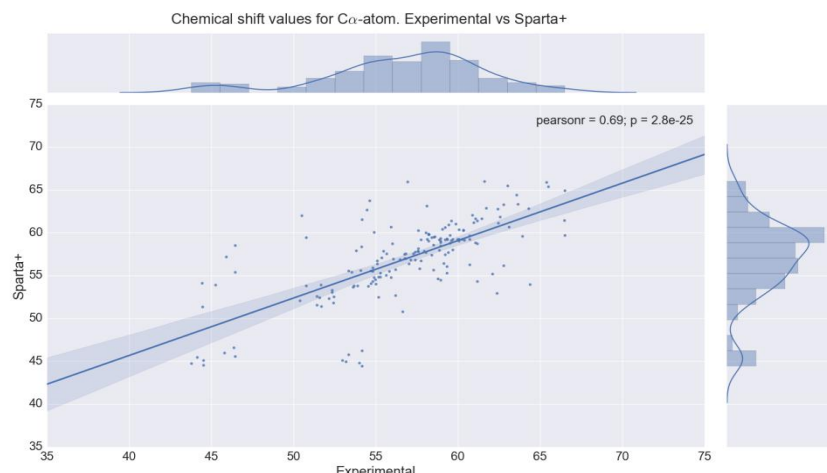*Figure 6: Chemical shifts of Cα- atoms. Experimental vs ShiftX2 predictions.*

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

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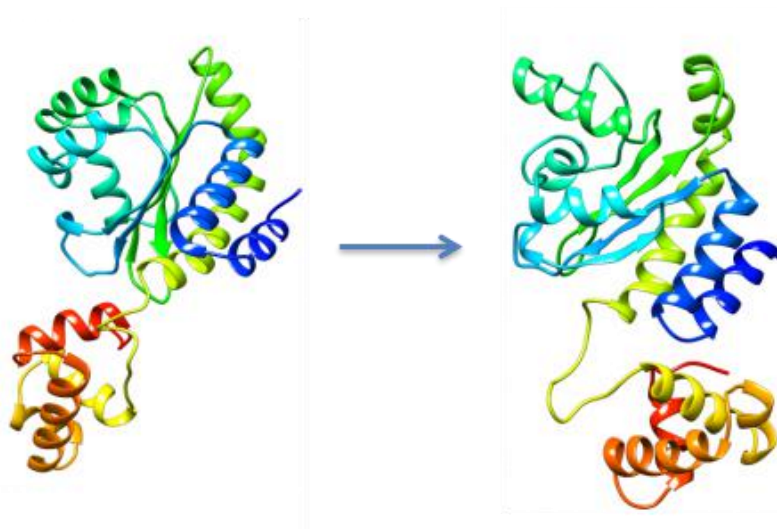

Figure 8: Left Homology Modeling, Right Representative Structure after 100ns MD run. Images generated with UCSF Chimera(46)

The representative structure of LasR system and its differences from the homology modeled structure are highlighted in Figure 8. Secondary structure analysis of the representative structure was performed as well using PDBSum (59) (Figure S10 of the Supporting Information).

To our knowledge, this is the most complete model used for long-time scale MD simulations to date, containing the full-length LasR molecule (residues 1 to 239). This is also the first study to include the dynamics of the complete C-terminal region of LasR,

which has proven to have interesting modulations of the N-terminal region of LasR dynamics, as will be discussed later.

3.2. 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 protein. PCA and cluster analysis were performed on docking data(Fig. 8). 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. (13), where it was also demonstrated that 3-O-C12-HSL binds to LBD.

We performed several rounds of K-Means clustering with different parameters (details can be found in the Computational methods section). The accuracy of the different clustering rounds was assessed using the DBI (55), Dunn Index (56), Silhouette score (57) and the pSF (58) metrics. These four metrics were plotted vs. the cluster count (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 representatives from each cluster are shown in Figure 10. These cluster representative structures were produced by finding the centroid conformations.
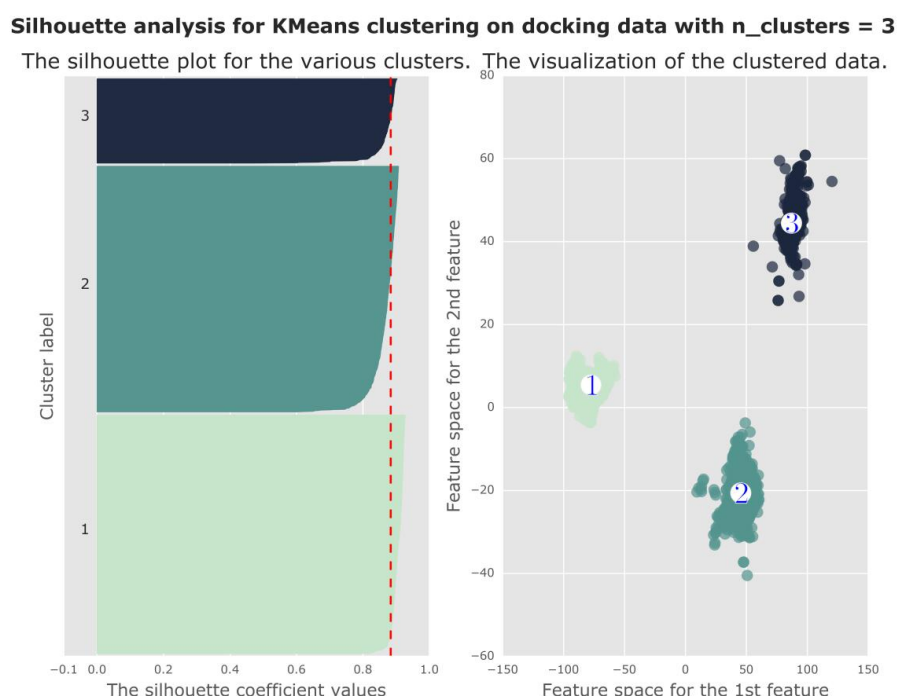

Figure 9: Clustering results of K-Means algorithm on 2d plane formed by first two PCs. Clustering was performed on the docking data.

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 (13). 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%, an interesting region that to our knowledge has never been previously studied.

There is significant literature suggesting that docking experiments are not suitable to predict the binding affinity of protein-ligand complexes, but they can still provide important information (60). For example, Warren et al. found that docking programs often identified the structure of the crystallographic ligand. However, the scoring functions were often not able to identify the structure of the crystallographic ligand as the lowest energy pose (60), as well as the work of Chen (35).
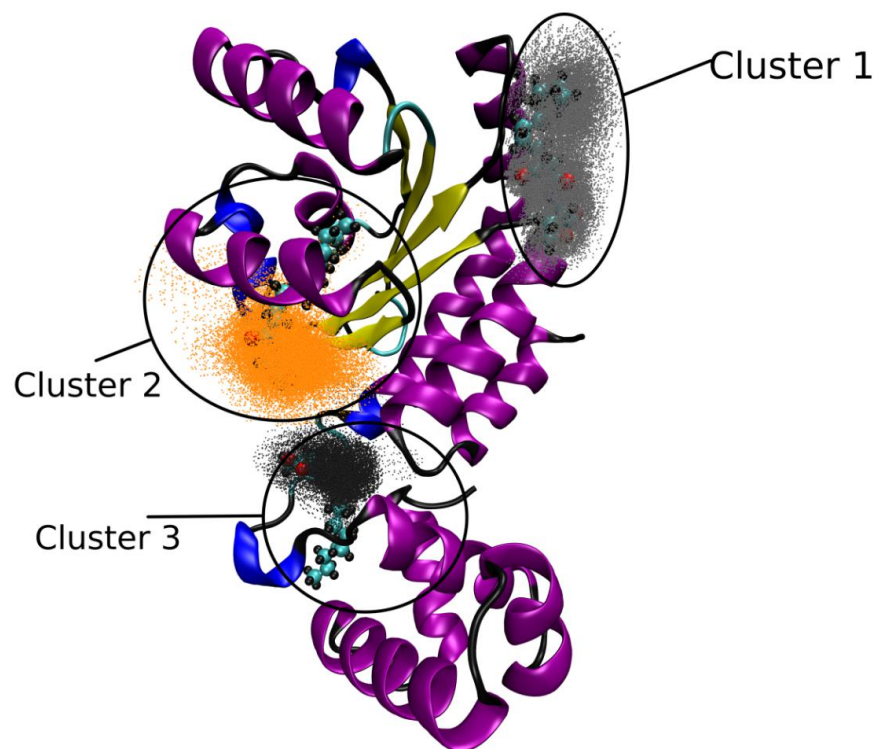

Figure 10: 3D visualization of the analyzed docking data with their representative structures and clusters.

3.3. Binding Modes of 3-O-C12-HSL.

We performed three 300 ns simulations using standard MD protocol. Overall, 900ns of aggregate simulation time was used for the data 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.

To assess the overall stability of the molecule, the mass-weighted root-mean-square-deviation (RMSD) of the backbone atoms was calculated with reference to the initial frame through time 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).

Simulation 2 shows that after 230ns, the structure becomes stable.While Simulation 3 changes its conformation in 100ns and becomes stable till 300ns.
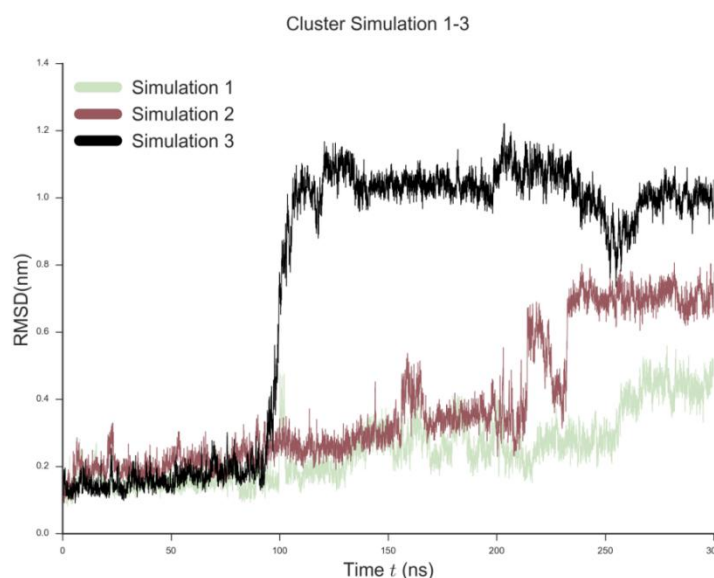

Figure 11: RMSD evolution during 300 ns of simulation for 3 independent runs using representative structures of 3-O-C12-HSL docking poses as starting points. The RMSD is calculated for the backbone atoms to the initial starting frame of each run.

Another more suitable measurement of the flexibility of the LasR monomer is the per-residue root-mean-square fluctuation (RMSF). In Figure 12 the average per-residue RMSF for each cluster simulation runs is plotted.
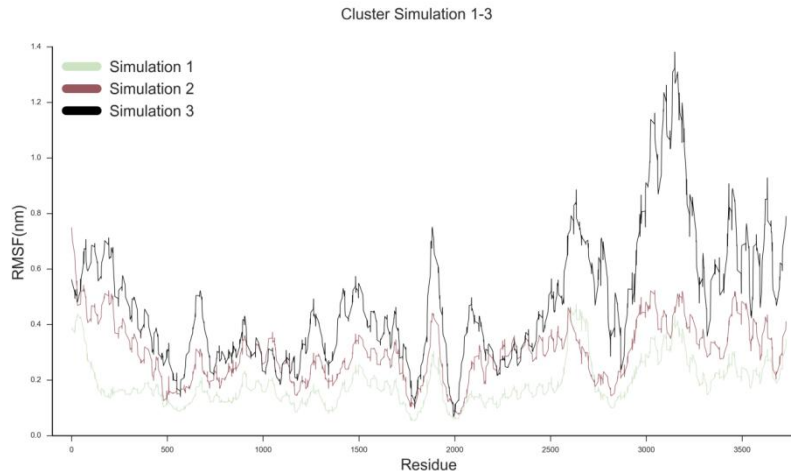
Figure 12: RMSF comparison for each individual run. The lines represent the mean RMSF value for each residue.

From Figure 12 it visible that the residues from 165 to 176, which correspond to beta turns in the viccinity of DBD in 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 viccinity of DNA binding domain.

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(61) using MDTraj (41). Over the course of cluster 1, 3-O-C12-HSL with LasR established an average of 0.655 ± 0.651 of hydogen 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 in 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 aminoacids from H8,H10 and A strand(Fig. S13 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(Fig. 13).
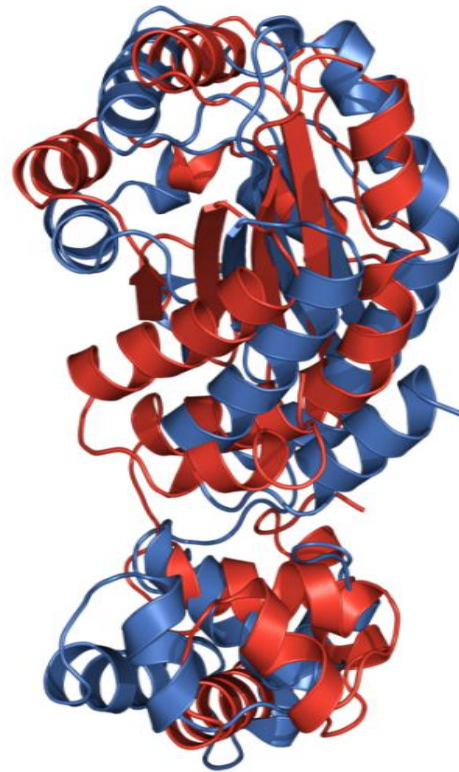
Figure 13: LasR monomer structures blue color) LasR monomer without 3-O-C12-HSL, red color) LasR monomer with 3-O-C12-HSL bound to LBD.

Residues that participate in hydrophobic interactions are shown on Figure 14 and Figure S13.
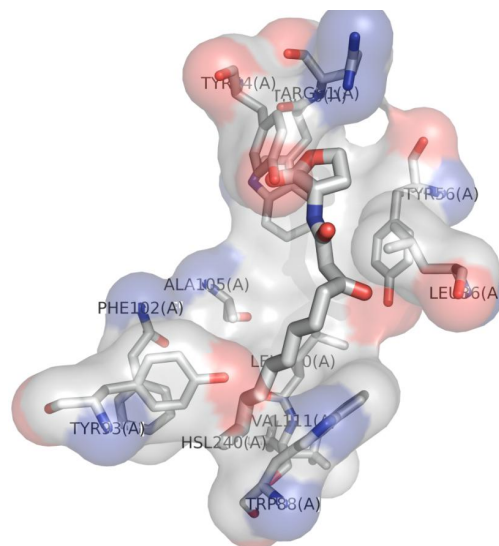

Figure 14: Insertion of 3-O-C12-HSL in LBD of LasR, from centroid snapshot of cluster 2 in simulation II(Fig. S9). Images were generated with PyMOL (44).

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

PCA, cluster and hydrogen bond analysis were also performed on simulation 3(Fig. S14). 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(Fig. S13),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 viccinity of DNA binding domain 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 viccinity of DNA binding domain(Fig. S15 from supplementary material). 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).
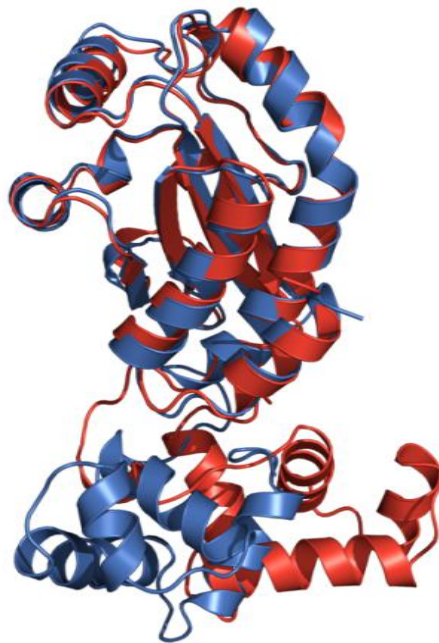

Figure 15: LasR monomer structures blue color) LasR monomer without 3-O-C12-HSL, red color) LasR monomer with 3-O-C12-HSL bound to beta turns in the viccinity of DBD.

Residues that participate in hydrogen and hydrophobic interactions are shown on Fig. 16.
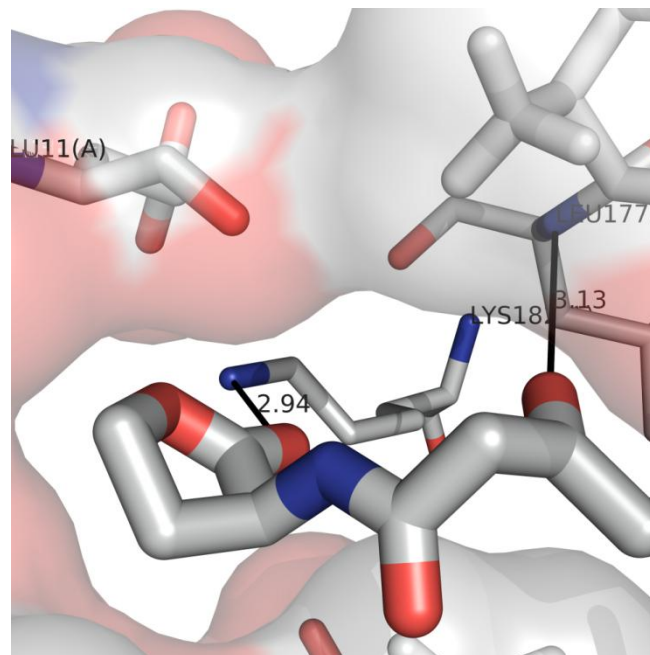
Figure 16: Formation of hydrogen bonds of OdDHL with Lys182 and Leu177, from a centroid snapshot of cluster 2 in simulation II(Figure S9). Putative hydrogen bonds are shown as black lines. Image was generated with PyMOL (44).

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

Validation of MD models by protein docking

Given that our findings are based on computational simulations, it is therefore necessary to apply the complementary method to validate these results. It is known that 3-O-C12-HSL to LasR helps in the process of dimerization and binds to the corresponding promoter DNA region as a dimer. Thus, the monomeric LasR–3-O-C12-HSL complexes were subjected to dimerization. We performed the protein docking experiment using ClusPro 2.0 (53, 62–64) due to 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 (53), suggesting to find the most populated clusters rather than the lowest-energy structures and use the centers of the clusters as putative models.

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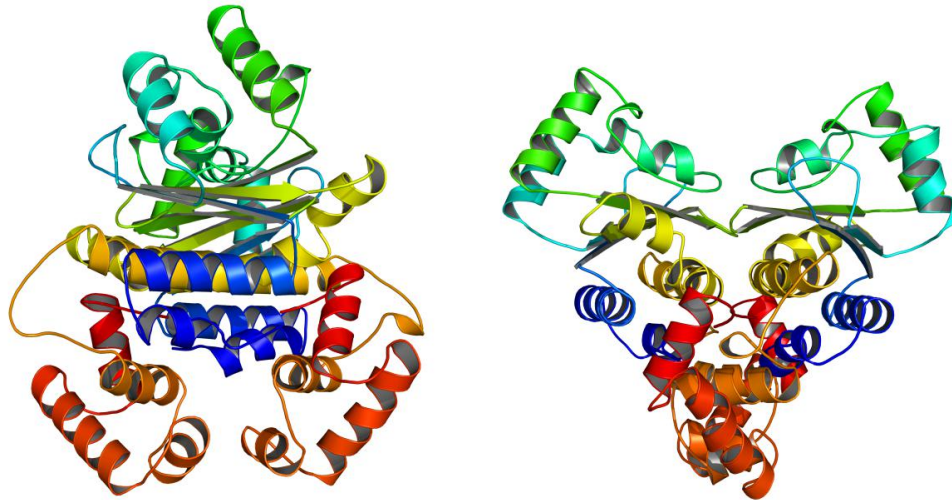
Figure 17: LasR Protein-Protein docking result based on centroid conformation from cluster 2 of simulation 2(3-O-C12-HSL with LBD). Left and front views.

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.(Fig. 18)
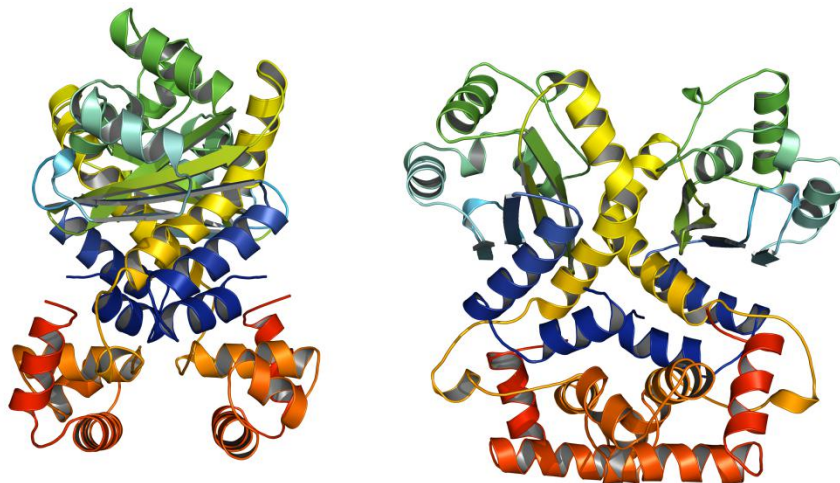

Figure 18: LasR Protein-Protein docking result based on centroid conformation from cluster 3 of simulation 3(3-O-C12-HSL with beta turns in the viccinity of DBD). Left and front views.

The interaction of 3-O-C12-HSL with the beta turns in the region of 170-175 creates a favourable conformation for dimerization. From our analysis, it is observed that OdDHL 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 viccinity of DBD of LasR protein. The binding of 3-O-C12-HSL provokes conformational transitions. The two binding conformations are capable of dimerization.

5. 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 viccinity of DBD of LasR protein. The interaction with the beta turn is a novel site. Both conformational transitions favour 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 also sheds light on the patterns of interactions of the autoinducer 3-O-C12-HSL, ligand with LasR protein. Results from this study may be used as a first-hand guide for future drug development endeavors.

References

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

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

7. Laville, J., Blumer, C., Von Schroetter, C., Gaia, V., Défago, G., Keel, C., & 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.

8. Solomonson, L. P. (1981). Cyanide as a metabolic inhibitor. Cyanide in biology/edited by B. Vennesland...[et al.].

9. Niquet, J., Baldwin, R. A., Allen, S. G., Fujikawa, D. G., & 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.

10. Pessi, G., & 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.

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

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

13. Bottomley, M. J., Muraglia, E., Bazzo, R., & 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.

14. Fukushima, J., Ishiwata, T., Kurata, M., You, Z., & 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.

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

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

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

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

19. Eisenberg, D., Lüthy, R., & Bowie, J. U. (1997). [20] VERIFY3D: Assessment of protein models with three-dimensional profiles. Methods in enzymology, 277, 396-404.

20. Wiederstein, M., & 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.

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

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

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

24. Jakalian, A., Jack, D. B., & 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.

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

26. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., & 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.

27. Darden, T., York, D., & 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.

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

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

30. Berendsen, H. J., Postma, J. P., van Gunsteren, W. F., & Hermans, J. (1981). Interaction models for water in relation to protein hydration. In Intermolecular forces (pp. 331-342). Springer Netherlands.

31. Berendsen, H. J., Postma, J. V., van Gunsteren, W. F., DiNola, A. R. H. J., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. The Journal of chemical physics, 81(8), 3684-3690.

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

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

34. Nosé, S., & Klein, M. L. (1983). Constant pressure molecular dynamics for molecular systems. Molecular Physics, 50(5), 1055-1076.

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

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

38. Hartigan, J. A., & 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.

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

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

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

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

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

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

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

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

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

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

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

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

51. Shen, Y., & 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.

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

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

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

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

56. 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.

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

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

59. Laskowski, R. A., Hutchinson, E. G., Michie, A. D., Wallace, A. C., Jones, M. L., & 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.

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

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

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

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

64. Kozakov D, Brenke R, Comeau SR, Vajda S. PIPER: an FFT-based protein docking program with pairwise potentials. Proteins. 2006;65(2):392–406. 48.