MDPI

Article

Searching Hit Potential Antimicrobials in Natural Compounds Space Against Biofilm Formation

Roberto Pestana-Nobles [1‡], Jorge A. Leyva-Rojas [1,‡] and Juvenal Yosa [1,‡]*

[1] Universidad Simón Bolivar, Facultad de Ciencias Básicas y Biomédicas, Laboratorio de Simulación Molecular y Bioinformática, Barranquilla, 080002, Colombia.

* Correspondence: juvenal.yosa@unisimonbolivar.edu.co

‡ These authors contributed equally to this work.

Version September 2, 2020 submitted to Molecules

Abstract: Biofilms are communities of microorganisms that can colonize biotic and abiotic surfaces playing a significant role in the persistence of bacterial infection and antibiotic resistance. About 65% and 80% of microbial and chronic infections are produced by biofilm formation. The increase in infections by multi-resistant bacteria draws attention to the discovery of new drugs based on natural inhibitory molecules. The inhibition of diguanylate cyclases (DGCs), the enzyme implicated in the synthesis of the second messenger, cyclic diguanylate (c-di-GMP), involved the biofilm formation, represents a potential method for preventing the biofilm development. It has been extensively studied using PleD protein as a model of DGC for *in silico* studies as virtual screening and as a model for *in vitro* studies in biofilms formation. In the present study 224205 molecules from natural products database, ZINC15 has been evaluated through molecular docking and molecular dynamic simulation, our result suggests *trans*-Aconitic acid (TAA) as a possible starting point for hit-to-lead methodologies to obtain new molecules capable of inhibiting the PleD protein and hence blocking the biofilm formation.

Keywords: Biofilms; virtual screening; molecular dynamics; natural products; binding energy; *trans*-Aconitic acid; hit-to-lead

1. Introduction

Biofilms are communities of microorganisms that can colonize surfaces, they have been found on both biotic and abiotic surfaces playing a significant role in the persistence of bacterial infection. The microorganisms are embedded in a self-produced extracellular matrix of extracellular polymeric substances (EPS) [1,2]. The biofilms formation provides advantages for the microorganisms. The change from free-living bacteria to biofilm involves the production of adhesins and extracellular matrix compound interconnecting the cells in the biofilms [3]. The extracellular matrix helps to support and connect the cell-to-cell interaction in the biofilms and plays an important purpose in several processes including cell attachment, cell-to-cell connecting, structural function, and antibiotic tolerance. This matrix producing by the bacteria is mainly composed of proteins, enzymes, polysaccharides, signalling molecules and extracellular DNA [4].

This microorganism assemblage provides a protected mode of growth with some resistance to antimicrobial and to killing by the host immune system. The biofilm development makes possible the surviving of the pathogen microorganisms in hostile environments and to disperse and colonize other niches also. [5].

In the medical field, the biofilm phenotype has been recognized as an active agent in the development of many infections, being the most common, those related to the use of medical devices[6]. The National Institutes of Health have reported that, among all microbial and chronic infections, 65%,

and 80%, respectively, are associated with biofilm formation [7]. The understanding of the molecular and physiological mechanisms of biofilm formation will bring a tool to inhibit its establishment.

The cyclic diguanylate (c-di-GMP) is an important second messenger involved in the biofilm organization[8]. This signalling molecule modulates bacterial growth phenotypes but not only limited to biofilm formation, virulence factor production, and motility. [8,9]. The synthesis of c-di-GMP depends on diguanylate cyclases (DGCs) which take two molecules of guanosine-5′-triphosphate (GTP) to obtain c-di-GMP in a two-step reaction. While, phosphodiesterases (PDEs) hydrolyse c-di-GMP to linear di-GMP [9,10]. The inhibition of DGCs as a strategy for preventing the biofilm formation represents a potential method, it has been extensively studied using PleD protein (EC: 2.7.7.65) as a model of DGC for *in silico* studies, using virtual screening and as a model for *in vitro* studies in biofilms formation [8–12].

The increase in infections by multi-resistant bacteria draws attention to the discovery of new drugs. Bioactive molecules are efficient natural compounds against living organisms, they are synthesized by microorganisms, plants and animals. These substances act as a chemical defence in various competitive environments [13]. An alternative for selecting DGC natural inhibitory-molecules is the search in the database of natural products, which could supply substances, especially smalls compounds, with good binding to the target enzyme.

Since the discovery of penicillin natural products plays an important role in the discovery and development of new drugs [14,15]. Between 1981 and 2014 were approved 43.6% of anti-infective drugs and 40.7% of anticancer agents where was based on a natural product or derivatives thereof [15], many natural products have become a template for drug design because of often present ligand-protein binding motifs [16]. Some natural products have shown antibacterial activity like the described by Nofianiet al [17] or Emiru et al [18].

In the present study, an evaluation of the of diguanylate cyclase, having as a model the PleD protein from *Caulobacter crescentus* against 224205 molecules from natural resources from the ZINC15 database is carried out. Here, virtual screening VS, molecular dynamics (MD) simulations and binding free energy calculation adopting the Molecular Mechanics Poisson-Boltzmann Surface Area method (MM/PBSA) were used for the hit searching. A summary of the complete procedure is shown in Figure 1. This work aimed to propose potential hit substance from natural compounds database that can be optimized to generate a promising lead candidate compounds as a competitive ligand for the diguanylate cyclase proteins.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 September 2020 doi:10.20944/preprints202009.0046.v1

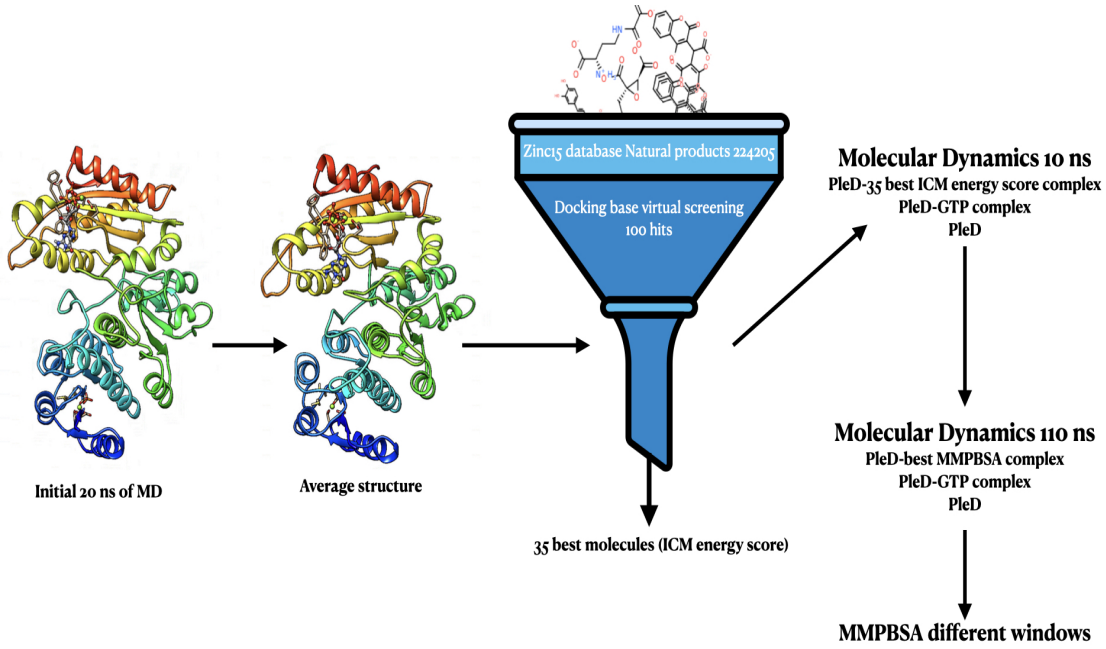

Figure 1. Summary of the procedure used for hit identification. An initial 20 ns of MD was performed with the PleD protein (inactive form) in complex with GTP molecule and Mg^{2+} ions, following the crystal structure of the active form 2V0N to coupling the GTP at the catalytic site. From equilibrated trajectory starting at 1.2 ns, an average structure was produced using chimera software [19]. A library of about 224205 natural products compounds was obtained from ZINC15 database. Bond orders, tautomeric forms, stereochemistry, hydrogen atoms and protonation states were assigned automatically by the ICM package with default parameters for each compound to prepare the ligands for docking. Ligands were docked using a VS scheme from ICM using the PleD protein average structure obtained in the above procedure. The GTP geometric space in the protein was used as a reference for the calculation of grid dimensions. 100 hits were obtained from VS procedure using the ICM energy scoring function for ranking it. From 100 hits, 35 with the best ICM energy score, were selected for 10 ns of MD simulations. For each molecule, MM/PBSA was computed using 200 frames from the last 2 ns of each MD. Finally, with the best complex (Ligand 4)-PleD, ranked from MM/PBSA binding free energy, a 110 ns of MD was done, also for GTP-PleD protein complex and PleD protein. From the above three trajectories, MM/PBSA was computed again at different windows starting from 20 ns until 60 ns, using for each window 200 frames.

2. Results and Discussion

The diguanylate cyclase PleD protein of *Caulobacter crescentus* belongs to response regulator family, its activity is controlled by phosphorylation where two cognate kinases are involved; DivJ and PleC. PleD is required for polar differentiation in the bacterial cell cycle. Bacterial cells without functional PleD are hypermobile and fail to accomplish the swarmer-to-stalked cell transition [20–23]. PleD protein contains an intrinsic nucleotide cyclase activity, which converts two molecules of GTP into 3′,5′-cyclic diguanylic acid (c-di-GMP).[21]. c-di-GMP is a molecule of great interest which regulates surface-adhesion properties and motility in bacteria [24,25]. In addition, c-di-GMP is involved in bacterial biofilm formation and persistence[26]. Due to the fact that c-di-GMP is found exclusively in bacteria, makes it a potential target for medicinal applications.

From the pseudo-active structure (PDB-ID: 2V0N) the mode of substrate binding as far as the position of the terminal phosphates close to the P-loop, together with two Mg^{2+} ions, which are coordinated by the phosphates and two invariant carboxylates (ASP327 and GLU370). Chain A from

an inactive structure (PDB-ID: 1W25), was chosen for MD and VS simulation, a GTP molecule and two Mg^{2+} ions were added to the catalytic site of the protein, keeping the position of the substrate (GTP) mention above, from the pseudo-active structure. Also, ASP52 was phosphorylated, parameters for the phosphate group were obtained following the procedure mentioned elsewhere [27].

Molecular dynamics was done for the complex PleD-GTP for 20 ns. The monomer of PleD (chain A) was chosen and GTP molecule was added to the structure in the catalytic site, together with Mg^{2+} ions and phosphorylation at ASP52. Charges and atom types, for each ligand, were assigned from the gaff force field.

The system converges at 1.2 ns of simulation(Figure S1 supplementary material). From 1.2 ns an average structure for the rest of the trajectory was computed for virtual screening (VS) simulation using Chimera package [19] Figure 2. First, molecular docking was done with the GTP molecule, obtained an ICM energy score of about -27.64 kcal/mol (Table S1 supplementary material). Virtual screening was done with natural compounds library from ZINC15, having as a reference the GTP molecule from the average structure. The coordinates of the ligands were extracted from the sdf file from ZINC15 natural compounds library. Bond orders, tautomeric forms, stereochemistry, hydrogen atoms and protonation states were assigned automatically by the ICM package with default parameters [28]. Atom types, charges and parameters were obtained from MMFF (Merck Molecular Force Field) force field for each ligand in the VS procedure. Geometry optimization was performed using the automatic converting procedure in ICM [28]). The ICM score for the best 100 hits molecules is summarized in table S1 supplementary material. These have ICM score energies ranging from -47.13 kcal/mol to -31.62 kcal/mol. As results from the VS procedure, the first 100 natural compounds have better ICM score energy than the GTP, Figure 3. The main idea of this work is to find a compound from natural product source, which can block the active site of the PleD protein, in that way, blocking the formation of the c-di-GMP. To do that, a more accurate calculation of the binding free energy was done. MM/PBSA was used in order to compute the binding free energies for the best 35 ICM energy score molecules from VS.

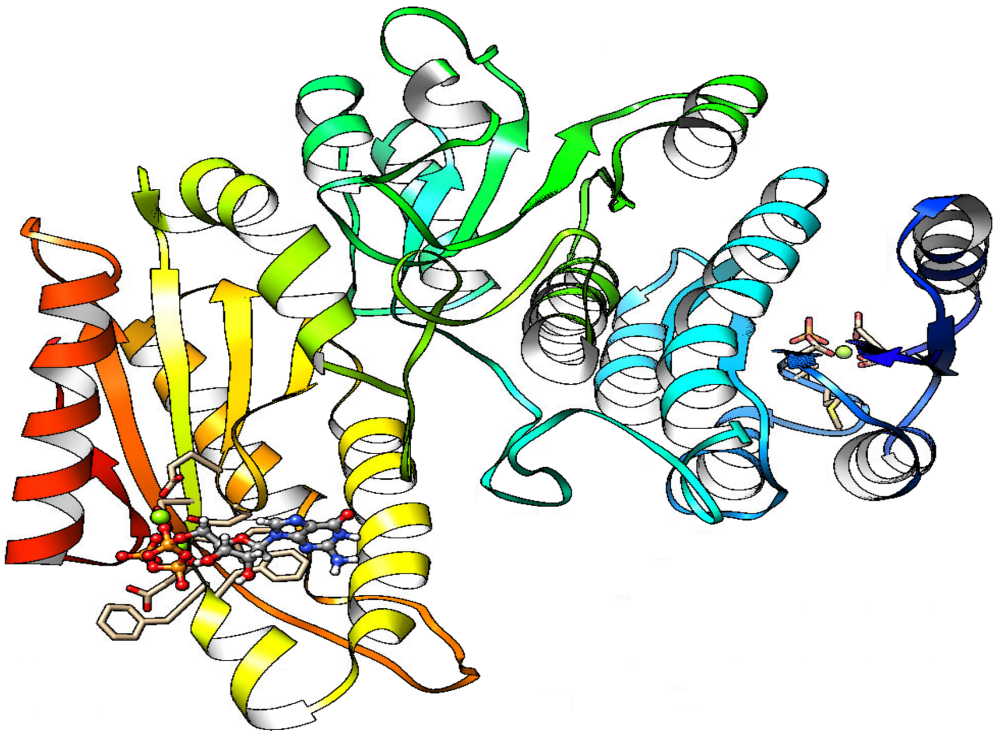

Figure 2. Average structure of PleD protein obtained from 20 ns of MD simulation. GTP, ASP52-phosphate moiety and Mg^{2+} ions are presented in the ball and stick representation.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 September 2020 doi:10.20944/preprints202009.0046.v1

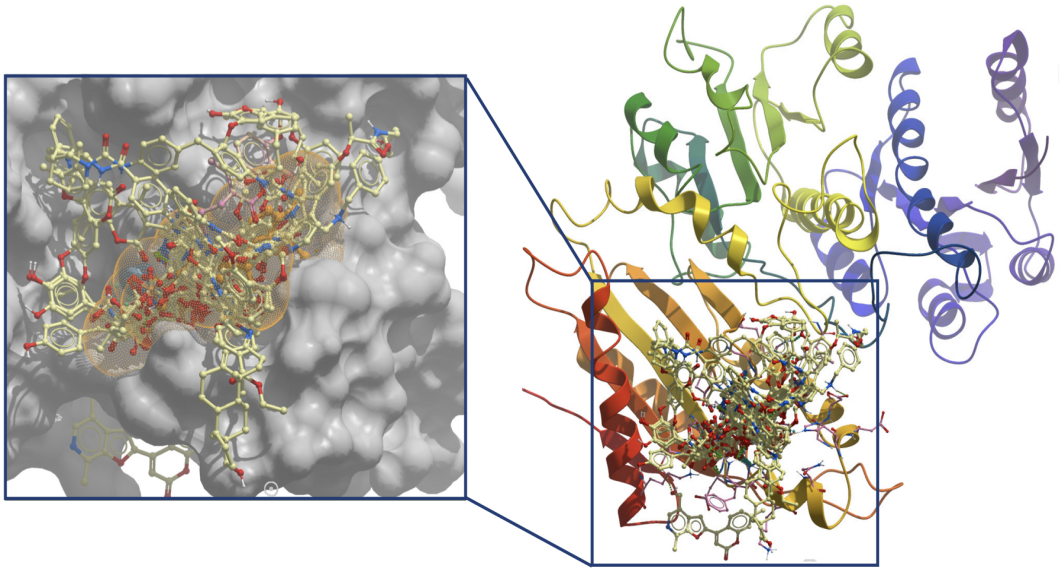

Figure 3. Best poses for the first 100 hits ranked according to ICM energy score. In orange dot envelope representation of the space found in the average structure for GTP.

2.1. Molecular Mechanics Poisson-Boltzmann Surface Area (MM/PBSA)

The best 35 ligands from the VS simulations were selected to perform a 10 ns of molecular dynamics simulation. From this results, the best 6 molecules with the best binding free energy MM/PBSA were studied. For the 35 ligands, the energy values were ranked between -121.39 (ZINC04501392) and -12.84 kcal/mol (ZINC15956889), the complete list is shown in the table S2 supplementary materials. On the other hand, the GTP molecule has a binding free energy of about -178.09 kcal/mol. The difference in energy from VS and MM/PBSA calculation lies not only in the method, but the parameter and charges used for this type of calculation (see methods).

The six ligands with the best binding free energy according to MM/PBSA calculations, are listed in table 1 with the ZINC15 id. In Figure 4 a 2D structure representation for the GTP and molecules are shown. Previous studies report antibacterial activity using one of these ligands or derivatives of them, for example, the citrate can be used to functionalize and synthesize nanoparticles with antibacterial activity [29,30]. In searching for new nematicidal factors from *Bacillus thuringiensis* against the plant-parasitic nematode *Meloidogyne incognita*, a component identified as *trans*-Aconitic acid (TAA) was found. The acid showed high nematicidal activity, suggesting that TAA is specifically synthesized by the bacillus as a virulence factor, TAA acid has shown activity nematicide for *Bacillus thuringiensis* bacterium [31].

Table 1. Show the ZINC15 id and the common name for the best six MM/PBSA binding free energy score compounds.

Zinc15 id	Name in this work	Common name
ZINC00895081	Ligand1	Citrate
ZINC03870145	Ligand2	Phosphoenolpyruvic acid
ZINC04028701	Ligand3	3-carboxy-2-(carboxymethyl)oxirane-2-carboxylate
ZINC04501392	Ligand4	*trans*-Aconitic acid
ZINC19336068	Ligand5	bis(4-hydroxy-2-oxo-2H-chromen-3-yl)acetic acid
ZINC27558828	Ligand6	2,3-Bis[(2E)-3-(4-hydroxyphenyl)-2-propenoyl]oxysuccinic acid

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 September 2020 doi:10.20944/preprints202009.0046.v1

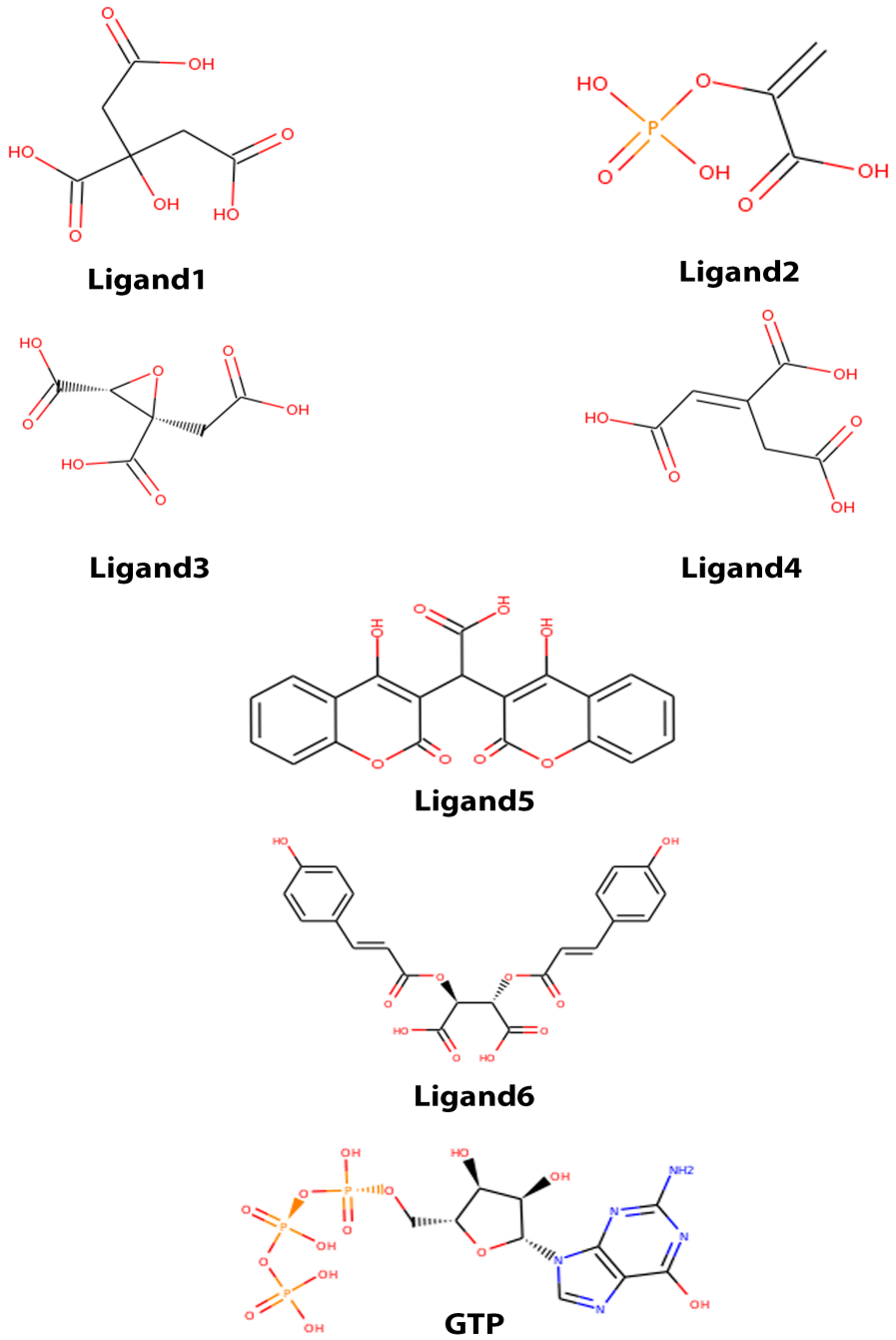

Figure 4. 2D structure of the six ligands with the best MM/PBSA performance.

The MM/PBSA calculation brings the information of how strong is the binding between protein and ligand, for this, the algorithm considers the binding free energy equal to the free energy of the complex minus the sum of the free energy of the protein plus the free energy of the ligand, see equation (1).

$$\Delta G = G_{complex} - (G_{protein} + G_{ligand}) \tag{1}$$

Each G component from equation (1) can be calculated as:

$$\Delta G = \langle E_{MM} \rangle - TS + \langle G_{solvation} \rangle \tag{2}$$

Where $\langle E_{MM} \rangle$ is the average molecular mechanics potential energy in vacuum, T is the temperature, S is the entropy and $\langle G_{solvation} \rangle$ is the free energy of solvation. The term $\langle E_{MM} \rangle$ include energy of both bonded and nonbonded interactions, E_{bonded} and $E_{nonbonded}$. The E_{bonded} are consider of bond, angle, dihedral and improper interactions. E_{bonded} consider electrostatic (E_{elec}) and van der Waals (E_{vdW}) interactions, see equation (3). $\langle G_{solvation} \rangle$ are included polar and non polar free energies, polar contribution is calculated using Poisson-Boltzmann equation and nonpolar contribution includes repulsive and attractive forces between solute and solvent generated by cavity, E_{cavity} formation and van der Waals interactions, E_vdW, see equation (4), for more information about how the equation works and how it is calculated consult the work of Kumari et al. [32], Baker et al. [33], Wisser et al. [34] and Konecny et al. [35].

$$\langle E_{MM} \rangle = E_{bonded} + E_{nonbonded} = E_{bonded} + E_{vdW} + E_{elec} \tag{3}$$

$$\langle G_{solvation} \rangle = G_{polar} + G_{nonpolar} = G_{polar} + E_{cavity} + E_{vdW} \tag{4}$$

A lower value of ΔG means a better coupling between protein and ligand. The table 2 shows the values obtained for the 6 ligands with the best binding free energy. Results show that non of the molecules tested get the same or close binding free energy as GTP, the difference between GTP and ligands energies are; Ligand1 78 kcal/mol, Ligand2 74.77 kcal/mol, Ligand3 58.51 kcal/mol, Ligand4 56.76 kcal/mol, Ligand5 70.16 kcal/mol and Ligand6 77.11 kcal/mol, the biggest difference is presented by Ligand1 and the lowest difference is with Ligand4. In order to have a better insight of which residues contribute more to the ΔG binding free energy, energy decomposition was performed (see methodology section for details), with this, the total ΔG energy is presented as a contribution per residue, with this information we can classify residues into favourable and non-favourable interactions.

Table 2. Contribution of energies in the MM/PBSA calculation, all energy values are in kcal/mol.

Ligand	VDWAALS	EEL	EPB	ENPOLAR	DELTA G
Ligand1	2.92 ± 4.30	-867.68 ± 21.67	766.94 ± 18.20	-1.64 ± 0.06	-99.46 ± 5.67
Ligand2	19.03 ± 4.87	-736.68 ± 24.60	615.55 ± 21.83	-1.22 ± 0.06	-103.32 ± 4.81
Ligand3	8.72 ± 5.14	-776.77 ± 38.28	650.15 ± 35.30	-1.67 ± 0.06	-119.58 ± 7.11
Ligand4	7.86 ± 4.83	-807.77 ± 33.78	680.01 ± 29.12	-1.43 ± 0.09	-121.33 ± 9.13
Ligand5	-2.20 ± 5.73	-686.56 ± 28.76	583.50 ± 27.10	-2.67 ± 0.13	-107.93 ± 8.92
Ligand6	-20.27 ± 4.39	-533.33 ± 19.94	456.46 ± 16.33	-3.84 ± 0.11	-100.98 ± 7.20
GTP	-10.86 ± 6.45	-1287.03 ± 35.53	1123.46 ± 29.55	3.49 ± 0.17	-178.09 ± 10.82

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 September 2020 doi:10.20944/preprints202009.0046.v1

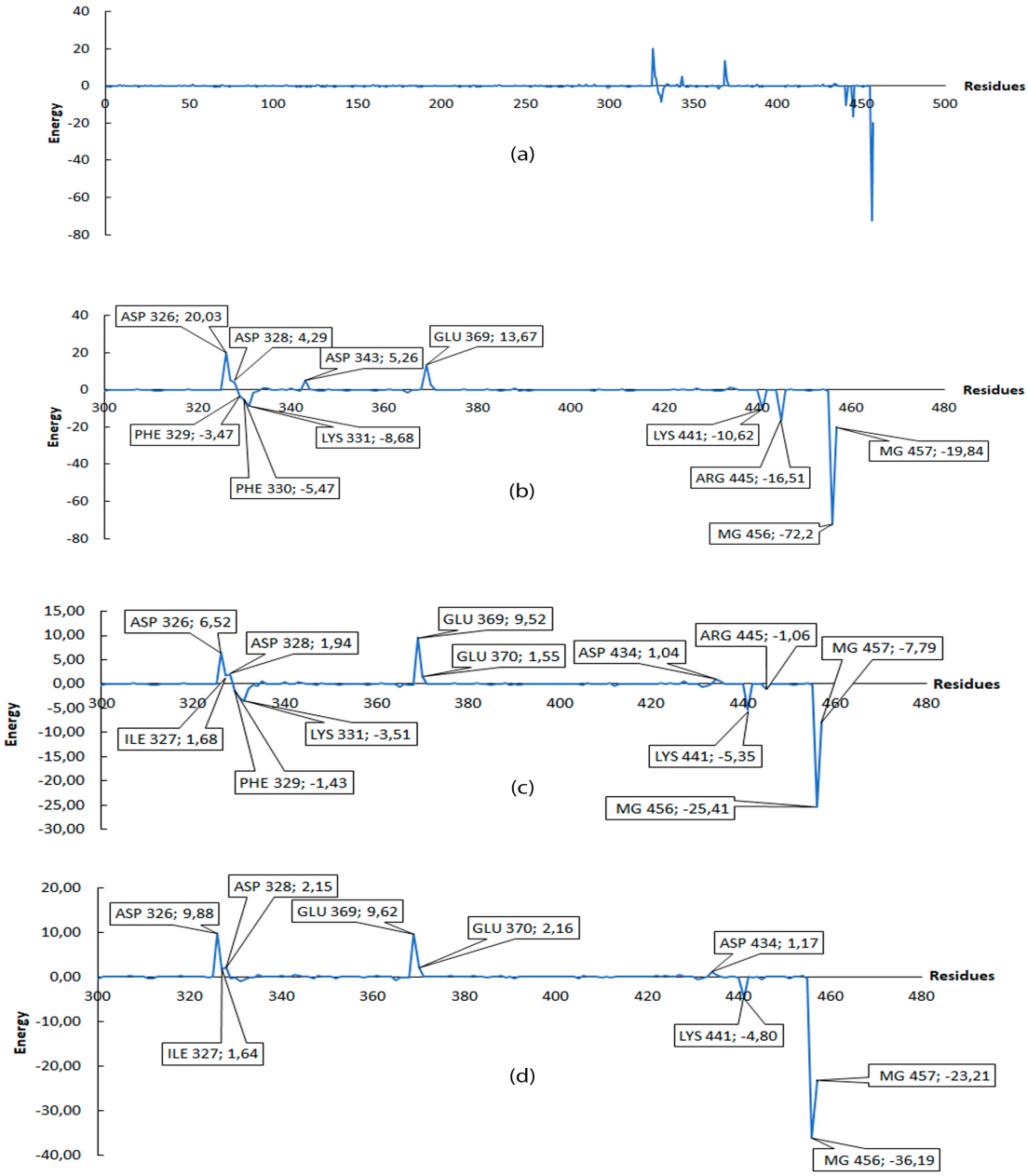

Figure 5. Interaction per residue, from residue 300 to 457, all energies were calculated in kcal/mol. **(a)** Contribution of all residues on the MM/PBSA energy for GTP. **(b)** Zoom from residue 300 to 457 for the GTP, **(c)** Ligand1, **(d)** Ligand2. Residues with a contribution higher than 1 or lower than -1 are marked.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 September 2020 doi:10.20944/preprints202009.0046.v1

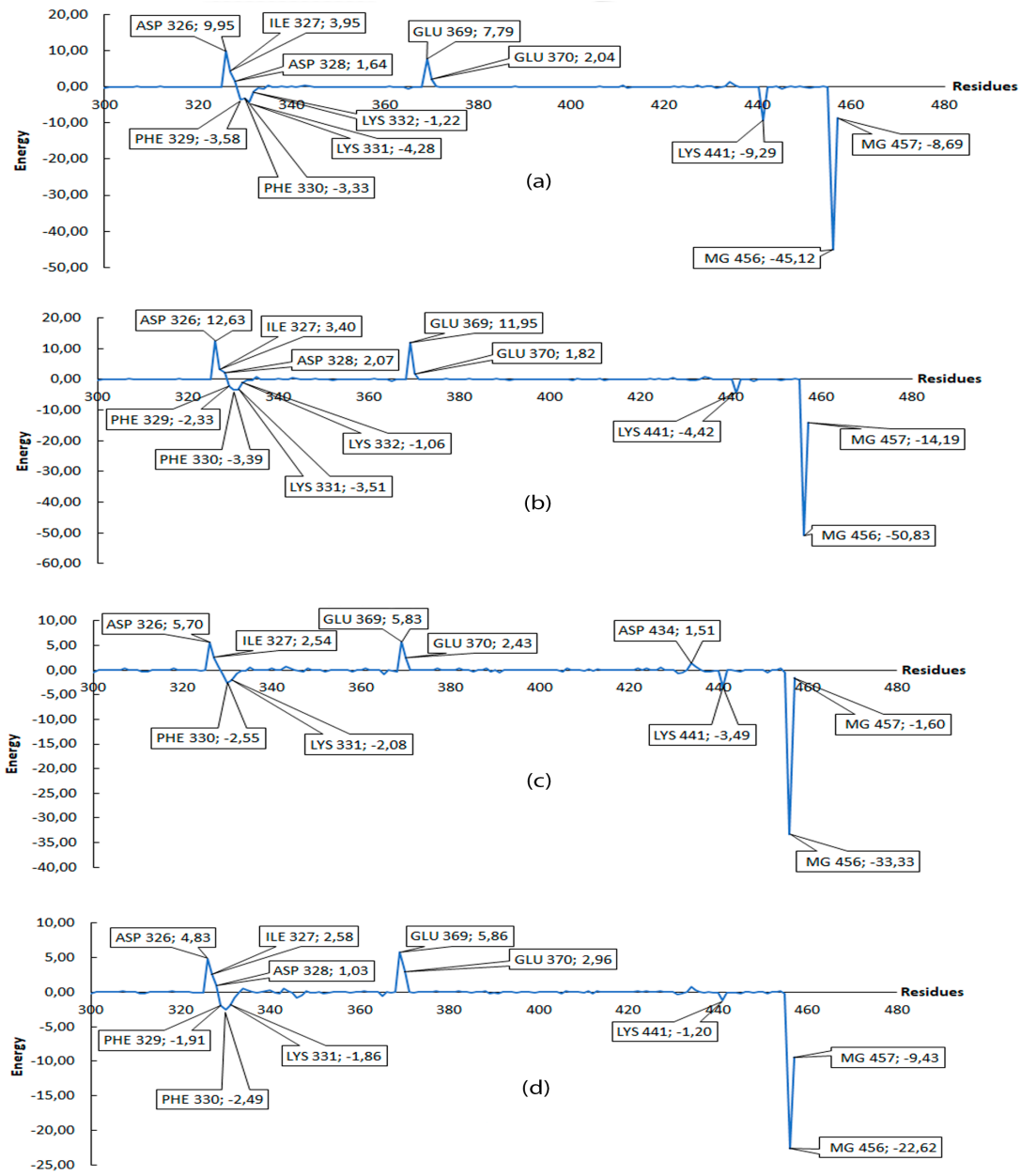

Figure 6. Interaction per residue from residue 300 to 457, all energies were calculated in kcal/mol. **(a)** Ligand3, **(b)** Ligand4, **(c)** Ligand5 and **(d)** from Ligand6. Residues with a contribution higher than 1 or lower than -1 are marked.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 September 2020 doi:10.20944/preprints202009.0046.v1

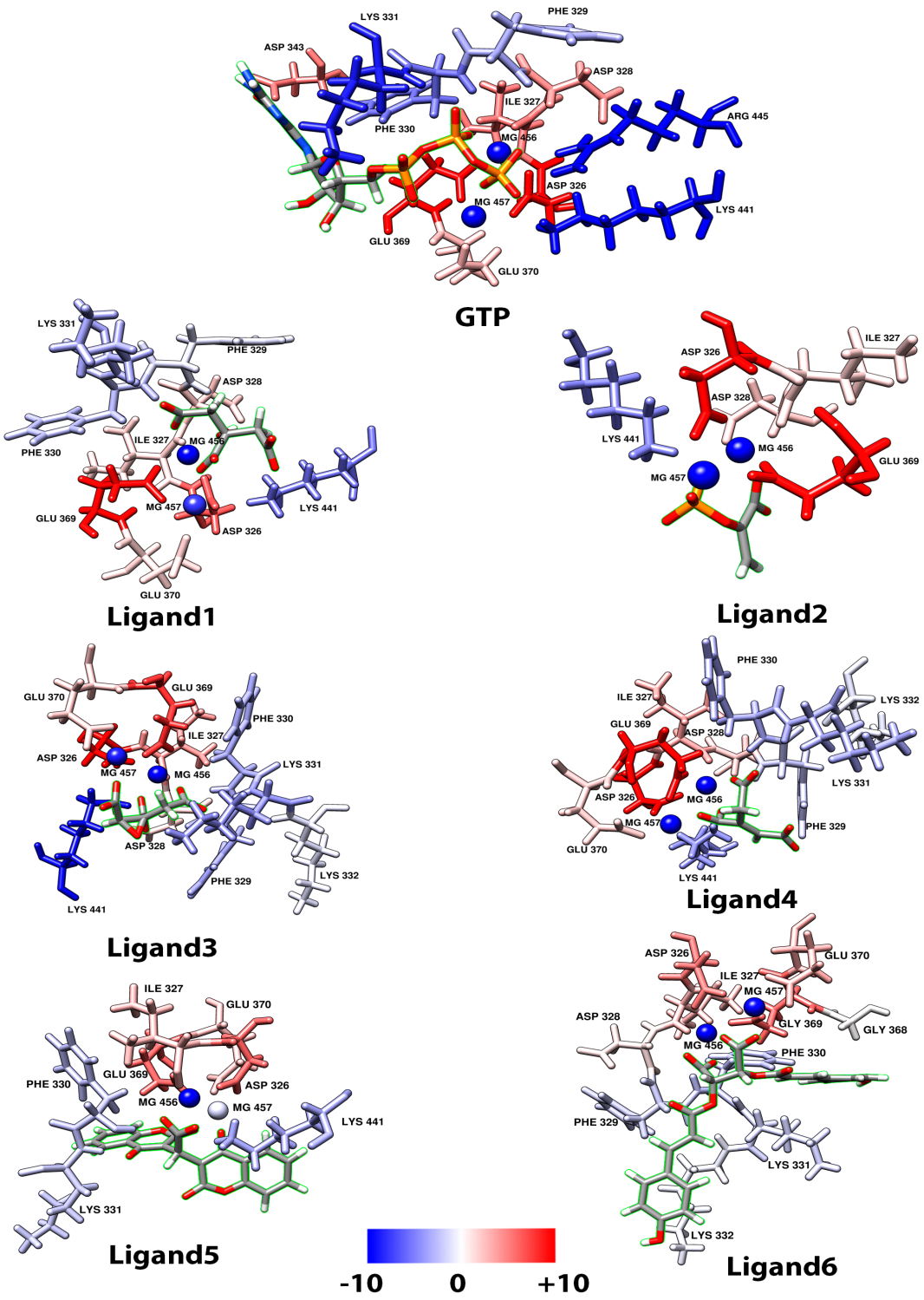

Figure 7. Contribution per residues around 3 Å from the ligand in the MM/PBSA calculation for each ligand and GTP, blue colour means more favourable interaction meanwhile red colour means non-favourable interaction, the ligand is highlighted in green.

Figure 5 only take into account the residues between 300 and 457, because as is shown in Figure 5a the first 300 residues do not show contribution to the binding free energy. For Figure 5 and Figure 6 negative values are considered as favourable interaction, the interactions between two of the three magnesium present in the protein play a role coordinating the carboxyl groups in all 6 ligands

including GTP which help to gain more stability. Taking the GTP as a reference, an interaction is observed with ARG445, this is a highly favourable interaction which is not present in any of the 6 ligands evaluated, suggesting this residue as an important residue in the binding pocket which help to reach a more stable conformational configuration. The ligands show common interactions with LYS331, LYS441, PHE329 and PHE330 as favourable except for the Ligand2 presumably because this ligand is the shortest one limiting its interaction with the protein. The result of positive energy suggest a non-favourable interaction with the protein residues, all ligands show non-favourable interaction with ASP326, ASP328, GLU369, ILE327 and GLU370, these four residues are common and have the higher positive values among the ligands. All ligands except for Ligand4 and Ligand6 show non-favourable interaction with ASP434. The non-favourable interaction for the GTP is mainly with ASP326 and GLU369.

2.1.1. Root Mean Square Deviation (RMSD)

From the 10 ns molecular dynamic simulation used for the calculation of MM/PBSA, a RMSD analysis was done. The RMSD was used to verify the protein stability during the entire simulation [36], less deviation in these value means more stability in the protein. For these, the table 3 show the average of the RMSD for the 6 ligands and the GTP. As the Figure 8 show the 6 ligands stabilize around 7 ns of simulation and GTP around 4 ns. These values were obtained using cpptraj tool [37] aligning to the first frame of the production part of the simulation, the result suggests a stable configuration of the protein with the ligands as we can see the same behaviour in the Figure 8, GTP shows less variation compared to the rest of the ligands.

Table 3. The average RMSD for each Pled-ligand and Pled-GTP complexes, in angstroms with its standard deviation considering the entire production.

Ligand	RMSD	Difference between GTP and ligand
Ligand1	2.261 ± 0.365	0.284
Ligand2	2.189 ± 0.336	0.212
Ligand3	2.165 ± 0.411	0.188
Ligand4	2.480 ± 0.255	0.503
Ligand5	2.506 ± 0.350	0.529
Ligand6	2.305 ± 0.389	0.328
GTP	1.977 ± 0.289	0

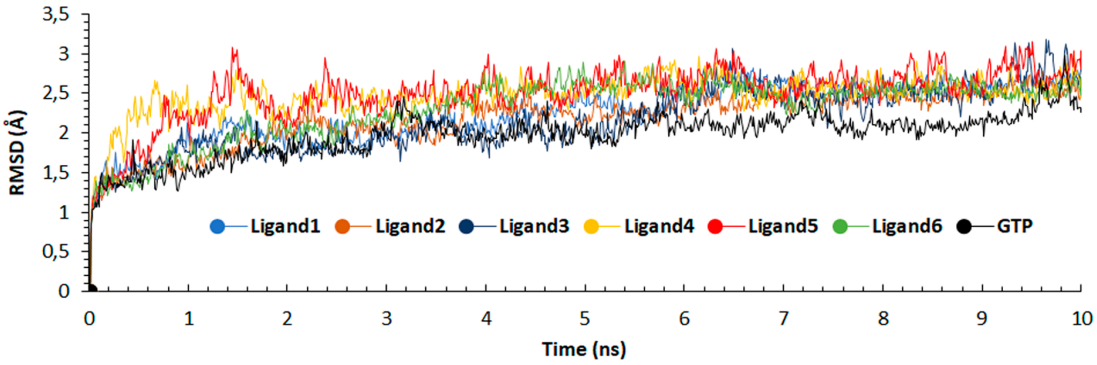

Figure 8. RMSD calculation for the 10 ns of production for each ligand and GTP.

2.1.2. Hydrogen bonds (H-bonds)

Hydrogen bonds represent an important interaction between protein-ligand complexes and the main reason for protein-ligand selectivity, this is why hydrogen bond analysis is often used to describe affinity between protein and ligand [38]. Using hbond function in the cpptraj tool we analyze

the average amount of time at which hydrogen bond was present during the simulation, from this information, it can be seen, which ones among all the residues, in the protein have more H-bond formation with the ligands. For these, the H-bond analysis for each ligand (including GTP) was done using a threshold of 3 Å around the ligands during all the production taking in account frame by frame.

Table 4. Residues involved in the H-bond formation with the percentage of how long this bond was present during the entire 10 ns of simulation.

Ligand	Residues involve in the h-bond formation with the major contribution
Ligand1	LYS441(54.20%), LYS331(35.88%), PHE330(7.11%), PHE329(2.80%)
Ligand2	LYS441(85.28%), PHE329(9.37%), PHE330 (4.97%)
Ligand3	LYS441(38.70%), PHE329(24.17%), PHE330(19.74%), LYS331(17.39%)
Ligand4	LYS441(52.68%), LYS331(24.05%), PHE330(19.28%), PHE329(3.99%)
Ligand5	LYS441(54.31%), PHE330(35.45%), LYS331(9.31%)
Ligand6	LYS441(36.13%), LYS331(23.48%), PHE330(17.81%), ASN334(17.03%), LYS332(4.77%)
GTP	ARG445(39.05%), LYS331(34.63%), LYS441(17.35%), PHE330(8.12%)

Among all ligands evaluated, the LYS441 is the more important residue in the H-bond formation. In some cases being present in more than the 80% of the production (Ligand2). Residues PHE331, PHE330 and LYS331 were also present as relevant. For the GTP the residue with more interactions was the ARG445, this also is seen the MM/PBSA decomposition results.

A second analysis was performed to observe the behaviour of the best ligand obtained in the MM/PBSA calculation, which is Ligand4. For this, 100 ns of molecular dynamics with the same condition of the previous 10 ns were done. The starting point for this new simulation was the ending of the previous one giving a total of 110 ns of simulation. In this second analysis, the PleD-protein without any ligand, to serve as a reference, the complex PleD-GTP and the complex PleD-Ligand4, were taken for simulations. MM/PBSA calculations were done, using different time windows, 20, 30, 40, 50 and 60 ns of the molecular simulation were consider, in each window, the same amount of snapshots (200) were taken.

2.2. Trans-Aconitic Acid (Ligand4)

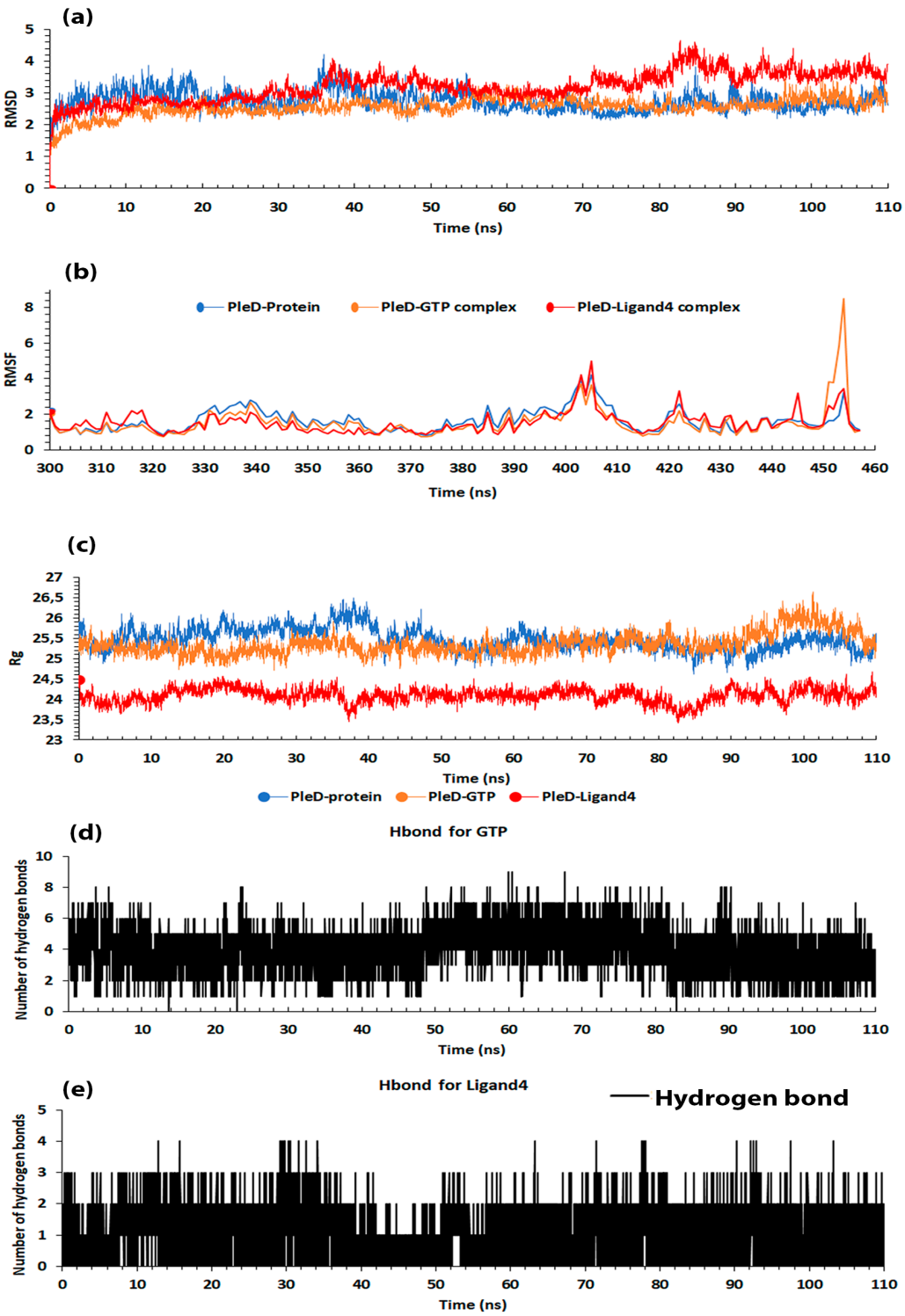

Figure 9. Analysis results from the second production of 110 ns, **(a)** RMSD in angstrom, **(b)** RMSF in angstrom, **(c)** Radius of gyration (Rg) in angstrom, **(d)** Hydrogen bonds for GTP and **(e)** Hydrogen bonds for Ligand4.

2.2.1. Root Mean Square Deviation (RMSD)

The RMSD show high stability in the three cases (PleD, GTP and Ligand4), 9.a, the average value with it standard deviation for the PleD-protein 2.75 ± 0.31, PleD-GTP complex 2.57 ± 0.28 and PleD-Ligand4 complex 3.20 ± 0.45. The PleD-protein present a higher value than PleD-GTP meaning GTP help the protein itself to reach a more stable configuration. The PleD-Ligand4 is the one who presents more variation although has high stability during the dynamic as is represented in the low standard deviation.

2.2.2. Root Mean Square Fluctuation (RMSF)

For the RMSF the same approach as the MM/PBSA was made, the graph just take into account residues 300-457 which are the residues were the interaction are establish, Figure 9b, the same fluctuation is observed in the protein structure for the three cases, specifically in the PleD-GTP, a major variation in residues 450-455 is present, the average variation with the standard deviation is 1.65 ± 0.63 for the PleD-protein, 1.53 ± 0.87 for PleD-GTP complex and 1.53 ± 0.62 for PleD-Ligand4 complex. The RMSF analysis agrees with the RMSD showing the PleD by itself present lower stability compared to a binding PleD with GTP.

2.2.3. Radius of gyration

The Radius of gyration (Rg) can give an idea of the compactness of a protein structure [39], the Figure 9c suggest structural stability in the 3 systems here considered however the complex PleD-Ligand4 less Rg value in comparison of the other two structures. The average value with the standard deviation for the three systems are 25.48 ± 0.27 for PleD-protein, 25.36 ± 0.26 PleD-GTP and 24.1 ± 0.17 PleD-Ligand4.

2.2.4. Hydrogen bond (H-bond)

The H-bond analysis results for the 110 ns production (5) show differences compared to the results from the table 4, for the GTP is observed the same tendency as the previous result, where the residue ARG445 is the one which is present for more time during the simulation, in this case, the order is maintained, the LYS331, LYS441 and PHE330 residues come after the ARG445. In the case of the Ligand4 there are different order, in this case, the residue PHE330 is the one with more presence followed by LYS331 and LYS441.

Table 5. Residues involved in the H-bond formation with the percentage of how long this bond was presented during the entire simulation for the 110 ns of simulation.

Ligand	Residues involve in the h-bond formation with the major contribution
PleD-Ligand4	PHE330(9.7%) ,LYS331(6.01%), LYS441(5.69%)
PleD-GTP	ARG445(41.44%), LYS331(15.14%), LYS441(12.07%), PHE330(11.95%)

The H-bond formation shows that, during the production, the GTP is making even nine H-bonds while the Ligand4 just make only four as is shown in Figure 9d and 9e. The GTP present the double of H-bond compared to the Ligand4, this is due to the phosphate group (the GTP has 3 of them) which provide more possible points to establish H-bonds thanks to the hydroxyl group (10 in the whole structure), instead of the Ligand4 which only have 6 hydroxyl group.

2.2.5. Molecular Mechanics Poisson-Boltzmann Surface Area (MM/PBSA)

The MM/PBSA values here presented are averages of the windows establish in this step, the values are (with the standard deviation); -169.93 ± 4.03 for PleD-GTP and -98.79 ± 0.43 for PleD-Ligand4. In the Figure 10 are represented the average values of the residues from the different windows (20, 30, 40, 50 and 60), compared to the previous results (Figure 5) for the GTP, here can be seen a similarity,

which indicates the same tendency for the residues like ASP326, ASP328 and GLU369 which have non-favourable interaction while ARG445, LYS441, LYS331, PHE330 and PHE329 have the major favourable contribution to the interaction with the GTP, suggesting the importance of these residues in the binding pocket. In the case of Ligand4, the residues showed the same tendency as in the simulation of 10 ns, ASP326, ILE327, ASP328, GLU369 and GLU370 residues, show non-favourable interaction while residues PHE329, LYS331 and LYS441 show favourable interaction. In both cases, the magnesium atoms present an important contribution to the binding free energy.

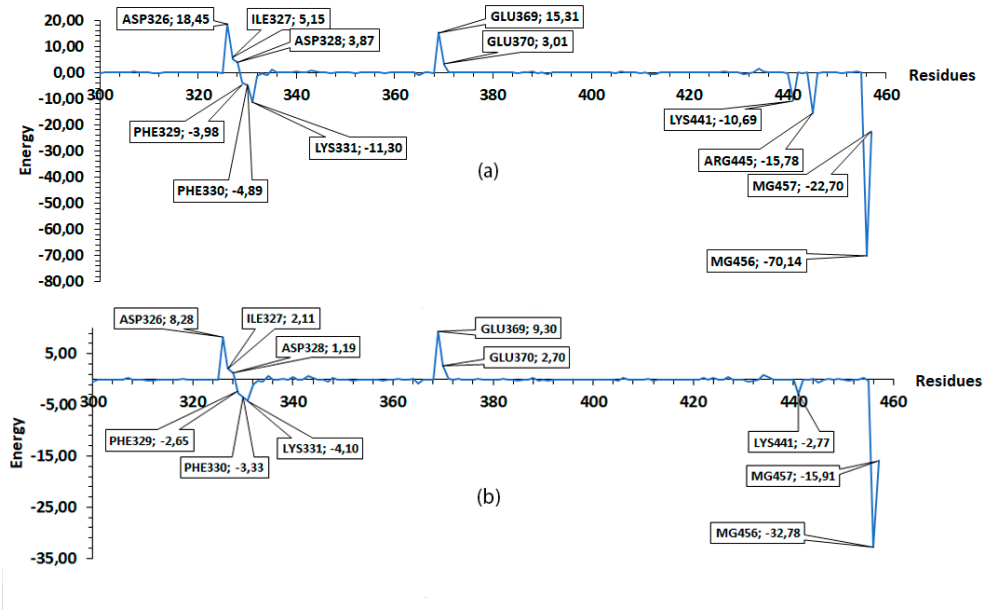

Figure 10. MM/PBSA Average ligand binding free energies for the 110 ns of simulation. Here, the residues 300 to 457 were taken into account. **(a)** GTP and **(b)** Ligand4.

The low standard deviation for the reported result at different windows is interpreted as high stability of the complexes, comparing the two MM/PBSA result (first and second analysis) some variation in energy are presented and the emergence or disappearance of new residues as GLU370 or ASP343 for GTP, and LYS332 for the Ligand4, is due to better conformational sampling obtained with more simulation time (110 ns). The result of the MM/PBSA with a short dynamic gives us an idea of the overall interaction and also of possible favourable or non-favourable ones.

For the GTP results, where the major contribution for the MM/PBSA binding free energy is the H-bond formation from ARG445, suggest this residue as a key one for the ligand-protein interaction. The absence of this interaction with the rest of the ligands can indicate a lower binding free energy in comparison with the GTP. Both MM/PBSA results, the one obtained by 10 ns and the average from the five windows, showed that the Ligand4 has good stability, suggesting this molecule as a possible template for design new ligand who can surpass the GTP binding free energy and effectively inhibit the PleD protein.

In this research the binding free energy was calculated by MM/PBSA method and was used to define a possible compound capable of inhibiting the PleD protein and hence blocking the formation of biofilms, the *trans*-aconitic acid performed the best among the 224205 ligands here studied. This compounds is found in plants like *Zea mays*, *Triticum aestivum*, *Avena sativa*, *Brachiaria plantaginea* and *Saccharum officinarum* [40]. Methodologies of extraction are described by Schnitzler et al. [41] and Kanitkar et al. [42]. In our results *trans*-aconitic acid does not surpass GTP binding free energy, however it does not mean that can not be used, our results suggest good interaction with the PleD protein, further optimization can be made to improve the results, strategies like Fragment-Based Drug

Discovery (FBDD) [43] or Fragment Molecular Orbital (FMO) [44,45] can be use to optimized the hit to become an effective lead.

3. Materials and Methods

3.1. Data collection

To search for antimicrobial agents from natural compounds, the ZINC15 [46] database was chosen. ZINC15 is a database with public access established to allow ready access to compounds for virtual screening. Molecules were downloaded in SDF format. ZINC 15 has a natural product-related compounds database of about 224205. All records from naturals products were download and stores in a single file.

3.2. Virtual screening (VS)

Structure-based virtual screening applying docking simulations was performed using ICM software package [28,47–54]. Virtual Ligand Screening (VLS) in ICM was performed by docking the database downloaded from ZINC15 against PleD protein structure obtained from protein data bank PDB (PDB:ID 1W25) followed by an evaluation of the docked conformation with a binding-score function. The search space encompassed the catalytic domain, the DGC (GGDEF) domain consists of a five-stranded central β-sheet surrounded by α helices (Figure 2). Water molecules and ligands had been removed from the original crystal structure. Phosphoryl analog BeF^{-3} attached to ASP52, was changed by phosphate group which was parameterized using the same procedure as was mention elsewhere [27]. A grid for space search was set up using as a reference the GTPαS and Mg^{2+} from 2V0N crystal structure, with the following dimensions in Å: centre (x, y, z) = (52.14.6, 76.2), dimensions (x, y, z) = (21.2, 17.9, 17.9). The docking simulation was then run at exhaustiveness of 5 and set to only output the lowest energy pose.

3.3. Ligand binding free energy calculations

For the ligands, no geometry optimization was applied because we kept the orientation obtained from the docking results. Charge for ligands were calculated with *ab initio* methods using B3LYP and 6-31G as basis set in GAUSSIAN09 software [55]. With the output file from gaussian09, charges and topology file (prepi and frcmod AMBER files respectively) were prepared using antechamber and parmchk2 from Ambertools18 [56]. Although antechamber generates a file with the charge information of ligands these charges were changes for the ones calculated with gaussian.

3.4. System Preparation

The system consists of the protein-ligand complex embedded in a rectangular box with waters and ions, LEaP module of Amber was used for the preparation of all systems [56]. Initially, the mbondi2 parameter for PBRadii were set [57]. The ff14SB force field was used for the protein which is an improvement of ff99SB [58]. 35 ligands with the best ICM energy score were chosen. The GAFF force field was used for the ligands[59], all the system were embedded in a box of size 100, 79, 109 Å (approximately) for X, Y, Z respectively, we use TIP3P water type molecules [60,61] and Na^{+} ions were add to neutralize the charge.

3.5. Molecular Dynamics (MD)

For all systems, first, a minimization of only waters molecules was done, for this, we restrain all the solute (protein and ligand) with a force of 100 kcal/mol-$Å^2$, in this minimization we establish the maxcyc and the ncyc to 50000 and 1500 respectively, the ntc was 1 (the shake was off) this is recommended to do it during the minimization [56], constant volume was established. Then a

minimization of the entire system using this time maxcyc=100000 and ncyc=1000, again constant volume was done.

After energy minimization the system was gradually heated until 300 K for 50 ps, in this part the solute was restrained with a force of 2.0 kcal/mol-Å^2, for the last step the system was equilibrated for 5 nanoseconds (ns), we establish ntt=3, ntp=1, ntc=2 and ntf=2, with constant pressure. Finally, for the first analysis, we perform the production for 10 ns in NPT ensemble.

The second analysis of PleD protein, PleD-GTP and PleD-Ligand4 complexes, a 100 ns more were simulated. For the PleD-GTP and PleD-Ligand4, the 100 ns production start at the end of the previous 10 ns, using the same conditions as the 10 ns production.

3.6. Molecular Mechanics Poisson-Boltzmann Surface Area (MM/PBSA) Calculation

The binding free energy was calculated using the MM/PBSA script from AMBER package [62], in the first analysis a total of 200 snapshots were taken for the MM/PBSA analysis, these snapshots are equal to the last 2 ns of simulation, Su et al suggest that working with data of MD trajectories longer than 2 ns decreases the accuracy of the prediction [57]. The parameters igb = 2, inp = 1 and radiopt = 0 were used, the igb parameter is chosen depending on the mbondi parameters following the amber manual indications, the inp uses the solvent-accessible-surface area to correlate total nonpolar solvation free energy and radiopt indicating to use the default atomic radius of the input file [56]. In the calculation of MM/PBSA binding free energy entropy was not considerate, this due to the high computational cost of it and this contribution just improve the result slightly or even can decrease the effectiveness of the MM/PBSA method [63–65]. For the decomposition per residue of the MM/PBSA result, the &decomp general option was established with idecomp=1 in the same input file.

In the second analysis, we take the last 20, 30, 40, 50, 60 ns of production and 200 snapshots, with the same condition as the first MM/PBSA calculation.

3.7. RMSD, RMSF, Radius of gyration and H-bond

The RMSD (Root Mean Square Deviation), the RMSF (Root Mean Square Fluctuation), the radius of gyration (Rg) and H-bond (hydrogen bond formation) were calculated with cpptraj tool [37], we use the entire production to make this analysis and the first frame was taken as reference for alignment in the RMSD and RMSF analysis.

4. Conclusions

Computational methods, using virtual screening, play a significant role in drug discovery. The identification of the drug interaction with the target macromolecule can help to clarify the potential molecule′s action mechanisms.

In our computational study, the binding free energy between PleD protein and ligands was calculated through MM/PBSA to rank the best molecules obtained from virtual screening of naturals products from ZINC15 database, we identify ARG445 as a key residue in the binding pocket, residues LYS441, LYS331, PHE 330 and PHE329 present favourable interaction and GLU369 present a common non-favourable interaction among the ligands here studied. Our results suggest *trans*-aconitic acid as a possible starting point for hit-to-lead methodologies.

Supplementary Materials: The following are available online at http://www.mdpi.com/1420-3049/xx/1/5/s1, Figure S1: title, Table S1: title, Video S1: title.

Author Contributions: Conceptualization, R.P.N, J.A.L.R and J.Y.; methodology, R.P.N and J.Y.; software, R.P.N and J.Y.; formal analysis, R.P.N and J.Y.; writing–original draft preparation, R.P.N.; writing–review and editing, R.P.N, J.A.L.R and J.Y.; visualization, R.P.N.; supervision, J.A.L.R and J.Y.;

Funding: This work was supported by "Convocatoria del Fondo de Ciencia, Tecnología e Innovación del Sistema General de Regalías para la conformación de una lista de proyectos elegibles para ser viabilizados, priorizados y aprobados por el OCAD en el marco del Programa de Becas de Excelencia Doctoral del Bicentenario 2019" which make possible the PhD studies of Roberto Pestana-Nobles.

Acknowledgments: This work was supported by "Convocatoria del Fondo de Ciencia, Tecnología e Innovación del Sistema General de Regalías para la conformación de una lista de proyectos elegibles para ser viabilizados, priorizados y aprobados por el OCAD en el marco del Programa de Becas de Excelencia Doctoral del Bicentenario 2019" which make possible the PhD studies of Roberto Pestana-Nobles.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

EPS	Extracellular polymeric substances
c-di-GMP	Cyclic diguanylate
DGC	Diguanylate cyclases
PDEs	Phosphodiesterases
GTP	Guanosine-5′-triphosphate
TTA	*trans*-aconitic acid
VS	Virtual screening
MD	molecular dynamics
MM/PBSA	Molecular Mechanics Poisson-Boltzmann Surface Area
MMFF	Merck Molecular Force Field

References

1. Flemming, H.C.; Wingender, J.; Szewzyk, U.; Steinberg, P.; Rice, S.A.; Kjelleberg, S. Biofilms: an emergent form of bacterial life. *Nature Reviews Microbiology* **2016**, *14*, 563–575. doi:10.1038/nrmicro.2016.94.
2. Yin, W.; Wang, Y.; Liu, L.; He, J. Biofilms: The Microbial "Protective Clothing" in Extreme Environments. *International journal of molecular sciences* **2019**, *20*, 3423. doi:10.3390/ijms20143423.
3. Tolker-Nielsen, T. Biofilm Development. *Microbiology Spectrum* **2015**, *3*. doi:10.1128/microbiolspec.MB-0001-2014.
4. Nadell, C.D.; Drescher, K.; Foster, K.R. Spatial structure, cooperation and competition in biofilms. *Nature Reviews Microbiology* **2016**, *14*, 589–600. doi:10.1038/nrmicro.2016.84.
5. Del Pozo, J.L. Biofilm-related disease. *Expert Review of Anti-Infective Therapy* **2018**, *16*, 51–65. doi:10.1080/14787210.2018.1417036.
6. Fleming, D.; Rumbaugh, K.P. Approaches to Dispersing Medical Biofilms. *Microorganisms* **2017**, *5*. doi:10.3390/microorganisms5020015.
7. Jamal, M.; Ahmad, W.; Andleeb, S.; Jalil, F.; Imran, M.; Nawaz, M.A.; Hussain, T.; Ali, M.; Rafiq, M.; Kamil, M.A. Bacterial biofilm and associated infections. *Journal of the Chinese Medical Association* **2018**, *81*, 7–11. doi:10.1016/j.jcma.2017.07.012.
8. Fernicola, S.; Paiardini, A.; Giardina, G.; Rampioni, G.; Leoni, L.; Cutruzzolà, F.; Rinaldo, S. In Silico Discovery and In Vitro Validation of Catechol-Containing Sulfonohydrazide Compounds as Potent Inhibitors of the Diguanylate Cyclase PleD. *Journal of Bacteriology* **2016**, *198*, 147–156. doi:10.1128/JB.00742-15.
9. Cai, Y.M.; Hutchin, A.; Craddock, J.; Walsh, M.A.; Webb, J.S.; Tews, I. Differential impact on motility and biofilm dispersal of closely related phosphodiesterases in Pseudomonas aeruginosa. *Scientific reports* **2020**, *10*, 6232–6232. doi:10.1038/s41598-020-63008-5.
10. Römling, U.; Galperin, M.Y.; Gomelsky, M. Cyclic di-GMP: the First 25 Years of a Universal Bacterial Second Messenger. *Microbiology and Molecular Biology Reviews* **2013**, *77*, 1–52. doi:10.1128/MMBR.00043-12.
11. Feirer, N.; Kim, D.; Xu, J.; Fernandez, N.; Waters, C.M.; Fuqua, C. The Agrobacterium tumefaciens CheY-like protein ClaR regulates biofilm formation. *Microbiology (Reading, England)* **2017**, *163*, 1680–1691. doi:10.1099/mic.0.000558.
12. Alviz-Gazitua, P.; Fuentes-Alburquenque, S.; Rojas, L.A.; Turner, R.J.; Guiliani, N.; Seeger, M. The Response of Cupriavidus metallidurans CH34 to Cadmium Involves Inhibition of the Initiation of Biofilm Formation, Decrease in Intracellular c-di-GMP Levels, and a Novel Metal Regulated Phosphodiesterase. *Frontiers in Microbiology* **2019**, *10*, 1499. doi:10.3389/fmicb.2019.01499.

13. Lage, O.M.; Ramos, M.C.; Calisto, R.; Almeida, E.; Vasconcelos, V.; Vicente, F. Current screening methodologies in drug discovery for selected human diseases. *Marine Drugs* **2018**, *16*, 1–31. doi:10.3390/md16080279.
14. Rossiter, S.E.; Fletcher, M.H.; Wuest, W.M. Natural Products as Platforms to Overcome Antibiotic Resistance. *Chemical Reviews* **2017**, *117*, 12415–12474. doi:10.1021/acs.chemrev.7b00283.
15. Herrmann, J.; Fayad, A.A.; Müller, R. Natural products from myxobacteria: Novel metabolites and bioactivities. *Natural Product Reports* **2017**, *34*, 135–160. doi:10.1039/c6np00106h.
16. Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G. Counting on natural products for drug design. *Nature Chemistry* **2016**, *8*, 531–541. doi:10.1038/nchem.2479.
17. Nofiani, R.; Weisberg, A.J.; Tsunoda, T.; Panjaitan, R.G.P.; Brilliantoro, R.; Chang, J.H.; Philmus, B.; Mahmud, T. Antibacterial Potential of Secondary Metabolites from Indonesian Marine Bacterial Symbionts. *International Journal of Microbiology* **2020**, *2020*, 8898631. doi:10.1155/2020/8898631.
18. Emiru, Y.K.; Siraj, E.A.; Teklehaimanot, T.T.; Amare, G.G. Antibacterial Potential of Aloe weloensis (Aloeacea) Leaf Latex against Gram-Positive and Gram-Negative Bacteria Strains. *International Journal of Microbiology* **2019**, *2019*, 5328238. doi:10.1155/2019/5328238.
19. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004**, *25*, 1605–1612. doi:10.1002/jcc.20084.
20. Burton, G.J.; Hecht, G.B.; Newton, A. Roles of the histidine protein kinase pleC in Caulobacter crescentus motility and chemotaxis. *Journal Of Bacteriology* **1997**, *179*, 5849–5853. doi:10.1128/jb.179.18.5849-5853.1997.
21. Paul, R.; Weiser, S.; Amiot, N.C.; Chan, C.; Schirmer, T.; Giese, B.; Jenal, U. Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. *Genes & Development* **2004**, *18*, 715–727. doi:10.1101/gad.289504.
22. Aldridge, P.; Paul, R.; Goymer, P.; Rainey, P.; Jenal, U. Role of the GGDEF regulator PleD in polar development of Caulobacter crescentus. *Molecular Microbiology* **2003**, *47*, 1695–1708. doi:10.1046/j.1365-2958.2003.03401.x.
23. Aldridge, P.; Jenal, U. Cell cycle-dependent degradation of a flagellar motor component requires a novel-type response regulator. *Molecular Microbiology* **1999**, *32*, 379–391. doi:10.1046/j.1365-2958.1999.01358.x.
24. Jenal, U. Cyclic di-guanosine-monophosphate comes of age: a novel secondary messenger involved in modulating cell surface structures in bacteria? *Current Opinion in Microbiology* **2004**, *7*, 185–191. doi:10.1016/j.mib.2004.02.007.
25. Jenal, U.; Malone, J. Mechanisms of cyclic-di-GMP signaling in bacteria. *Annual Review of Genetics* **2006**, *40*, 385–407. doi:10.1146/annurev.genet.40.110405.090423.
26. Tamayo, R.; Pratt, J.T.; Camilli, A. Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. *Annual Review of Microbiology* **2007**, *61*, 131–48. doi:10.1146/annurev.micro.61.080706.093426.
27. Yosa Reyes, J.; Nagy, T.; Meuwly, M. Competitive reaction pathways in vibrationally induced photodissociation of H2SO4. *Phys. Chem. Chem. Phys.* **2014**, *16*, 18533–18544. doi:10.1039/C4CP01832J.
28. Neves, M.A.C.; Totrov, M.; Abagyan, R. Docking and scoring with ICM: the benchmarking results and strategies for improvement. *Journal of computer-aided molecular design* **2012**, *26*, 675–686. doi:10.1007/s10822-012-9547-0.
29. Khatoon, U.T.; Nageswara Rao, G.V.S.; Mohan, K.M.; Ramanaviciene, A.; Ramanavicius, A. Antibacterial and antifungal activity of silver nanospheres synthesized by tri-sodium citrate assisted chemical approach. *Vacuum* **2017**, *146*, 259–265. doi:10.1016/j.vacuum.2017.10.003.
30. Choudhury, R.; Majumdar, M.; Biswas, P.; Khan, S.; Misra, T.K. Kinetic study of functionalization of citrate stabilized silver nanoparticles with catechol and its anti-biofilm activity. *Nano-Structures & Nano-Objects* **2019**, *19*, 100326. doi:10.1016/j.nanoso.2019.100326.
31. Du, C.; Cao, S.; Shi, X.; Nie, X.; Zheng, J.; Deng, Y.; Ruan, L.; Peng, D.; Sun, M. Genetic and Biochemical Characterization of a Gene Operon for trans-Aconitic Acid, a Novel Nematicide from Bacillus thuringiensis. *Journal of Biological Chemistry* **2017**, *292*, 3517–3530. doi:10.1074/jbc.M116.762666.
32. Kumari, R.; Kumar, R.; Lynn, A. g_mmpbsa—A GROMACS Tool for High-Throughput MM-PBSA Calculations. *Journal of Chemical Information and Modeling* **2014**, *54*, 1951–1962. doi:10.1021/ci500020m.

33. Baker, N.A.; Sept, D.; Holst, M.J.; McCammon, J.A. The adaptive multilevel finite element solution of the Poisson-Boltzmann equation on massively parallel computers. *IBM Journal of Research and Development* **2001**, *45*, 427–438. doi:10.1147/rd.453.0427.
34. Weiser, J.; Shenkin, P.S.; Still, W.C. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *Journal of Computational Chemistry* **1999**, *20*, 217–230. doi:10.1002/(SICI)1096-987X(19990130)20:2<217::AID-JCC4>3.0.CO;2-A.
35. Konecny, R.B.; McCammon, N.A..; Andrew, J. iAPBS: a programming interface to the adaptive Poisson-Boltzmann solver. *Comput. Sci. Disc., 2012* **2012**, *5*. doi:10.1088/1749-4699/5/1/015005.
36. Sargsyan, K.; Grauffel, C.; Lim, C. How Molecular Size Impacts RMSD Applications in Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation* **2017**, *13*, 1518–1524. doi:10.1021/acs.jctc.7b00028.
37. Roe, D.R.; Cheatham, T.E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **2013**, *9*, 3084–3095. doi:10.1021/ct400341p.
38. Nittinger, E.; Inhester, T.; Bietz, S.; Meyder, A.; Schomburg, K.T.; Lange, G.; Klein, R.; Rarey, M. Large-Scale Analysis of Hydrogen Bond Interaction Patterns in Protein–Ligand Interfaces. *Journal of Medicinal Chemistry* **2017**, *60*, 4245–4257. doi:10.1021/acs.jmedchem.7b00101.
39. Lobanov, M.Y.; Bogatyreva, N.S.; Galzitskaya, O.V. Radius of gyration as an indicator of protein structure compactness. *Molecular Biology* **2008**, *42*, 623–628. doi:10.1134/S0026893308040195.
40. Bortolo, T.d.S.C.; Marchiosi, R.; Viganó, J.; Siqueira-Soares, R.d.C.; Ferro, A.P.; Barreto, G.E.; Bido, G.d.S.; Abrahão, J.; dos Santos, W.D.; Ferrarese-Filho, O. Trans-aconitic acid inhibits the growth and photosynthesis of Glycine max. *Plant Physiology and Biochemistry* **2018**, *132*, 490–496. doi:10.1016/j.plaphy.2018.09.036.
41. Schnitzler, M.; Petereit, F.; Nahrstedt, A. Trans-Aconitic acid, glucosylflavones and hydroxycinnamoyltartaric acids from the leaves of Echinodorus grandiflorus ssp. aureus, a Brazilian medicinal plant. *Revista Brasileira de Farmacognosia* **2007**, *17*, 149–154. doi:10.1590/S0102-695X2007000200002.
42. Kanitkar, A.; Aita, G.; Madsen, L. The recovery of polymerization grade aconitic acid from sugarcane molasses. *Journal of Chemical Technology & Biotechnology* **2013**, *88*, 2188–2192. doi:10.1002/jctb.4084.
43. De Souza Neto, L.R.; Moreira-Filho, J.T.; Neves, B.J.; Maidana, R.L.B.R.; Guimarães, A.C.R.; Furnham, N.; Andrade, C.H.; Silva Jr., F.P. In silico Strategies to Support Fragment-to-Lead Optimization in Drug Discovery. *Frontiers in chemistry* **2020**, *8*, 93–93. doi:10.3389/fchem.2020.00093.
44. Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment molecular orbital method: an approximate computational method for large molecules. *Chemical Physics Letters* **1999**, *313*, 701 – 706. doi:10.1016/S0009-2614(99)00874-X.
45. Hevener, K.E.; Pesavento, R.; Ren, J.; Lee, H.; Ratia, K.; Johnson, M.E. Chapter Twelve - Hit-to-Lead: Hit Validation and Assessment. In *Modern Approaches in Drug Discovery*; Lesburg, C.A., Ed.; Academic Press, 2018; Vol. 610, *Methods in Enzymology*, pp. 265 – 309. doi:https://doi.org/10.1016/bs.mie.2018.09.022.
46. Sterling, T.; Irwin, J.J. ZINC 15 – Ligand Discovery for Everyone. *Journal of Chemical Information and Modeling* **2015**, *55*, 2324–2337. doi:10.1021/acs.jcim.5b00559.
47. Abagyan, R.; Totrov, M.; Kuznetsov, D. ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *Journal of Computational Chemistry* **1994**, *15*, 488–506. doi:10.1002/jcc.540150503.
48. Abagyan, R.; Totrov, M. Biased Probability Monte Carlo Conformational Searches and Electrostatic Calculations for Peptides and Proteins. *Journal of Molecular Biology* **1994**, *235*, 983 – 1002. doi:10.1006/jmbi.1994.1052.
49. Totrov, M.; Abagyan, R. Rapid boundary element solvation electrostatics calculations in folding simulations: Successful folding of a 23-residue peptide. *Peptide Science* **2001**, *60*, 124–133. doi:10.1002/1097-0282(2001)60:2<124::AID-BIP1008>3.0.CO;2-S.
50. An, J.; Totrov, M.; Abagyan, R. Pocketome via Comprehensive Identification and Classification of Ligand Binding Envelopes. *Molecular & Cellular Proteomics* **2005**, *4*, 752–761. doi:10.1074/mcp.M400159-MCP200.
51. Fernandez-Recio, J.; Totrov, M.; Skorodumov, C.; Abagyan, R. Optimal docking area: A new method for predicting protein–protein interaction sites. *Proteins: Structure, Function, and Bioinformatics* **2005**, *58*, 134–143. doi:10.1002/prot.20285.

52. Fernandez-Recio, J.; Totrov, M.; Abagyan, R. ICM-DISCO docking by global energy optimization with fully flexible side-chains. *Proteins: Structure, Function, and Bioinformatics* **2003**, *52*, 113–117. doi:10.1002/prot.10383.
53. Méndez, R.; Leplae, R.; Lensink, M.F.; Wodak, S.J. Assessment of CAPRI predictions in rounds 3–5 shows progress in docking procedures. *Proteins: Structure, Function, and Bioinformatics* **2005**, *60*, 150–169. doi:10.1002/prot.20551.
54. Méndez, R.; Leplae, R.; De Maria, L.; Wodak, S.J. Assessment of blind predictions of protein–protein interactions: Current status of docking methods. *Proteins: Structure, Function, and Bioinformatics* **2003**, *52*, 51–67. doi:10.1002/prot.10393.
55. Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H.; Izmaylov, A.; Bloino, J.; Zheng, G.; Sonnenberg, J.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J.; Peralta, J.; Ogliaro, F.; Bearpark, M.; Heyd, J.; Brothers, E.; Kudin, K.; Staroverov, V.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J.; Iyengar, S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J.; Klene, M.; Knox, J.; Cross, J.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R.; Yazyev, O.; Austin, A.; Cammi, R.; Pomelli, C.; Ochterski, J.; Martin, R.; Morokuma, K.; Zakrzewski, V.; Voth, G.; Salvador, P.; Dannenberg, J.; Dapprich, S.; Daniels, A.; Farkas, O.; Foresman, J.; Ortiz, J.; Cioslowski, J.; Fox, D. *Gaussian 09*; Gaussian, Inc: Wallingford, CT, 2009.
56. Case, D.; Ben-Shalom, I.; Brozell, S.; Cerutti, D.; Cheatham, III, T.; Cruzeiro, V.; Darden, T.; Duke, R.; Ghoreishi, D.; Gilson, M.; Gohlke, H.; Goetz, A.; Greene, D.; Harris, R.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D.; Merz, K.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R.; Wang, J.; Wei, H.; Wolf, R.; Wu, X.; Xiao, L.; York, D.; Kollman, P. *AMBER 2018*; University of California: San Francisco, CA, 2018.
57. Su, P.C.; Tsai, C.C.; Mehboob, S.; Hevener, K.E.; Johnson, M.E. Comparison of radii sets, entropy, QM methods, and sampling on MM-PBSA, MM-GBSA, and QM/MM-GBSA ligand binding energies of F. tularensis enoyl-ACP reductase (FabI). *Journal of Computational Chemistry* **2015**, *36*, 1859–1873. doi:10.1002/jcc.24011.
58. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. ff 14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff 99SB. *Journal of chemical theory and computation* **2015**, *18*, 3696–3713. doi:10.1021/acs.jctc.5b00255.
59. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general Amber force field. *Journal of Computational Chemistry* **2004**, *25*, 1157–1174. doi:10.1002/jcc.20035.
60. Onufriev, A.V.; Izadi, S. Water models for biomolecular simulations. *WIREs Computational Molecular Science* **2018**, *8*, e1347. doi:10.1002/wcms.1347.
61. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926–935. doi:10.1063/1.445869.
62. Miller, B.R.; McGee, T.D.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. MMPBSA.py: An efficient program for end-state free energy calculations. *Journal of Chemical Theory and Computation* **2012**, *8*, 3314–3321. doi:10.1021/ct300418h.
63. Ben-Shalom, I.Y.; Pfeiffer-Marek, S.; Baringhaus, K.H.; Gohlke, H. Efficient Approximation of Ligand Rotational and Translational Entropy Changes upon Binding for Use in MM-PBSA Calculations. *Journal of Chemical Information and Modeling* **2017**, *57*, 170–189. doi:10.1021/acs.jcim.6b00373.
64. Genheden, S.; Ryde, U. Comparison of the Efficiency of the LIE and MM/GBSA Methods to Calculate Ligand-Binding Energies. *Journal of Chemical Theory and Computation* **2011**, *7*, 3768–3778. doi:10.1021/ct200163c.
65. Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of chemical information and modeling* **2011**, *51*, 69–82. doi:10.1021/ci100275a.