



Article

# Structural Requirements of *N*-alpha-Mercaptoacetyl Dipeptide (NAMdP) Inhibitors of *Pseudomonas Aeruginosa* Virulence Factor LasB: 3D-QSAR, Molecular Docking, and Interaction Fingerprint Studies

José Luis Velázquez-Libera <sup>1</sup>, Juliana Andrea Murillo-López <sup>2</sup>, Alexander F. de la Torre <sup>3,\*</sup> and Julio Caballero <sup>1,\*</sup>

<sup>1</sup> Centro de Bioinformática y Simulación Molecular (CBSM), Universidad de Talca, Casilla 747, Talca 3460000, Chile; josevlibera2010@gmail.com

<sup>2</sup> Departamento de Ciencias Químicas, Facultad de Ciencias Exactas, Universidad Andres Bello, Autopista Concepción–Talcahuano 7100, Talcahuano 4260000, Chile; jamurillo1987@gmail.com

<sup>3</sup> Departamento de Química Orgánica, Facultad de Ciencias Químicas, Universidad de Concepción, Concepción 4030000, Chile

\* Correspondence: jcaballero@utalca.cl (J.C.); afernandezd@udec.cl (A.F.d.l.T.); Tel.: +56-7124-188-50 (J.C.); +56-4122-033-17 (A.F.d.l.T.)

Received: 18 November 2019; Accepted: 3 December 2019; Published: 5 December 2019



**Abstract:** The zinc metallopeptidase *Pseudomonas* elastase (LasB) is a virulence factor of *Pseudomonas aeruginosa* (*P. aeruginosa*), a pathogenic bacterium that can cause nosocomial infections. The present study relates the structural analysis of 118 *N*-alpha-mercaptoacetyl dipeptides (NAMdPs) as LasB inhibitors. Field-based 3D-QSAR and molecular docking methods were employed to describe the essential interactions between NAMdPs and LasB binding sites, and the chemical features that determine their differential activities. We report a predictive 3D-QSAR model that was developed according to the internal and external validation tests. The best model, including steric, electrostatic, hydrogen bond donor, hydrogen bond acceptor, and hydrophobic fields, was found to depict a three-dimensional map with the local positive and negative effects of these chemotypes on the LasB inhibitory activities. Furthermore, molecular docking experiments yielded bioactive conformations of NAMdPs inside the LasB binding site. The series of NAMdPs adopted a similar orientation with respect to phosphoramidon within the LasB binding site (crystallographic reference), where the backbone atoms of NAMdPs are hydrogen-bonded to the LasB residues N112, A113, and R198, similarly to phosphoramidon. Our study also included a deep description of the residues involved in the protein–ligand interaction patterns for the whole set of NAMdPs, through the use of interaction fingerprints (IFPs).

**Keywords:** *Pseudomonas aeruginosa* elastase; *N*-alpha-mercaptoacetyl dipeptides; 3D-QSAR; docking; interaction fingerprints

## 1. Introduction

Nosocomial infections—also known as hospital acquired infections—are general or localized infectious processes in organs or anatomical regions, acquired by patients during hospitalization or through visits to other local health care centers. These undesired infections are considered to be a big global problem due to their social and economic impact [1]. The ongoing abuse of antibiotics to treat nosocomial infections has led to the problem of multidrug resistance [1], which is one of the major

challenges of antimicrobial discovery [2–5]. In the past, the assembly required for antibiotic drugs development [6,7] was aimed to obtain compounds that inactivate bacterial targets through specific mechanisms (i.e., inhibition of cell growth or causing cell death) [8]. As is known, pathogenic bacteria can produce a wide range of virulence factors that specifically participate in host functions, in order to allow colonization. Research is currently directed to seek specific antimicrobial agents focused on the action mechanisms of virulence factors [9].

*Pseudomonas aeruginosa* (*P. aeruginosa*) is a Gram-negative opportunistic pathogenic bacterium that is responsible for many nosocomial infections [10,11]. It has multiple virulence factors, such as the toxin metalloprotease *Pseudomonas elastase* (LasB), which is responsible for lung hemorrhages and corneal tissue destruction. Additionally, LasB can inactivate the alpha 1 proteinase inhibitor that controls tissue destruction through the degradation of a series of proteins, including elastin, collagen, and fibrin [12]. The idea that decreasing the function of virulence factors leads to less resistance of *P. aeruginosa* to antibiotics is still valid and was previously afforded [5,9]. In this context, some scientific endeavors have focused on developing new molecules that contain specific functional groups capable of interacting with the zinc ion of LasB metalloprotease [13–15], and additional chemical groups that are optimized to form specific interactions with the residues in the binding site of this target protein.

To support these tasks, theoretical studies oriented to characterize inhibitor interactions with the LasB crystal enzyme [16] could help with the development of new specific drugs to avoid antibiotic resistance in *P. aeruginosa*. Despite the widespread use of computational methods for drug design, there are a few studies related to the LasB inhibitors [17–19]. Fortunately, LasB–ligand complexes have been reported by X-ray crystallography [19]. Notwithstanding of several biological evaluations of sets of LasB inhibitors [13–15,17,18,20,21], a quantitative structure–activity relationship (QSAR) to predict and correlate the efficiency of the molecules reported is not present in the literature; neither is a deep description of the LasB binding site.

Inspired by the low number of theoretical studies dedicated to LasB inhibitors, we carried out QSAR and docking studies of the congeneric family of 118 *N*-alpha-mercaptoacetyl dipeptides (NAMdPs) reported by Cathcart et al. [17,20], providing interesting information about their binding poses and the causes of their differential activities. We assumed that this information could be useful for the design of new potential LasB inhibitors.

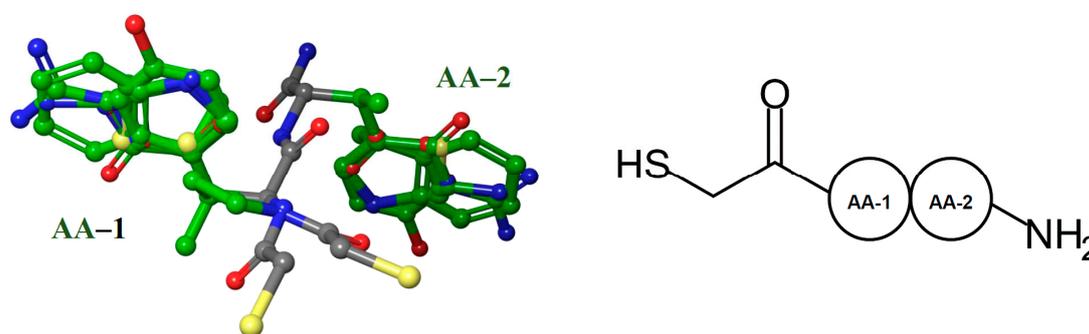
## 2. Results and Discussion

### 2.1. Results of the QSAR Models

First, the 118 NAMdPs structures were aligned (Figure 1). In order to better understand and visualize the structure–activity relationship (SAR), the amino acid residues (AA) of NAMdPs were denoted as AA-1 and AA-2, as shown in Figure 1. The alignment was done manually by using the Maestro program. In this way, we set the HSCH<sub>2</sub>–substituent and the amide groups as the backbone of the skeleton of peptides, and the substituents as AA-1 and AA-2, which were all different. When the AA-1 was proline, the HSCH<sub>2</sub>–groups were not typical because of the steric restrictions of the proline ring.

The field-based 3D-QSAR models were constructed by including the five available field descriptors in Phase (Phase, Schrodinger, LLC, New York, NY, USA, 2016) and by exploring the different number of principal components (PCs) through the partial least square (PLS) method. The best model, including two PCs, was statistically adequate:  $R^2 = 0.617$ , standard deviation (SD) of 0.6,  $Q^2 > 0.5$  ( $Q^2 = 0.529$ ), and the stability was close to 1 (stability = 0.985). This model was also predictive, where  $R^2_{\text{test}} > 0.6$  ( $R^2_{\text{test}} = 0.615$ ) and standard deviation of the test set predictions ( $SD_{\text{test}}$ ) was 0.532. The predictions of the pKi values for the 95 compounds from the training set and the 23 compounds from the test set were reported in Table 1. The values of the predicted pKi values were plotted against the experimental values (Figure 2) for the training set, and the cross-validated model was plotted over the training set and the test set. As can be seen in Figure 2, the plotted predictions were well-distributed across the

activity domain; although, the selected model seemed to have some problems in describing the SAR of the most active NAMdPs.



**Figure 1.** Structural alignment representation of the 118 LasB inhibitors; (left) backbone carbon atoms in gray and side chain carbon atoms in green, and (right) schematic structure of the NAMdPs.

Although within the limit, our model complied with QSAR statistics defined by Golbraikh and Tropsha ( $Q^2 > 0.5$ ,  $R^2_{\text{test}} > 0.6$ ) [22]. We also performed further tests on the external validation, according to the Roy and Roy criteria [23]. This test was based on the following criteria for a QSAR model to have predictive power—(i) at least one of the correlation coefficients for regressions through the origin (predicted versus observed activities, or observed versus predicted activities), specifically:  $[(R^2_{\text{test}} - R_0^2)/R^2_{\text{test}}]$  or  $[(R^2_{\text{test}} - R_0'^2)/R^2_{\text{test}}] < 0.1$ ; (ii) at least one slope ( $k$  or  $k'$ ) of the regression lines through the origin should be close to 1, i.e.,  $k$  or  $k'$  should satisfy:  $0.85 \leq k \leq 1.15$ , or  $0.85 \leq k' \leq 1.15$ ; (iii) a high value of  $R^2_m$  ( $R^2_m > 0.5$ ) was required, where  $R^2_m = R^2_{\text{test}} \times (1 - (R^2_{\text{test}} - R_0^2)^{1/2})$ . Our model complied with these criteria since  $[(R^2_{\text{test}} - R_0^2)/R^2_{\text{test}}] = 0.03$ ,  $k = 0.987$ , and  $R^2_m = 0.530$ .

Despite the calculated statistic being adequate, the  $R^2$  and  $Q^2$  values did not represent the high-fitted rates, which was commonly associated to a lower QSAR predictive ability. However, this asseveration is a point of discussion in the literature [22,24,25]. For instance, in the paper ‘3D-QSAR illusions’ Doweiko [25] considered that a higher  $Q^2$  reflects that the model identified the redundancy in the training set and this has nothing to do with its predictability. Under this criterion, a low  $Q^2$  reflects that each member of the training set is important for the model. In any case, our purpose in this work was the interpretation of the model to describe the differential activity in the dataset, following the idea of Doweiko, which suggested the use of 3D-QSAR models as a retrospective analytical tool, instead a predictive tool [25].

In the constructed 3D-QSAR model, the steric component had a 34.0% contribution, electrostatic had only 4.0%, hydrogen bond (HB)–donor had 13.4%, HB–acceptor had 21.9%, and hydrophobic had 26.7%. We utilized the contour isopleths projected in the most active NAMdP compound **103** ( $\text{HSCH}_2\text{CO-Trp-Tyr-NH}_2$ ) to mechanistically interpret the best 3D-QSAR model and to predict the most favorable AA residues in each position AA-1 and AA-2. The five 3D-QSAR field contour plots are shown in Figure 3, labeled as (A) steric, (B) electrostatic, (C) hydrophobic, (D) hydrogen bond (HB)–acceptor, and (E) HB–donor.

**Table 1.** List of NAMdPs used in this study, their experimental and predicted pK<sub>i</sub> values using the best QSAR model, their role in the QSAR construction (training or test set). Glide scores of their selected docking poses, and root mean square deviation (RMSD) values with respect to phosphoramidon in Protein Data Bank (PDB) structure 3DBK.

	NAMdP	Experimental pK <sub>i</sub>	Predicted pK <sub>i</sub>	QSAR Set	Glide Score (kcal/mol)	RMSD (Å)	%RefMatch	%MolMatch
1	HSCH <sub>2</sub> CO–Ala–Arg–NH <sub>2</sub>	−2.0607	−2.3215	Training	−7.652	0.95	38	67
2	HSCH <sub>2</sub> CO–Ala–Asp–NH <sub>2</sub>	−2.4997	−2.1689	Training	−5.239	0.95	38	78
3	HSCH <sub>2</sub> CO–Ala–Leu–NH <sub>2</sub>	−1.3222	−2.1287	Training	−5.067	0.96	38	78
4	HSCH <sub>2</sub> CO–Ala–Lys–NH <sub>2</sub>	−2.1847	−2.3482	Training	−5.975	0.94	38	74
5	HSCH <sub>2</sub> CO–Ala–Trp–NH <sub>2</sub>	−2.5966	−2.3854	Training	−5.161	3.09	62	96
6	HSCH <sub>2</sub> CO–Ala–Val–NH <sub>2</sub>	−1.7076	−2.1121	Test	−4.774	0.95	38	82
7	HSCH <sub>2</sub> CO–Arg–Asp–NH <sub>2</sub>	−2.8129	−2.3112	Test	−5.663	0.94	43	67
8	HSCH <sub>2</sub> CO–Arg–Lys–NH <sub>2</sub>	−2.1303	−2.4909	Training	−7.267	0.91	43	64
9	HSCH <sub>2</sub> CO–Arg–Phe–NH <sub>2</sub>	−2.3502	−2.2621	Training	−6.318	1.05	46	63
10	HSCH <sub>2</sub> CO–Arg–Trp–NH <sub>2</sub>	−2.0969	−2.5236	Training	−6.212	3.01	68	83
11	HSCH <sub>2</sub> CO–Asn–Arg–NH <sub>2</sub>	−2.4472	−2.4689	Training	−8.582	0.98	46	71
12	HSCH <sub>2</sub> CO–Asn–Leu–NH <sub>2</sub>	−2.7059	−2.2637	Training	−5.526	0.97	46	81
13	HSCH <sub>2</sub> CO–Asn–Lys–NH <sub>2</sub>	−2.4609	−2.4835	Test	−6.439	0.97	46	77
14	HSCH <sub>2</sub> CO–Asn–Phe–NH <sub>2</sub>	−1.5682	−2.2525	Training	−6.015	1.00	49	75
15	HSCH <sub>2</sub> CO–Asn–Trp–NH <sub>2</sub>	−1.8451	−2.5190	Training	−5.518	2.87	70	96
16	HSCH <sub>2</sub> CO–Asn–Val–NH <sub>2</sub>	−2.2553	−2.2475	Training	−5.682	0.99	46	85
17	HSCH <sub>2</sub> CO–Asp–Arg–NH <sub>2</sub>	−2.7412	−2.5774	Training	−8.351	0.98	46	71
18	HSCH <sub>2</sub> CO–Asp–Leu–NH <sub>2</sub>	−2.0719	−2.3721	Test	−5.413	0.95	46	81
19	HSCH <sub>2</sub> CO–Asp–Lys–NH <sub>2</sub>	−2.6911	−2.5914	Training	−5.397	0.96	46	77
20	HSCH <sub>2</sub> CO–Asp–Trp–NH <sub>2</sub>	−2.3160	−2.6275	Training	−4.564	2.86	70	96
21	HSCH <sub>2</sub> CO–Cys–Arg–NH <sub>2</sub>	−2.8102	−2.4454	Training	−5.857	1.18	41	68
22	HSCH <sub>2</sub> CO–Cys–Lys–NH <sub>2</sub>	−2.4378	−2.4601	Training	−3.691	1.18	41	75
23	HSCH <sub>2</sub> CO–Cys–Phe–NH <sub>2</sub>	−2.1173	−2.2302	Test	−3.155	0.97	43	73
24	HSCH <sub>2</sub> CO–Cys–Trp–NH <sub>2</sub>	−2.0334	−2.4966	Test	−3.091	3.05	65	96

Table 1. Cont.

	NAMdP	Experimental pK <sub>i</sub>	Predicted pK <sub>i</sub>	QSAR Set	Glide Score (kcal/mol)	RMSD (Å)	%RefMatch	%MolMatch
25	HSCH <sub>2</sub> CO–Cys–Val–NH <sub>2</sub>	–2.2068	–2.2235	Training	–2.802	0.95	41	83
26	HSCH <sub>2</sub> CO–Gln–Arg–NH <sub>2</sub>	–2.3365	–2.4675	Test	–8.794	0.97	43	64
27	HSCH <sub>2</sub> CO–Gln–Leu–NH <sub>2</sub>	–2.7324	–2.2627	Test	–6.350	0.95	43	73
28	HSCH <sub>2</sub> CO–Gln–Lys–NH <sub>2</sub>	–2.5798	–2.4815	Test	–6.717	0.94	43	70
29	HSCH <sub>2</sub> CO–Gln–Trp–NH <sub>2</sub>	–1.9590	–2.5192	Training	–6.093	2.93	68	89
30	HSCH <sub>2</sub> CO–Gln–Val–NH <sub>2</sub>	–2.9717	–2.2451	Training	–6.254	0.93	43	76
31	HSCH <sub>2</sub> CO–Glu–Lys–NH <sub>2</sub>	–2.7427	–2.6041	Training	–5.895	0.93	43	70
32	HSCH <sub>2</sub> CO–Glu–Phe–NH <sub>2</sub>	–2.1644	–2.3753	Training	–4.813	1.15	46	68
33	HSCH <sub>2</sub> CO–Glu–Trp–NH <sub>2</sub>	–2.8476	–2.6418	Training	–5.828	2.97	68	89
34	HSCH <sub>2</sub> CO–Glu–Val–NH <sub>2</sub>	–2.8633	–2.3677	Training	–6.019	0.96	43	76
35	HSCH <sub>2</sub> CO–Gly–Arg–NH <sub>2</sub>	–2.8069	–2.5394	Training	–8.393	0.94	35	65
36	HSCH <sub>2</sub> CO–Gly–Leu–NH <sub>2</sub>	–2.1399	–2.3336	Training	–5.225	0.94	35	76
37	HSCH <sub>2</sub> CO–Gly–Lys–NH <sub>2</sub>	–2.6542	–2.5543	Training	–6.274	0.94	35	72
38	HSCH <sub>2</sub> CO–Gly–Phe–NH <sub>2</sub>	–1.7076	–2.3247	Training	–5.604	0.97	38	70
39	HSCH <sub>2</sub> CO–Gly–Trp–NH <sub>2</sub>	–2.0864	–2.5911	Training	–5.584	3.16	59	96
40	HSCH <sub>2</sub> CO–Gly–Val–NH <sub>2</sub>	–2.6599	–2.3163	Training	–5.022	0.94	35	81
41	HSCH <sub>2</sub> CO–His–Ala–NH <sub>2</sub>	–0.5587	–1.6658	Training	–3.155	0.93	38	70
42	HSCH <sub>2</sub> CO–His–Leu–NH <sub>2</sub>	–2.4857	–1.68	Training	–3.252	0.92	38	61
43	HSCH <sub>2</sub> CO–His–Lys–NH <sub>2</sub>	–2.5211	–1.8976	Training	–3.563	0.96	38	58
44	HSCH <sub>2</sub> CO–His–Phe–NH <sub>2</sub>	–1.3222	–1.6688	Training	–3.128	0.95	41	58
45	HSCH <sub>2</sub> CO–His–Trp–NH <sub>2</sub>	–1.2553	–1.9353	Training	–3.052	3.10	62	79
46	HSCH <sub>2</sub> CO–His–Val–NH <sub>2</sub>	–1.6721	–1.6632	Training	–2.932	0.94	38	64
47	HSCH <sub>2</sub> CO–Ile–Asp–NH <sub>2</sub>	–2.1523	–1.5871	Training	–4.167	1.00	43	76
48	HSCH <sub>2</sub> CO–Ile–Gln–NH <sub>2</sub>	–0.5717	–0.5951	Training	–6.036	0.92	43	73
49	HSCH <sub>2</sub> CO–Ile–Leu–NH <sub>2</sub>	–0.1004	–1.5472	Training	–4.684	0.94	43	76
50	HSCH <sub>2</sub> CO–Ile–Lys–NH <sub>2</sub>	–2.2788	–1.7662	Training	–5.633	0.95	43	73

Table 1. Cont.

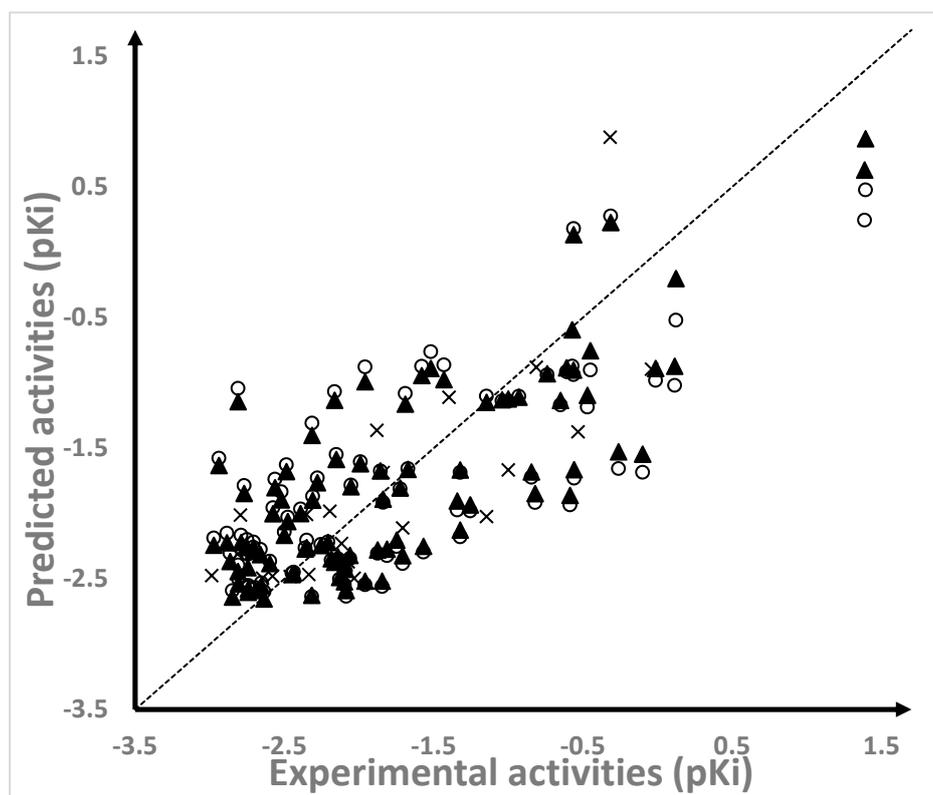
	NAMdP	Experimental pK <sub>i</sub>	Predicted pK <sub>i</sub>	QSAR Set	Glide Score (kcal/mol)	RMSD (Å)	%RefMatch	%MolMatch
51	HSCH <sub>2</sub> CO-Ile-Thr-NH <sub>2</sub>	-0.5328	-1.3756	Test	-6.040	0.93	43	80
52	HSCH <sub>2</sub> CO-Ile-Trp-NH <sub>2</sub>	-2.5635	-1.8041	Training	-5.822	2.92	68	93
53	HSCH <sub>2</sub> CO-Ile-Tyr-NH <sub>2</sub>	-0.3139	0.226	Training	-5.897	0.94	46	68
54	HSCH <sub>2</sub> CO-Ile-Val-NH <sub>2</sub>	-0.2625	-1.5298	Training	-4.851	0.93	43	80
55	HSCH <sub>2</sub> CO-Leu-Arg-NH <sub>2</sub>	-2.7945	-2.013	Test	-7.750	1.16	46	71
56	HSCH <sub>2</sub> CO-Leu-Asp-NH <sub>2</sub>	-2.7686	-1.8479	Training	-4.075	1.10	46	81
57	HSCH <sub>2</sub> CO-Leu-Leu-NH <sub>2</sub>	-1.7243	-1.8061	Training	-5.417	1.23	46	81
58	HSCH <sub>2</sub> CO-Leu-Lys-NH <sub>2</sub>	-1.1461	-2.0253	Test	-5.291	1.10	46	77
59	HSCH <sub>2</sub> CO-Leu-Phe-NH <sub>2</sub>	-2.0531	-1.7966	Training	-3.977	1.26	49	75
60	HSCH <sub>2</sub> CO-Leu-Trp-NH <sub>2</sub>	-2.4771	-2.0634	Training	-4.790	2.98	70	96
61	HSCH <sub>2</sub> CO-Lys-Asp-NH <sub>2</sub>	-2.9872	-2.4763	Test	-6.095	0.97	43	73
62	HSCH <sub>2</sub> CO-Lys-Leu-NH <sub>2</sub>	-2.0899	-2.4284	Training	-6.349	0.96	43	73
63	HSCH <sub>2</sub> CO-Lys-Lys-NH <sub>2</sub>	-2.6365	-2.6554	Training	-7.149	0.96	43	70
64	HSCH <sub>2</sub> CO-Lys-Phe-NH <sub>2</sub>	-2.1004	-2.4267	Training	-6.502	0.97	46	68
65	HSCH <sub>2</sub> CO-Lys-Val-NH <sub>2</sub>	-2.7443	-2.4166	Training	-6.072	0.96	43	76
66	HSCH <sub>2</sub> CO-Met-Arg-NH <sub>2</sub>	-0.8195	-1.8487	Training	-7.798	1.00	43	67
67	HSCH <sub>2</sub> CO-Met-Asp-NH <sub>2</sub>	-0.8451	-1.6836	Training	-4.891	0.99	43	76
68	HSCH <sub>2</sub> CO-Met-Lys-NH <sub>2</sub>	-0.5866	-1.8627	Training	-6.684	0.97	43	73
69	HSCH <sub>2</sub> CO-Met-Phe-NH <sub>2</sub>	-2.9380	-1.6339	Training	-5.564	1.19	46	71
70	HSCH <sub>2</sub> CO-Met-Trp-NH <sub>2</sub>	-2.3096	-1.9004	Training	-4.916	2.89	68	93
71	HSCH <sub>2</sub> CO-Met-Tyr-NH <sub>2</sub>	-0.5623	0.1316	Training	-5.955	1.26	46	68
72	HSCH <sub>2</sub> CO-Met-Val-NH <sub>2</sub>	-1.9912	-1.6213	Training	-5.815	0.97	43	80
73	HSCH <sub>2</sub> CO-Phe-Gln-NH <sub>2</sub>	0.1226	-0.2028	Training	-6.245	0.96	38	56
74	HSCH <sub>2</sub> CO-Phe-Ile-NH <sub>2</sub>	-0.6503	-1.1362	Training	-5.120	0.93	38	58
75	HSCH <sub>2</sub> CO-Phe-Leu-NH <sub>2</sub>	-2.8096	-1.146	Training	-4.450	0.95	38	58
76	HSCH <sub>2</sub> CO-Phe-Lys-NH <sub>2</sub>	-1.8808	-1.3642	Test	-5.413	0.95	38	56

Table 1. Cont.

	NAMdP	Experimental pK <sub>i</sub>	Predicted pK <sub>i</sub>	QSAR Set	Glide Score (kcal/mol)	RMSD (Å)	%RefMatch	%MolMatch
77	HSCH <sub>2</sub> CO-Phe-Met-NH <sub>2</sub>	-0.4502	-0.7559	Training	-6.914	0.91	38	58
78	HSCH <sub>2</sub> CO-Phe-Phe-NH <sub>2</sub>	-2.1644	-1.1354	Training	-5.606	1.23	41	56
79	HSCH <sub>2</sub> CO-Phe-Trp-NH <sub>2</sub>	-2.3139	-1.4029	Training	-7.568	1.76	62	77
80	HSCH <sub>2</sub> CO-Phe-Tyr-NH <sub>2</sub>	1.3872	0.6276	Training	-6.918	0.95	41	54
81	HSCH <sub>2</sub> CO-Phe-Val-NH <sub>2</sub>	-1.0414	-1.1315	Training	-4.537	0.96	38	61
82	HSCH <sub>2</sub> CO-Pro-Arg-NH <sub>2</sub>	-1.7482	-2.2055	Training	-7.338	1.53	38	73
83	HSCH <sub>2</sub> CO-Pro-Leu-NH <sub>2</sub>	-2.3909	-2.0021	Training	-7.612	1.40	38	85
84	HSCH <sub>2</sub> CO-Pro-Lys-NH <sub>2</sub>	-2.8842	-2.2260	Training	-8.277	1.75	38	81
85	HSCH <sub>2</sub> CO-Pro-Trp-NH <sub>2</sub>	-2.7497	-2.2623	Training	-5.630	2.96	62	96
86	HSCH <sub>2</sub> CO-Pro-Val-NH <sub>2</sub>	-2.1959	-1.9837	Test	-7.628	1.48	38	92
87	HSCH <sub>2</sub> CO-Ser-Arg-NH <sub>2</sub>	-2.6474	-2.4949	Test	-8.952	0.94	41	68
88	HSCH <sub>2</sub> CO-Ser-Leu-NH <sub>2</sub>	-2.7076	-2.2896	Training	-6.262	1.14	41	79
89	HSCH <sub>2</sub> CO-Ser-Phe-NH <sub>2</sub>	-1.8751	-2.2801	Training	-6.316	1.15	43	73
90	HSCH <sub>2</sub> CO-Ser-Val-NH <sub>2</sub>	-2.3598	-2.2737	Training	-5.362	1.12	41	83
91	HSCH <sub>2</sub> CO-Thr-Arg-NH <sub>2</sub>	-2.7896	-2.2227	Training	-9.096	1.07	41	65
92	HSCH <sub>2</sub> CO-Thr-Phe-NH <sub>2</sub>	-2.3522	-2.0078	Test	-5.473	1.13	43	70
93	HSCH <sub>2</sub> CO-Thr-Trp-NH <sub>2</sub>	-1.8129	-2.2743	Training	-6.166	2.99	65	92
94	HSCH <sub>2</sub> CO-Thr-Val-NH <sub>2</sub>	-2.5763	-2.0035	Training	-5.808	1.06	41	79
95	HSCH <sub>2</sub> CO-Trp-Arg-NH <sub>2</sub>	-1.3979	-1.1114	Test	-8.417	0.94	38	47
96	HSCH <sub>2</sub> CO-Trp-Asp-NH <sub>2</sub>	-1.5798	-0.9463	Training	-6.044	0.96	38	52
97	HSCH <sub>2</sub> CO-Trp-Glu-NH <sub>2</sub>	-1.9590	-0.9930	Training	-6.483	0.99	38	50
98	HSCH <sub>2</sub> CO-Trp-Ile-NH <sub>2</sub>	-0.0128	-0.8916	Training	-7.636	0.51	38	52
99	HSCH <sub>2</sub> CO-Trp-Leu-NH <sub>2</sub>	-0.5635	-0.9039	Training	-5.190	0.95	38	52
100	HSCH <sub>2</sub> CO-Trp-Lys-NH <sub>2</sub>	-1.0000	-1.1254	Training	-8.613	0.66	38	50
101	HSCH <sub>2</sub> CO-Trp-Phe-NH <sub>2</sub>	-0.0414	-0.8966	Test	-6.127	1.04	41	50
102	HSCH <sub>2</sub> CO-Trp-Trp-NH <sub>2</sub>	-1.6902	-1.1641	Training	-5.961	3.04	62	70
103	HSCH <sub>2</sub> CO-Trp-Tyr-NH <sub>2</sub>	1.3925	0.8661	Training	-5.518	0.98	41	48

Table 1. Cont.

	NAMdP	Experimental pK <sub>i</sub>	Predicted pK <sub>i</sub>	QSAR Set	Glide Score (kcal/mol)	RMSD (Å)	%RefMatch	%MolMatch
104	HSCH <sub>2</sub> CO-Trp-Val-NH <sub>2</sub>	-0.6096	-0.8870	Training	-5.893	0.98	38	54
105	HSCH <sub>2</sub> CO-Tyr-Arg-NH <sub>2</sub>	-0.4698	-1.0968	Training	-10.083	0.59	38	50
106	HSCH <sub>2</sub> CO-Tyr-Asp-NH <sub>2</sub>	-0.7404	-0.9317	Training	-6.700	0.43	38	56
107	HSCH <sub>2</sub> CO-Tyr-Glu-NH <sub>2</sub>	-1.4314	-0.9771	Training	-7.712	0.47	38	54
108	HSCH <sub>2</sub> CO-Tyr-Leu-NH <sub>2</sub>	-1.5185	-0.8919	Training	-7.353	0.43	38	56
109	HSCH <sub>2</sub> CO-Tyr-Lys-NH <sub>2</sub>	-0.9294	-1.1111	Training	-8.097	0.52	38	54
110	HSCH <sub>2</sub> CO-Tyr-Phe-NH <sub>2</sub>	-0.8129	-0.8823	Test	-7.380	0.50	41	54
111	HSCH <sub>2</sub> CO-Tyr-Trp-NH <sub>2</sub>	-1.1461	-1.1498	Training	-7.392	2.81	62	74
112	HSCH <sub>2</sub> CO-Tyr-Tyr-NH <sub>2</sub>	-0.3181	0.8790	Test	-7.686	0.87	41	52
113	HSCH <sub>2</sub> CO-Tyr-Val-NH <sub>2</sub>	0.1146	-0.8746	Training	-7.274	0.46	38	58
114	HSCH <sub>2</sub> CO-Val-Arg-NH <sub>2</sub>	-1.8388	-1.8918	Training	-8.183	1.00	41	65
115	HSCH <sub>2</sub> CO-Val-Leu-NH <sub>2</sub>	-1.8388	-1.6874	Test	-5.767	0.94	41	75
116	HSCH <sub>2</sub> CO-Val-Lys-NH <sub>2</sub>	-1.3424	-1.9057	Training	-5.501	0.95	41	71
117	HSCH <sub>2</sub> CO-Val-Phe-NH <sub>2</sub>	-1.8573	-1.6755	Training	-5.001	0.95	43	70
118	HSCH <sub>2</sub> CO-Val-Val-NH <sub>2</sub>	-1.0000	-1.6676	Test	-5.225	0.95	41	79



**Figure 2.** Scatter plots of the predicted versus experimental pKi values for the training set (▲), the cross-validated models over the training set (○), and the test set (×).

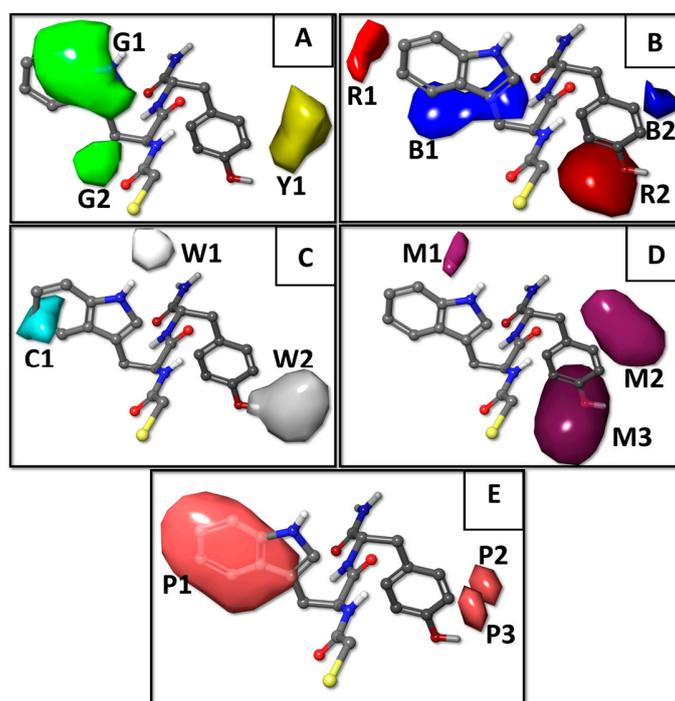
The green (G1 and G2) and yellow (Y1) isopleths represent favorable and unfavorable components of the steric field (Figure 3A). Bulky groups were tolerated as AA-1 residue (green isopleths G1 and G2) to increase the inhibitory activity. Thus, NAMdPs with Trp, Phe, His, and Tyr in the side chain of AA-1 (isopleth G1) were among the most potent inhibitors in the dataset (considering  $pK_i > -1.5$  as a threshold, a deeper analysis revealed that 8 of 10 compounds with Trp as AA-1, 5 of 9 compounds with Phe as AA-1, 3 of 6 compounds with His as AA-1, and 8 of 9 compounds with Tyr as AA-1, were among the most potent LasB inhibitors). Additionally, the presence of alkyl substituents at the C $\beta$  of AA-1 (isopleth G2), also increased the LasB inhibitory activity. In fact, the majority of NAMdPs increase their activities when Ala in AA-1 was replaced by Ile or Val residues. For instance, the pKi values increased when compounds **2**, **3**, and **6** (with AA-1 = Ala) were compared with **47**, **49**, and **54** (with AA-1 = Ile), respectively, and the same happened when compounds **1**, **4**, and **6** (with AA-1 = Ala) were compared with **114**, **116**, and **118** (with AA-1 = Val), respectively. Bulky substituents at the position AA-2 (yellow isopleth Y1) considerably decreased the biological activity. Thus, NAMdPs with Phe and Trp in the side chain of AA-2 (isopleth Y1) were among the less potent inhibitors in the dataset (considering  $pK_i < -1.5$  to be a threshold, a deeper analysis revealed that 11 of 15 compounds with Phe as AA-2 and 15 of 17 compounds with Trp as AA-2, were among the less potent LasB inhibitors).

Blue (B1 and B2) and red (R1 and R2) isopleths represent the favorable and unfavorable components of the electrostatic field. Blue isopleths were in the regions where the positive charges were favorable (or negative charges are unfavorable) for activity, and the red isopleths were in regions where more negative charges were favorable (or positive charges were unfavorable) for activity. Positively charged residues with a long side chain (e.g., Arg, Lys) as AA-1 and AA-2, had negative effects in the potency of the LasB inhibition (red isopleths R1 and R2 in Figure 3B). All compounds with Arg and Lys in AA-1 (isopleth R1) had  $pK_i < -2.0$  (Table 1). Meanwhile, NAMdPs with Arg and Lys in AA-2 (isopleth R2) were among the less potent inhibitors in the dataset (considering  $pK_i < -1.5$  as a threshold, a

deeper analysis revealed that 11 of the 14 compounds with Arg as AA-2 and 13 of 18 compounds with Lys as AA-2, were among the less potent LasB inhibitors). On the other hand, negatively charged residues (e.g., Asp, Glu) as AA-1 and AA-2, had negative effects on the potency of LasB inhibition (blue isopleths B1 and B2 in Figure 3B). All compounds with Asp and Glu in AA-1 (isopleth B1) had  $pK_i < -2.0$  (Table 1). Meanwhile, NAMdPs with Asp in AA-2 (isopleth B2) were among the less potent inhibitors in the dataset (considering  $pK_i < -1.5$  as a threshold, a deeper analysis revealed that 6 of 8 compounds with Asp as AA-2 were among the less potent LasB inhibitors).

White (W1 and W2) and cyan (C1) isopleths represent the favorable and unfavorable components of the HB-donor field (Figure 3C). The white isopleth W1 indicates that Trp and His residues in AA-1 favor the inhibitory activity (considering  $pK_i > -1.5$  as the threshold, it is possible to see that 3 of 6 compounds with His as AA-1 and 7 of 10 compounds with Trp as AA-1, were among the more potent LasB inhibitors). Meanwhile, isopleth W2 indicated that Tyr and Gln residues in AA-2 were essential for increasing the inhibitory activity (all compounds with Tyr and Gln in AA-2 had  $pK_i > -0.6$ ). On the other hand, the cyan isopleth C1 indicated that the HB-donor groups in this region (provided by residues Arg, Lys, and Gln in AA-1) were unfavorable to perform a good inhibitory activity.

Maroon isopleths (M1, M2, and M3) represent the favorable components of the HB-acceptor field (Figure 3D). Isopleths M1 and M3 indicate that Tyr residue in AA-1 and AA-2, respectively, favor the LasB inhibitory activity (it is possible to see in Table 1 that the vast majority of compounds that contain tyrosine are potent inhibitors). On the other hand, isopleth M2 indicate that Gln residue in AA-2 favors the inhibitory activity (all compounds with Gln in AA-2 had  $pK_i > -0.6$ ).



**Figure 3.** 3D-QSAR model contour maps obtained with the five force fields and represented around the most active inhibitor in the dataset (compound 103,  $\text{HSCH}_2\text{CO-Trp-Tyr-NH}_2$ ). (A) Steric—the favorable and unfavorable components are in green (G1 and G2) and in yellow (Y1), respectively. (B) Electrostatic—the favorable and unfavorable components are in blue (B1 and B2) and in red (R1 and R2), respectively. (C) Hydrogen bond (HB)-donor—the favorable and unfavorable components are in white (W1 and W2) and in cyan (C1), respectively, and (D) HB-acceptor—the favorable components are in maroon (M1, M2, and M3). (E) Hydrophobic—the favorable components are in pink (P1, P2, and P3).

Pink isopleths (P1, P2, and P3) represent the favorable components of the hydrophobic field (Figure 3E). These isopleths reflect that the presence of hydrophobic amino acids in AA-1 and AA-2 positions increases the LasB inhibition potency of NAMdPs. It is possible to see in Figure 3E that P1 encompasses the hydrophobic part of Trp in AA-1, and the side chain groups of Phe, Ile, Val, and Leu were also in this region. As mentioned above, there were instances of potent NAMdPs with these residues in AA-1. On the other hand, isopleths P2 and P3 indicate that the hydrophobic residues could increase the LasB inhibition potency. In fact, compounds with the residues Ile and Met had  $pK_i > -0.7$ .

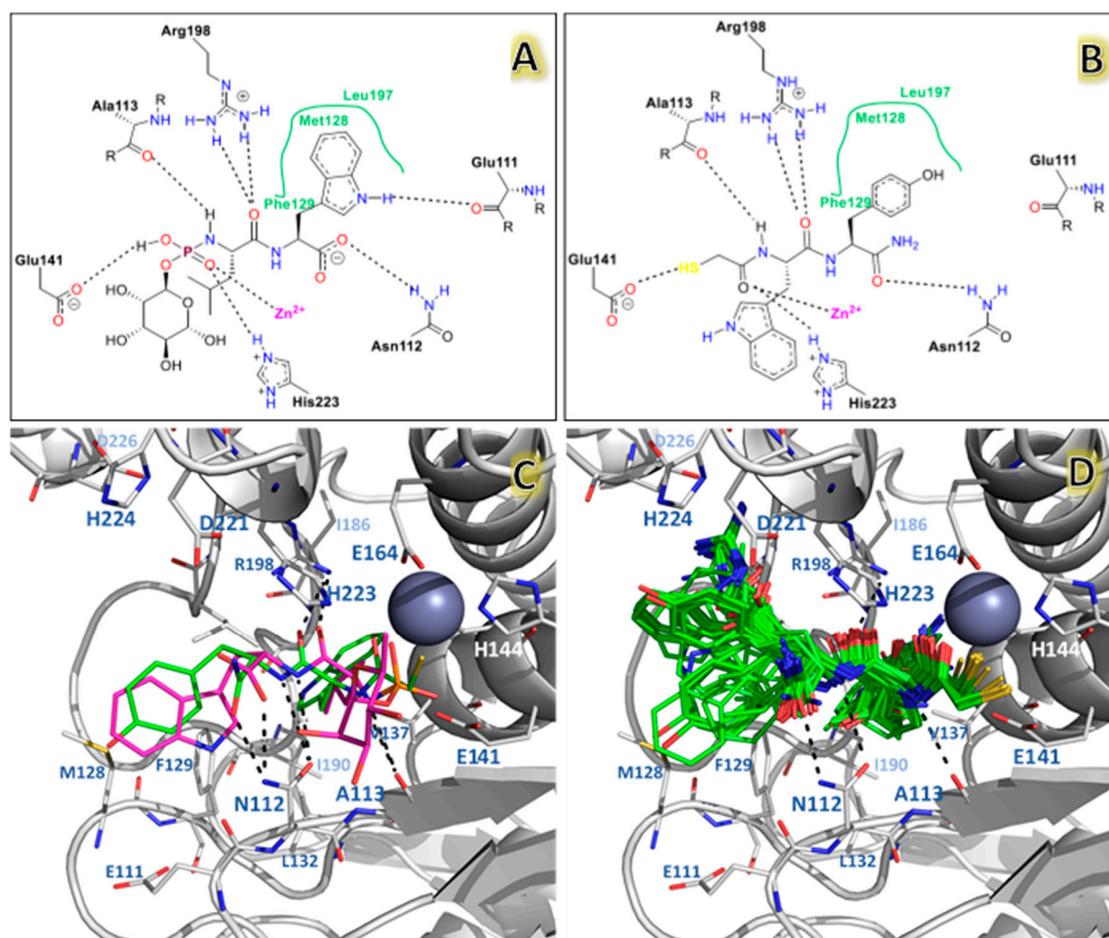
Thus, the 3D-QSAR let us conclude that the rational design of novel potential inhibitors should be directed to compounds that might have at least an aromatic and bulky group at the AA-1 position and a middle size with HB interactions at the AA-2 position. However, it is well-known that the 3D-QSAR models had limitations, since they did not consider protein–ligand interactions. To complement the 3D-QSAR results, *in silico* molecular docking experiments of all 118 NAMdPs were performed.

## 2.2. Molecular Docking Results

The docking method allows to create a protein–ligand interaction model for LasB inhibitors. The docking Glide scores of the selected poses per 118 ligands are reported in Table 1. These poses were first compared to the phosphoramidon inhibitor that was taken as a reference (Protein Data Bank (PDB) code 3DBK), since it was the only crystallized compound with a structure similar to that of NAMdPs. To get a better insight on the chemical environment surrounding the ligands, ligand interaction diagrams were sketched for phosphoramidon in the 3DBK crystal and the docking pose of the most active compound **103** (Figure 4).

The phosphate group of phosphoramidon in the crystallographic structure shows electrostatic interactions with the Zinc ion and the residue His223 of the LasB active site. The mercaptoacetyl group in all docked NAMdP poses was oriented to the same ion and histidine, in agreement with the previous report of Cathcart et al. [17]. These interactions keep the ligands fixed, allowing for better orientations of the whole structures to occupy the complete binding site. The backbone NH and CO groups of the AA-1 residue (Leu) in phosphoramidon and the AA-1 residues in 100% of the docked structures formed HB interactions with the residues A113 and R198. On the other hand, the backbone groups of the AA-2 residue (Trp) in phosphoramidon and the AA-2 residues of the docked structures formed HB interactions with the residues N112.

The docking poses of the entire set of 118 NAMdPs were compared with the conformation of phosphoramidon in the X-ray crystallographic structure (PDB: 3DBK). Figure 4c,d show that the docked structures fitted in an acceptable way with phosphoramidon. For a better insight, the above mentioned comparison was performed with an in-house script (named ligRMSD) [26], which identified the common graphs between molecules and calculated the root mean square deviation (RMSD) between the equivalent atoms of the common graphs. Since the NAMdPs were different from the reference compound (phosphoramidon), RMSD values were calculated by considering only the common graphs between molecules. In this context, %RefMatch and %MolMatch values were defined, where %RefMatch referred to the percent of common graphs between the docked compound and phosphoramidon, with respect to the total number of atoms of phosphoramidon; whereas, %MolMatch refers to the percent of common graphs between the docked compound, and phosphoramidon, with respect to the total number of atoms of the docked compound. These values allowed for identifying the maximal similitude between the docked compound and phosphoramidon; therefore, an RMSD value with high %RefMatch and %MolMatch values reflected that the compound under analysis bore a major resemblance with phosphoramidon.



**Figure 4.** The docking results for NAMdPs and comparison with phosphoramidon in X-ray crystallographic structure 3DBK. Diagram for phosphoramidon (A), and the most active compound **103** (B) inside LasB binding site. (C) The docking pose obtained for compound **103** (green) and comparison with the crystallographic structures of phosphoramidon (purple). (D) The binding modes of the 118 compounds.

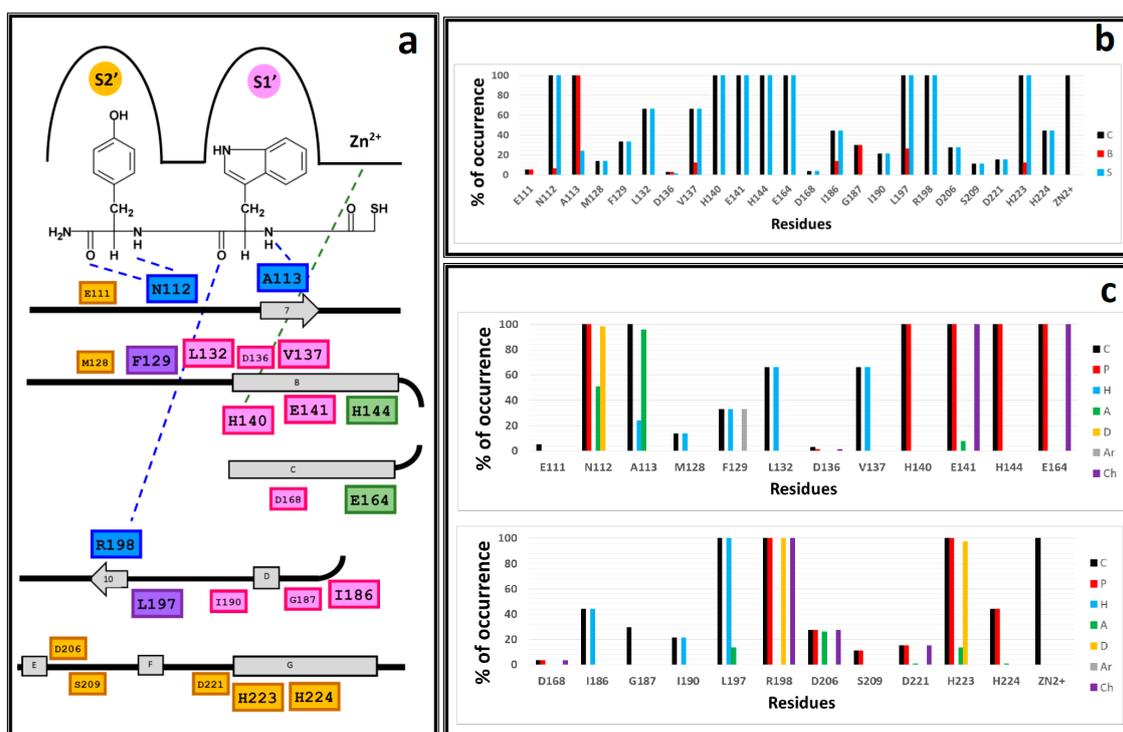
The calculated RMSD values, reported in Table 1, were in the range of 0.4–3.1 Å. It is accepted in literature that  $\text{RMSD} = 2.0 \text{ \AA}$  could be considered to be the threshold value that discriminates between the right and wrong docking solutions for identical compounds [27,28] (this threshold could be higher for non-identical compounds).

Among the structures with lower RMSD values, a match that included just the peptide backbone was found, which is why the match percentage exhibited small values for that set. On the contrary, the NAMdPs with higher %MolMatch values (96%) included the entire set of  $\text{HSCH}_2\text{CO}-(\text{AA}-1)\text{-Trp-NH}_2$  inhibitors, of which  $\text{HSCH}_2\text{CO-Leu-Trp-NH}_2$ ,  $\text{HSCH}_2\text{CO-Asp-Trp-NH}_2$ , and  $\text{HSCH}_2\text{CO-Asn-Trp-NH}_2$  showed better match values but were associated with high RMSD values (2.86–2.98 Å). These values reflected that the backbone atoms of the NAMdPs matched with the ones of phosphoramidon, but the side chain groups were differently oriented (Figure 4c,d). The NAMdP residues AA-2 contributed the most to increase the RMSD values. For instance, the side chain groups of the tryptophan residues for these inhibitors were rotated by some degrees with respect to this residue in phosphoramidon, due to the wide cavity of the binding pocket S2'. This higher flexibility led to higher RMSD values. The visual analysis of the structures (Figure 4d) indicated a similar orientation with respect to the reference phosphoramidon, which was reflected in the match between the backbone groups of both structures.

A systematic and detailed analysis of all possible interactions between LasB and the docked NAMdP could be performed by using Interaction Fingerprints (IFPs). Previous studies have demonstrated that IFP analysis is a valuable tool that allowed for a better schematization of protein–ligand interactions [26,29,30].

For a better comprehension of the interactions between the docked ligands and LasB, IFP analysis was performed. This analysis was a robust way of understanding the possible interactions between the receptor (enzyme) and the docked ligands. Then, the identification of those chemical interactions and the residues involved in them, could lead to a more detailed description of the LasB binding site available to the series of 118 NAMdP conformations, according to the docking results. This information could be considered for the design of novel potent inhibitors.

Plots of the chemical interactions types occurrence per residue are depicted in Figure 5. Residues with interactions and their position in the LasB sequence are depicted in Figure 5a. IFP analysis applied to the 118 LasB–NAMdP complexes is shown in Figure 5b,c. At first sight, there is a clear distinction between the residues belonging to both binding pockets previously described by Cathcart [17], called as S1' (where AA-1 was placed) and S2' (where AA-2 was placed). Topologically, the S1' pocket was more internal and the S2' pocket was formed by more superficial residues. Percent of occurrence plots showed that 23 residues of LasB were involved in the binding of NAMdPs.



**Figure 5.** Occurrence of interaction types at the LasB–NAMdP binding interfaces. (A) Residues with observed interactions, their position in the LasB sequence, and their position in the binding site (with compound 103 as reference)—residues in pocket S1' are in pink, residues in pocket S2' are in orange, residues in the interface between the two pockets are in violet, residues coordinated to Zn<sup>2+</sup> are in green, residues hydrogen-bonded to the backbone residues of the ligands are in blue. (B) Percentages of occurrence of contacts C, interactions with the backbone of the residue B, and interactions with the side chain of the residue S. (C) Percentages of occurrence of chemical interactions—contacts C, polar P, hydrophobic H, and HBs where the residue is acceptor A, HBs where the residue is donor D, aromatic Ar, and the electrostatic charged groups, Ch. The LasB–NAMdP structures obtained by docking were used for calculations of the percentages of occurrence represented here.

IFPs reflected that 100% of the docked structures formed the above mentioned HB interactions with the residues A113 and R198 that anchored the backbone of the AA-1 ligand residues to the pocket S1' and the HB interactions with the residue N112 that anchored the backbone of the AA-2 ligand residues to the pocket S2'. In this context, the plots in Figure 5c showed that A113 acted as the HB acceptor in more than 95% of the total structures, R198 acted as the HB donor and formed electrostatic interactions in 100% of the total structures, and N112 acted as the HB acceptor and HB donor in more than 50% and 98%, respectively, of the total structures. IFPs also showed that the residues H144 and E164, which were coordinated to  $Zn^{2+}$  and are close to the thiol groups of the docked NAMdPs, had polar contributions in 100% of the docked structures. E164 also had electrostatic contributions in 100% of them.

IFPs also reflected that the majority of the residues inside the pocket S1' were mainly hydrophobic (Figure 5); in fact, the residues L132, V137, I186, I190, and L197 provide hydrophobic contributions to this pocket. L197 (also with some contacts in pocket S2') showed these contributions in 100% of the docked poses. Additionally, L132 and V137 showed these contributions in more than 65% of the structures, and I186 and I190 showed them in more than 40% and 20% of the structures, respectively. It was noteworthy that I186 and I190 showed interactions with compounds that contained bulky AA-1 ligand residues such as Phe and Trp, and these residues were present in the AA-1 group of several of the most active NAMdPs. The residue E141, also in pocket S1' and close to the mercaptoacetyl group, showed polar and electrostatic contributions in 100% of the docked structures, and the residue H140, which was also in pocket S1' and was coordinated to  $Zn^{2+}$ , showed polar contributions in 100% of the docked structures. Finally, the residue G187, also in pocket S1', had contacts in around 30% of the docked structures.

Unlike pocket S1', IFPs reflected that the pocket S2' was more hydrophilic (Figure 5). The residues D206, S209, D221, H223, and H224 provided polar contributions to this pocket. The residue H223 had polar contributions in 100% of the docked poses. It also acted as the HB donor and the HB acceptor in more than 97% and 13% of them, respectively. The residue D206 had polar contributions and acted as the HB acceptor in more than 25% of the docked structures. The residues S209, D221, and H224 had polar contributions in 11.0%, 15.2%, and 44.0% of the docked structures. On the other hand, the residues M128 and F129, also in pocket S2', provided hydrophobic contributions in 13.5% and 33.0% of the docked structures, respectively. F129, which also had some contacts in pocket S1', showed aromatic contributions.

### 3. Materials and Methods

#### 3.1. Dataset Collection and Pre-Processing

A dataset of 118 LasB inhibitors was collected from the two series reported by Cathcart et al. in publications [17,20]. Compounds with unknown activity values were excluded, and those with values reported in both papers kept the activity value with more precision (in all cases the  $pK_i$  values were virtually identical).  $K_i$  values were in the range of  $4.05 \times 10^{-2} \mu\text{M}$ –971  $\mu\text{M}$ , and they were scaled using logarithmic scale as  $\log(1/K_i)$  (as  $pK_i$  units). The structures were pre-processed in Maestro using LigPrep (Maestro, Schrodinger LLC, New York, NY, USA, 2016), and the protonation states of titratable groups were predicted using Epik [31,32], at a pH value of 7.2. As the last preparation step, the molecular representations were visually inspected. The naming scheme of the ligands in terms of their two AA residues (i.e., AA-1 and AA-2 components) is represented in Table 1. The names are related to the position as well, such as a ligand with arginine as AA-1 and aspartic acid as AA-2 are represented as "HSC<sub>2</sub>CO–Arg–Asp–NH<sub>2</sub>".

#### 3.2. QSAR Methodology

The 118 structures in the data set were aligned by hand in Maestro's Molecular Editor (Maestro, Schrodinger LLC, New York, NY, USA). Moreover, the field-based 3D-QSAR models were trained in

Phase (Phase, Schrodinger LLC, New York, NY, USA, 2016) by the random splitting implemented in this tool, getting a relation training/test sets of approximately 80/20 (95 and 23 compounds in training and test sets, respectively). The training process of 3D-QSAR models was carried out over the available descriptors, using the OPLS\_2005 force field [33,34]. The fields were calculated on an orthohedral grid that enclosed the training set molecules, with a spacing of 1 Å and extended 3 Å beyond the limits of this set. The threshold for van der Waals and electrostatic interactions was set to 30 Kcal/mol, removing points closer than 2 Å to any of the training set atoms. During the PLS procedure, all variables (points in the grid) with a standard deviation of less than 0.01 were removed. Additionally, the variables whose regression coefficients were overly sensitive to small changes in the training set composition were removed using a  $|t\text{-value}| < 2.00$  filter, as implemented in the Phase interface. Finally, the maximum number of PLS factors was set to 10.

### 3.3. Molecular Docking

Molecular docking calculations were performed using Glide [35] from the Schrodinger Suite. The coordinates of LasB were extracted from the PDB crystal with code 3DBK. This crystal is a complex between LasB and the ligand phosphoramidon. This ligand was similar in dimensions with the studied ligands; which is why it was used as a reference for the box generation. The downloaded crystal was pre-processed using the Protein Preparation Wizard (Protein Preparation Wizard, Schrodinger LLC, New York, NY, USA, 2016) protocol. The set of previously pre-processed ligands were docked in a 30 Å × 30 Å × 30 Å box, centered on the center of mass of phosphoramidon, covering the entire active site of the receptor. Standard (SP) and extra-precision (XP) modes were run in Glide [35], but only the XP mode was used to find adequate poses for all ligands [36]. The Glide protocol and parameters were the same as reported in previous reports [37–39]. The selection of poses was based on looking for the observed protein–ligand interactions patterns in the reported PDB crystals of LasB, and in the selection of the lowest scoring energy from among the adequate poses. Protein–ligand interaction patterns were identified and defined as essential chemical interactions described for analogue ligands (ECIDALs) [27,40] with LasB inhibitory activities. Finally, one pose per ligand was chosen.

### 3.4. IFP Calculations

IFPs defined in Singh reports [41,42] were calculated in Maestro (Maestro, Schrodinger LLC, New York, NY, USA, 2016) over the poses of ligands selected in docking experiments. The method identifies the presence of different chemotypes, such as polar (P), hydrophobic (H), HBs with an acceptor (A) as a residue group, HBs with a donor (D) as a residue group, aromatic (Ar), and electrostatic interactions with charged groups (Ch). These chemotypes were accounted as interactions between the ligands and the binding site residues of the target receptor. An interaction (under the chemotype definition) was accounted when a residue contained atoms within a specified cut-off distance from the ligand atoms.

## 4. Conclusions

Summing up, a set of 118 NAMdP inhibitors was studied by using 3D-QSAR modeling and molecular docking. QSAR analysis reflected that the side chain of the residue AA-1, at the NAMdP skeleton, should be mainly hydrophobic with bulky aromatic substituents (e.g., Phe, Trp, and Tyr). Furthermore, polar residues with large side chains (Lys, Arg, Glu, Gln, Asp) were not tolerated as AA-1. Meanwhile, non-bulky aromatic groups with functional groups were able to act, as HB donors are preferred as AA-2 (e.g., Gln and Tyr). The best compound in the studied set (compound 103) had these preferred structural requirements to assure a good LasB inhibition—AA-1 was Trp, a bulky hydrophobic residue, and AA-2 was Tyr, which contained the hydroxyl group as an HB donor.

In addition, this work studied in detail the ligand–enzyme interactions of the whole set of compounds, and compared them with the ligand phosphoramidon (which is a crystal forming a complex with LasB). NAMdPs, according to the docking experiments, were oriented inside the LasB binding site by forming interactions with the Zn<sup>2+</sup> ion, pocket S1', and pocket S2', as expected. The

poses obtained by docking were used to carry out an IFP analysis, leading to a complete map of the LasB residues that interacted with NAMdPs.

The information provided here, through a combination of 3D-QSAR and docking experiments, might be taken into account by medicinal chemists interested in the synthesis and design of antimicrobials, specifically LasB inhibitors, with the goal of improving future rational drug design of specific potent therapeutics.

**Author Contributions:** Conceptualization, A.F.d.l.T., J.L.V.-L., and J.C.; Methodology, J.L.V.-L.; Software, J.L.V.-L.; Formal Analysis, J.C., J.L.V.-L., and J.A.M.-L.; and A.F.d.l.T.; Resources, J.C.; Writing—Original Draft Preparation, A.F.d.l.T., J.A.M.-L., and J.C.; Writing—Review & Editing, A.F.d.l.T., J.A.M.-L., and J.C.; Supervision, J.C.; Funding Acquisition, A.F.d.l.T. and J.C.

**Funding:** This research was funded by FONDECYT Regular grant number 1170718 (J.C.) and FONDECYT Postdoc grant number 3170003 (A.F.T). It also was founded by CONICYT-FONDEQUIP-EQM160063 (computational resources).

**Acknowledgments:** A. F. de la Torre is grateful to FONDECYT project No. 3170003 for financial support. J.A.M.-L. would like to thank Universidad Andres Bello for the facilities. We would also like to thank the computing resources provided by the Centro de Bioinformática y Simulación Molecular (CBSM), at the Universidad de Talca, Talca, Chile.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

## Abbreviations

3D	three-dimensional
AA	amino acid
HB	Hydrogen bonding
IFPs	interaction fingerprints analysis
LasB	<i>Pseudomonas</i> elastase
NAMdP	<i>N</i> -alpha-mercaptoacetyl dipeptide
PDB	Protein Data Bank
RMSD	root mean square deviation
SAR	structure–activity relationship

## References

- O'Neill, J. *Tackling Drug-resistant Infections Globally: Final Report and Recommendations*; IICA: San José, Costa Rica, 2016.
- Fischbach, M.A.; Walsh, C.T. Antibiotics for emerging pathogens. *Science* **2009**, *325*, 1089–1093. [[CrossRef](#)] [[PubMed](#)]
- Brown, E.D.; Wright, G.D. Antibacterial drug discovery in the resistance era. *Nature* **2016**, *529*, 336–343. [[CrossRef](#)] [[PubMed](#)]
- Silver, L.L. Challenges of antibacterial discovery. *Clin. Microbiol. Rev.* **2011**, *24*, 71–109. [[CrossRef](#)] [[PubMed](#)]
- Payne, D.J.; Gwynn, M.N.; Holmes, D.J.; Pompliano, D.L. Drugs for bad bugs: Confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* **2007**, *6*, 29–40. [[CrossRef](#)] [[PubMed](#)]
- Walsh, C.T.; Wenczewicz, T.A. Prospects for new antibiotics: A molecule-centered perspective. *J. Antibiot.* **2014**, *67*, 7–22. [[CrossRef](#)] [[PubMed](#)]
- Tommasi, R.; Iyer, R.; Miller, A.A. Antibacterial drug discovery: Some assembly required. *Acs Infect. Dis.* **2018**, *4*, 686–695. [[CrossRef](#)]
- Crofts, T.S.; Gasparrini, A.J.; Dantas, G. Next-generation approaches to understand and combat the antibiotic resistome. *Nat. Rev. Microbiol.* **2017**, *15*, 422–434. [[CrossRef](#)]
- Clatworthy, A.E.; Pierson, E.; Hung, D.T. Targeting virulence: A new paradigm for antimicrobial therapy. *Nat. Chem. Biol.* **2007**, *3*, 541–548. [[CrossRef](#)]
- Lyczak, J.B.; Cannon, C.L.; Pier, G.B. Establishment of *Pseudomonas aeruginosa* infection: Lessons from a versatile opportunist. *Microbes Infect.* **2000**, *2*, 1051–1060. [[CrossRef](#)]

11. Burrows, L.L. The Therapeutic pipeline for *Pseudomonas aeruginosa* infections. *Acs Infect. Dis.* **2018**, *4*, 1041–1047. [[CrossRef](#)]
12. Morihara, K.; Tsuzuki, H.; Oda, K. Protease and elastase of *Pseudomonas aeruginosa*: Inactivation of human plasma alpha 1-proteinase inhibitor. *Infect. Immun.* **1979**, *24*, 188–193. [[PubMed](#)]
13. Kessler, E.; Israel, M.; Landshman, N.; Chechick, A.; Blumberg, S. In vitro inhibition of *Pseudomonas aeruginosa* elastase by metal-chelating peptide derivatives. *Infect. Immun.* **1982**, *38*, 716–723. [[PubMed](#)]
14. Jacobsen, J.A.; Fullagar, J.L.; Miller, M.T.; Cohen, S.M. Identifying chelators for metalloprotein inhibitors using a fragment-based approach. *J. Med. Chem.* **2011**, *54*, 591–602. [[CrossRef](#)] [[PubMed](#)]
15. Fullagar, J.L.; Garner, A.L.; Struss, A.K.; Day, J.A.; Martin, D.P.; Yu, J.; Cai, X.; Janda, K.D.; Cohen, S.M. Antagonism of a zinc metalloprotease using a unique metal-chelating scaffold: Tropolones as inhibitors of *P. aeruginosa* elastase. *Chem. Commun.* **2013**, *49*, 3197. [[CrossRef](#)]
16. Thayer, M.M.; Flaherty, K.M.; McKay, D.B. Three-dimensional structure of the elastase of *Pseudomonas aeruginosa* at 1.5 angstroms resolution. *J. Biol. Chem.* **1991**, *266*, 2864–2871.
17. Cathcart, G.R.A.; Quinn, D.; Greer, B.; Harriott, P.; Lynas, J.F.; Gilmore, B.F.; Walker, B. Novel inhibitors of the *Pseudomonas aeruginosa* virulence factor LasB: A potential therapeutic approach for the attenuation of virulence mechanisms in Pseudomonas infection. *Antimicrob. Agents Chemother.* **2011**, *55*, 2670–2678. [[CrossRef](#)]
18. Zhu, J.; Cai, X.; Harris, T.L.; Gooyit, M.; Wood, M.; Lardy, M.; Janda, K.D. Disarming *Pseudomonas aeruginosa* virulence factor LasB by leveraging a caenorhabditis elegans infection model. *Chem. Biol.* **2015**, *22*, 483–491. [[CrossRef](#)]
19. Kany, A.M.; Sikandar, A.; Hauptenthal, J.; Yahiaoui, S.; Maurer, C.K.; Proschak, E.; Köhnke, J.; Hartmann, R.W. Binding mode characterization and early in vivo evaluation of fragment-like thiols as inhibitors of the virulence factor LasB from *Pseudomonas aeruginosa*. *ACS Infect. Dis.* **2018**, *4*, 988–997. [[CrossRef](#)]
20. Cathcart, G.R.; Gilmore, B.F.; Greer, B.; Harriott, P.; Walker, B. Inhibitor profiling of the *Pseudomonas aeruginosa* virulence factor LasB using N-alpha mercaptoamide template-based inhibitors. *Bioorganic Med. Chem. Lett.* **2009**, *19*, 6230–6232. [[CrossRef](#)]
21. Garner, A.L.; Struss, A.K.; Fullagar, J.L.; Agrawal, A.; Moreno, A.Y.; Cohen, S.M.; Janda, K.D. 3-Hydroxy-1-alkyl-2-methylpyridine-4(1 H)-thiones: Inhibition of the *Pseudomonas aeruginosa* Virulence Factor LasB. *Acs Med. Chem. Lett.* **2012**, *3*, 668–672. [[CrossRef](#)]
22. Golbraikh, A.; Tropsha, A. Beware of q<sup>2</sup>! *J. Mol. Graph. Model.* **2002**, *20*, 269–276. [[CrossRef](#)]
23. Roy, P.P.; Roy, K. On Some aspects of variable selection for partial least squares regression models. *Qsar Comb. Sci.* **2008**, *27*, 302–313. [[CrossRef](#)]
24. Aptula, A.O.; Jeliakova, N.G.; Schultz, T.W.; Cronin, M.T.D. The better predictive model: High q<sup>2</sup> for the training set or low root mean square error of prediction for the test set? *Qsar Comb. Sci.* **2005**, *24*, 385–396. [[CrossRef](#)]
25. Doweiko, A.M. 3D-QSAR illusions. *J. Comput. Aided Mol. Des.* **2004**, *18*, 587–596. [[CrossRef](#)] [[PubMed](#)]
26. Velázquez-Libera, J.L.; Navarro-Retamal, C.; Caballero, J. Insights into the structural requirements of 2(S)-amino-6-boronohexanoic acid derivatives as arginase I inhibitors: 3D-QSAR, docking, and interaction fingerprint studies. *Int. J. Mol. Sci.* **2018**, *19*, 2956. [[CrossRef](#)]
27. Ramírez, D.; Caballero, J. Is It Reliable to take the molecular docking top scoring position as the best solution without considering available structural data? *Molecules* **2018**, *23*, 1038. [[CrossRef](#)]
28. Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-based scoring function to predict protein-ligand interactions. *J. Mol. Biol.* **2000**, *295*, 337–356. [[CrossRef](#)]
29. Caballero, J.; Morales-Bayuelo, A.; Navarro-Retamal, C. Mycobacterium tuberculosis serine/threonine protein kinases: Structural information for the design of their specific ATP-competitive inhibitors. *J. Comput. Aided Mol. Des.* **2018**, *32*, 1315–1336. [[CrossRef](#)]
30. Velázquez-Libera, J.L.; Rossino, G.; Navarro-Retamal, C.; Collina, S.; Caballero, J. Docking, interaction fingerprint, and three-dimensional quantitative structure-activity relationship (3D-QSAR) of sigma1 receptor ligands, analogs of the neuroprotective agent RC-33. *Front. Chem.* **2019**, *7*, 496.
31. Shelley, J.C.; Cholleti, A.; Frye, L.L.; Greenwood, J.R.; Timlin, M.R.; Uchimaya, M. Epik: A software program for pK<sub>a</sub> prediction and protonation state generation for drug-like molecules. *J. Comput. -Aided Mol. Des.* **2007**, *21*, 681–691. [[CrossRef](#)]

32. Greenwood, J.R.; Calkins, D.; Sullivan, A.P.; Shelley, J.C. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *J. Comput. -Aided Mol. Des.* **2010**, *24*, 591–604. [[CrossRef](#)] [[PubMed](#)]
33. Jorgensen, W.L.; Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666. [[CrossRef](#)] [[PubMed](#)]
34. Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *J. Chem. Theory Comput.* **2010**, *6*, 1509–1519. [[CrossRef](#)] [[PubMed](#)]
35. Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shelley, M.; Perry, J.K.; et al. Glide: A new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749. [[CrossRef](#)]
36. Friesner, R.A.; Murphy, R.B.; Repasky, M.P.; Frye, L.L.; Greenwood, J.R.; Halgren, T.A.; Sanschagrin, P.C.; Mainz, D.T. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177–6196. [[CrossRef](#)]
37. Quesada-Romero, L.; Mena-Ulecia, K.; Tiznado, W.; Caballero, J. Insights into the interactions between maleimide derivatives and GSK3 $\beta$  combining molecular docking and QSAR. *PLoS ONE* **2014**, *9*, e102212. [[CrossRef](#)]
38. Quesada-Romero, L.; Caballero, J. Docking and quantitative structure–activity relationship of oxadiazole derivatives as inhibitors of GSK3 $\beta$ . *Mol. Divers.* **2014**, *18*, 149–159. [[CrossRef](#)]
39. Mena-Ulecia, K.; Tiznado, W.; Caballero, J. Study of the differential activity of thrombin inhibitors using docking, QSAR, molecular dynamics, and MM-GBSA. *PLoS ONE* **2015**, *10*, e0142774. [[CrossRef](#)]
40. Muñoz-Gutierrez, C.; Adasme-Carreño, F.; Fuentes, E.; Palomo, I.; Caballero, J. Computational study of the binding orientation and affinity of PPAR $\gamma$  agonists: Inclusion of ligand-induced fit by cross-docking. *Rsc. Adv.* **2016**, *6*, 64756–64768. [[CrossRef](#)]
41. Deng, Z.; Chuaqui, C.; Singh, J. Structural interaction fingerprint (SIFt): A novel method for analyzing three-dimensional protein–ligand binding interactions. *J. Med. Chem.* **2004**, *47*, 337–344. [[CrossRef](#)]
42. Singh, J.; Deng, Z.; Narale, G.; Chuaqui, C. Structural interaction fingerprints: A new approach to organizing, mining, analyzing, and designing protein–small molecule complexes. *Chem. Biol. Drug Des.* **2006**, *67*, 5–12. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).