

RESEARCH ARTICLE

# Study of the Differential Activity of Thrombin Inhibitors Using Docking, QSAR, Molecular Dynamics, and MM-GBSA

Karel Mena-Ulecia<sup>1,2</sup>, William Tiznado<sup>1</sup>, Julio Caballero<sup>2\*</sup>

**1** Departamento de Química, Facultad de Ciencias Exactas, Universidad Andres Bello, Avenida República 252, Santiago, Chile, **2** Centro de Bioinformática y Simulación Molecular, Facultad de Ingeniería, Universidad de Talca, 2 Norte 685, Casilla 721, Talca, Chile

\* [jcaballero@utalca.cl](mailto:jcaballero@utalca.cl) or [jmcr77@yahoo.com](mailto:jmcr77@yahoo.com)



**OPEN ACCESS**

**Citation:** Mena-Ulecia K, Tiznado W, Caballero J (2015) Study of the Differential Activity of Thrombin Inhibitors Using Docking, QSAR, Molecular Dynamics, and MM-GBSA. PLoS ONE 10(11): e0142774. doi:10.1371/journal.pone.0142774

**Editor:** Freddie Salsbury, Jr, Wake Forest University, UNITED STATES

**Received:** June 2, 2015

**Accepted:** October 27, 2015

**Published:** November 24, 2015

**Copyright:** © 2015 Mena-Ulecia et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by the grant no. 1130141 (J.C.) from Fondecyt, Chile. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

Non-peptidic thrombin inhibitors (TIs; 177 compounds) with diverse groups at motifs P<sub>1</sub> (such as oxyguanidine, amidinohydrazone, amidine, amidinopiperidine), P<sub>2</sub> (such as cyano-fluorophenylacetamide, 2-(2-chloro-6-fluorophenyl)acetamide), and P<sub>3</sub> (such as phenylethyl, arylsulfonate groups) were studied using molecular modeling to analyze their interactions with S<sub>1</sub>, S<sub>2</sub>, and S<sub>3</sub> subsites of the thrombin binding site. Firstly, a protocol combining docking and three dimensional quantitative structure–activity relationship was performed. We described the orientations and preferred active conformations of the studied inhibitors, and derived a predictive CoMSIA model including steric, donor hydrogen bond, and acceptor hydrogen bond fields. Secondly, the dynamic behaviors of some selected TIs (compounds **26**, **133**, **147**, **149**, **162**, and **177** in this manuscript) that contain different molecular features and different activities were analyzed by creating the solvated models and using molecular dynamics (MD) simulations. We used the conformational structures derived from MD to accomplish binding free energetic calculations using MM-GBSA. With this analysis, we theorized about the effect of van der Waals contacts, electrostatic interactions and solvation in the potency of TIs. In general, the contents reported in this article help to understand the physical and chemical characteristics of thrombin-inhibitor complexes.

## Introduction

Thromboembolic diseases are among the principal causes of mortality in the world. Vein thrombosis can progress to pulmonary embolism. These disorders, identified with the term venous thromboembolism (VTE), affect several million people around the world [1]. VTE is the third leading cause of cardiovascular-related death, after myocardial infarction and stroke [2].

The central role of the serine protease thrombin in thrombosis and haemostasis makes it an attractive target for antithrombotic therapy [3]. Thrombin catalyzes the conversion of soluble fibrinogen to insoluble fibrin in the clotting cascade, and also acts on other substrates such as

factor V, factor VIII, factor XI, and factor XIII. It is widely believed that an oral thrombin inhibitor (TI) could provide a new standard of care in anticoagulation therapy.

The discovery of small molecule TIs is an important goal for anti-thrombotic therapy [4]. In the last years, potent and selective inhibitors have been reported, such as pyridones, acetamides, oxyguanidines, aminopiperidines, amidines, and amidinohydrazones [5–13]. These sets have shown that subtle structural differences in compounds (due to the presence of similar scaffolds or the same scaffolds with different substituents) can lead to big differences in their thrombin inhibitory activities.

The knowledge of the relevant structural features that positively influence the activity of TIs is important for the design of potent compounds. Molecular modeling has demonstrated to be a powerful support to investigate bioactive compounds and their structure-activity relationship (SAR) with the main purpose of identifying the molecular features that contribute to a high bioactivity. These methods have been applied for studying TIs. Several quantitative structure-activity relationship (QSAR) models were reported using approaches such as classic QSAR [14], CoMFA/CoMSIA [15,16], topological descriptors [17], and artificial neural networks [18]. Other reports used docking and molecular dynamics (MD) simulations to study structural features of several TIs identified with a high activity [19–21]. In general, these reports do not include an analysis of the interactions between inhibitors and different subsites of the thrombin binding site. An exception is the work of Nilsson et al. [22]. These authors designed a set of compounds to bind to the  $S_2$  and  $S_3$  subsites with no interactions in  $S_1$ , and developed a classic QSAR model to study the chemical features that are important for the prediction of their binding constants. Other exception is the work of Bhunia et al. [23], which used three dimensional (3D) QSAR and MD simulations to profile structural determinants for the selectivity of representative diverse classes of thrombin-selective inhibitors.

In the current work, we applied some of the popular molecular modeling methods to study the interactions of TIs with  $S_1$ ,  $S_2$ , and  $S_3$  subsites of the thrombin binding site. We studied the orientations and SAR for 177 TIs by using a protocol that includes docking and the 3D-QSAR method CoMSIA. Additionally, we analyzed the dynamical behavior for some selected compounds (26, 133, 147, 149, 162, and 177) by using MD and free energy calculations. By means of a comparison of the selected systems, we gained further insight into the role played by different TI molecular constituents in the binding affinities due to interactions with different subsites in the thrombin active site.

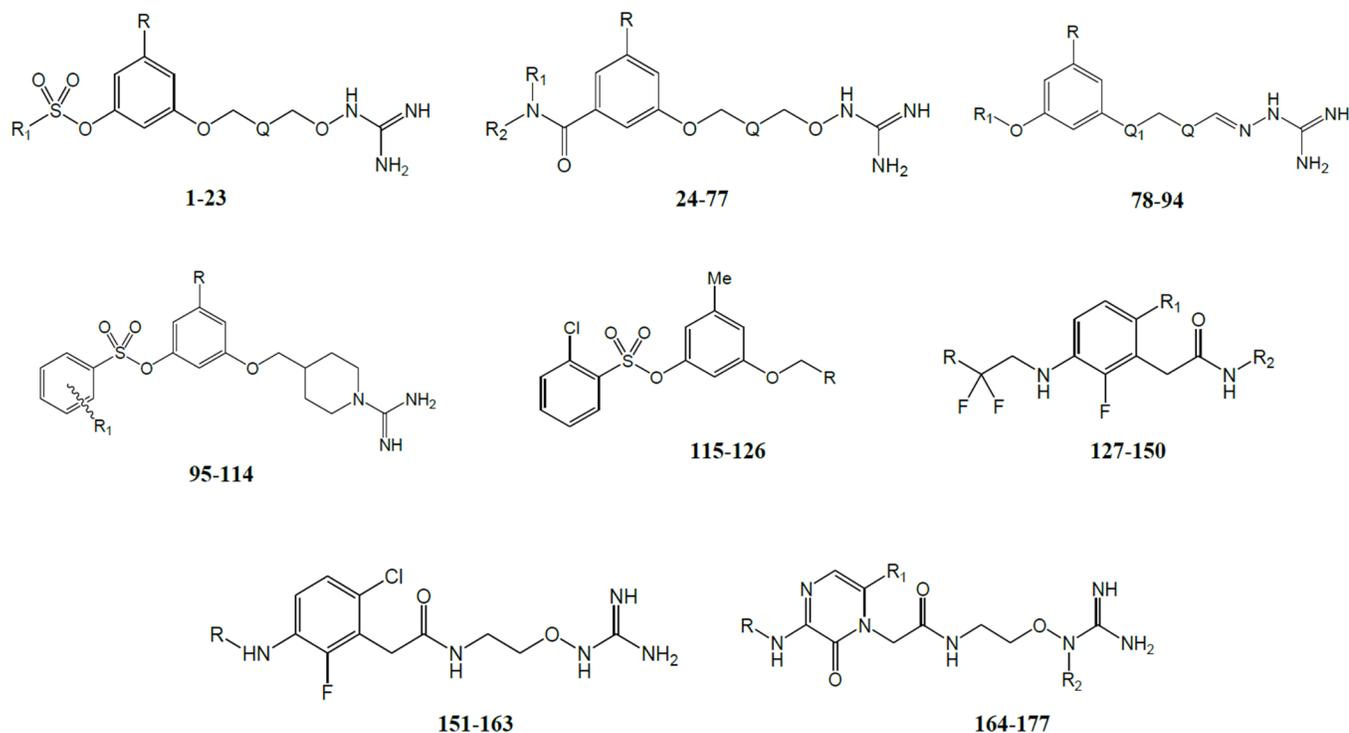
## Materials and Methods

### Data set

The primary structures and activities of 177 TIs were taken from the literature [5–13]. Inhibitory activities were collected and transformed into  $\log(10^3/K_i)$  values.  $K_i$  values are in nM and represent the enzyme inhibition constants. TI structures are in Fig 1 and their biological activities used in this study are summarized in Table 1. The chemical structures were sketched using the molecular editor of Maestro 9.0 software suite [24].

### Docking

The use of docking has extended in the last decades because successful applications to research in medicinal chemistry. Typically, docking programs are able to closely reproduce the binding orientations of the inhibitors in complexes for which crystallographic data are available [25–27]. Docking was performed using Glide [28], which is part of the Maestro 9.0 software suite [24]. Protein coordinates were extracted from the crystal structure of the thrombin–inhibitor complex with the code 1T4U in Protein Data Bank [6]; the protein structure was refined and



**Fig 1. Structures of TIs.**

doi:10.1371/journal.pone.0142774.g001

completed using the Protein Wizard Preparation module also available in Maestro software. A grid box of  $30\text{\AA} \times 30\text{\AA} \times 30\text{\AA}$  was centered on the center of mass of the inhibitor in this crystal structure. The module LigPrep 2.4 [29] was used to assign ionization states, stereochemistries, and ring conformations of the ligands. Docking parameters were used as in previous works [30], Glide standard (SP) and extra-precision (XP) modes were used. The better docking poses for each ligand were examined according to their relative total energy scores. Among docking poses, the more energetically favorable conformation was selected by considering the total energy value.

### CoMSIA calculations

Compound structures and inhibitory activities as  $\log(10^3/K_i)$  values were considered for CoMSIA calculations. The relative alignment of the compounds to set field calculations was inside the binding site. For this, the 3D conformations previously obtained by docking simulations were used. QSAR modeling was performed using the Sybyl 7.3 software of Tripos [31]. The data set was divided into two sub data sets (143 and 34 compounds were in training and test sets respectively) for external validation process. The molecules contained in the training set were placed in a rectangular grid extended beyond  $4\text{\AA}$  in each direction from the more external coordinates of each molecule. The interaction energies between a hypothetical atom (a  $sp^3$  hybridized carbon atom with +1 charge) and all compounds were computed at the defined points, using a volume-dependent lattice with  $2.0\text{\AA}$  grid spacing. Then, partial least squares (PLS) method was applied using standard Sybyl parameters, and considering an optimal number of components determined by optimization of leave-one-out (LOO) cross-validation  $Q^2$  value (using SAMPLS [32] sampling method). Similarity was expressed in terms of steric

**Table 1. Experimental and predicted thrombin inhibitory activities ( $\log(10^3/K_i)$  (nM)) using CoMSIA-SDA model.**

Compounds 1–23			
compound	R, R <sub>1</sub> , R <sub>2</sub> , Q	Exp. $\log(10^3/K_i)$	Calc. $\log(10^3/K_i)$
1	R = Me; R <sub>1</sub> = 2-Cl phenyl; Q = CH <sub>2</sub>	2.68	2.12
2	R = Me; R <sub>1</sub> = 2-CN phenyl; Q = CH <sub>2</sub>	2.00	1.56
3	R = Me; R <sub>1</sub> = 2-OMe phenyl; Q = CH <sub>2</sub>	2.17	2.10
4	R = Me; R <sub>1</sub> = 2-NH <sub>2</sub> phenyl; Q = CH <sub>2</sub>	1.06	1.40
5 <sup>a</sup>	R = Me; R <sub>1</sub> = 2-(MeSO <sub>2</sub> )Ph; Q = CH <sub>2</sub>	1.92	-
6	R = Me; R <sub>1</sub> = 2-(PheSO <sub>2</sub> )Ph; Q = CH <sub>2</sub>	2.32	2.84
7	R = Me; R <sub>1</sub> = 3-(MeSO <sub>2</sub> )Ph; Q = CH <sub>2</sub>	0.69	0.73
8	R = Me; R <sub>1</sub> = 2,4-(MeSO <sub>2</sub> ) <sub>2</sub> Ph; Q = CH <sub>2</sub>	0.13	0.14
9	R = Me; R <sub>1</sub> = 8-quinolinyl; Q = CH <sub>2</sub>	2.17	1.77
10	R = Me; R <sub>1</sub> = 5-isoquinolinyl; Q = CH <sub>2</sub>	1.88	2.05
11	R = Me; R <sub>1</sub> = 1,2,3,4-tetrahydro-8-quinolinyl; Q = CH <sub>2</sub>	1.51	1.37
12	R = Me; R <sub>1</sub> = 1,1-dioxido-2,3-dihydro-1-benzothien-6-yl; Q = CH <sub>2</sub>	0.99	0.82
13	R = Me; R <sub>1</sub> = 2-(MeSO <sub>2</sub> )Ph; Q = C(-C <sub>2</sub> H <sub>4</sub> -)	2.11	2.06
14	R = Me; R <sub>1</sub> = 2-(MeSO <sub>2</sub> )Ph; Q = CHOH	0.86	0.71
15	R = Me; R <sub>1</sub> = 2-(MeSO <sub>2</sub> )Ph; Q = CHF	0.94	0.83
16	R = Me; R <sub>1</sub> = 2-(MeSO <sub>2</sub> )Ph; Q = CF <sub>2</sub>	0.44	0.68
17	R = Me; R <sub>1</sub> = 2-(MeSO <sub>2</sub> )Ph; Q = C = CH <sub>2</sub>	1.49	1.20
18 <sup>a</sup>	R = Me; R <sub>1</sub> = 2-(MeSO <sub>2</sub> )Ph; Q = bond	-0.42	-
19	R = OMe; R <sub>1</sub> = 2-(MeSO <sub>2</sub> )Ph; Q = CH <sub>2</sub>	1.46	1.49
20	R = Et; R <sub>1</sub> = 2-(MeSO <sub>2</sub> )Ph; Q = CH <sub>2</sub>	1.53	1.91
21	R = Cl; R <sub>1</sub> = 2-(MeSO <sub>2</sub> )Ph; Q = CH <sub>2</sub>	1.41	1.40
22	R = CH <sub>2</sub> OH; R <sub>1</sub> = 2-(MeSO <sub>2</sub> )Ph; Q = C(-C <sub>2</sub> H <sub>4</sub> -)	0.40	0.43
23	R = F; R <sub>1</sub> = 2-(MeSO <sub>2</sub> )Ph; Q = C(-C <sub>2</sub> H <sub>4</sub> -)	1.31	1.34
Compounds 24–77			
compound	R, R <sub>1</sub> , R <sub>2</sub> , Q	Exp. $\log(10^3/K_i)$	Calc. $\log(10^3/K_i)$
24	R = Cl; R <sub>1</sub> = allyl; R <sub>2</sub> = cyclopentyl; Q = CH <sub>2</sub>	1.68	1.55
25	R = Cl; R <sub>1</sub> = Me; R <sub>2</sub> = cyclohexyl; Q = CH <sub>2</sub>	1.48	1.62
26 <sup>b,c</sup>	R = Cl; R <sub>1</sub> = Me; R <sub>2</sub> = cyclopentyl; Q = CH <sub>2</sub>	0.43	1.46
27	R = Cl; R <sub>1</sub> = Pr; R <sub>2</sub> = cyclopentylmethyl; Q = CH <sub>2</sub>	1.42	1.66
28	R = Cl; R <sub>1</sub> = Pr; R <sub>2</sub> = cyclobutylmethyl; Q = CH <sub>2</sub>	1.80	1.87
29 <sup>b</sup>	R = Cl; R <sub>1</sub> = Pr; R <sub>2</sub> = cyclopropylmethyl; Q = CH <sub>2</sub>	1.64	1.57
30 <sup>b</sup>	R = Cl; R <sub>1</sub> = 2-carboxyethyl; R <sub>2</sub> = cyclopropylmethyl; Q = CH <sub>2</sub>	0.67	1.61
31	R = Cl; R <sub>1</sub> = 3-methoxy-3-oxopropyl; R <sub>2</sub> = cyclopropylmethyl; Q = CH <sub>2</sub>	1.42	1.79
32	R = Cl; R <sub>1</sub> = 2-methoxyethyl; R <sub>2</sub> = cyclopropylmethyl; Q = CH <sub>2</sub>	0.81	0.58
33 <sup>a</sup>	R = Cl; R <sub>1</sub> = 2-(1-pyrrolidinyl)ethyl; R <sub>2</sub> = cyclopropylmethyl; Q = CH <sub>2</sub>	0.69	-
34 <sup>b</sup>	R = Cl; R <sub>1</sub> = 2-(4-morpholinyl)ethyl; R <sub>2</sub> = cyclopropylmethyl; Q = CH <sub>2</sub>	0.80	1.45
35	R = Cl; R <sub>1</sub> = 2-(2-pyridinyl)ethyl; R <sub>2</sub> = cyclopropylmethyl; Q = CH <sub>2</sub>	1.23	1.13
36	R = Cl; R <sub>1</sub> = 2-(3-pyridinyl)ethyl; R <sub>2</sub> = cyclopropylmethyl; Q = CH <sub>2</sub>	1.33	1.37
37	R = Cl; R <sub>1</sub> = 2-(4-pyridinyl)ethyl; R <sub>2</sub> = cyclopropylmethyl; Q = CH <sub>2</sub>	1.28	1.33
38	R = Cl; R <sub>1</sub> = allyl; R <sub>2</sub> = 3-furylmethyl; Q = CH <sub>2</sub>	1.70	1.16
39	R = Cl; R <sub>1</sub> = 3-furylmethyl; R <sub>2</sub> = cyclopropylmethyl; Q = CH <sub>2</sub>	1.77	1.73
40	R = Cl; R <sub>1</sub> = allyl; R <sub>2</sub> = 2-thienylmethyl; Q = CH <sub>2</sub>	0.83	0.94
41	R = Cl; R <sub>1</sub> = Et; R <sub>2</sub> = 4-pyridinylmethyl; Q = CH <sub>2</sub>	-0.20	-0.34
42	R = Cl; R <sub>1</sub> = Et; R <sub>2</sub> = Bn; Q = CH <sub>2</sub>	1.04	1.68
43	R = Cl; R <sub>1</sub> , R <sub>2</sub> = -(CH <sub>2</sub> ) <sub>5</sub> -; Q = CH <sub>2</sub>	0.06	0.60
44	R = Cl; R <sub>1</sub> , R <sub>2</sub> = -CH <sub>2</sub> -CH <sub>2</sub> -O-CH <sub>2</sub> -CH <sub>2</sub> -; Q = CH <sub>2</sub>	-0.88	-1.25

(Continued)

Table 1. (Continued)

45	R = Cl; R <sub>1</sub> , R <sub>2</sub> = -(CH <sub>2</sub> ) <sub>6</sub> -; Q = CH <sub>2</sub>	0.33	0.73
46	R = Cl; R <sub>1</sub> , R <sub>2</sub> = -CH <sub>2</sub> -CH <sub>2</sub> -CH(Bn)-CH <sub>2</sub> -CH <sub>2</sub> -; Q = CH <sub>2</sub>	-0.18	0.40
47	R = Cl; R <sub>1</sub> , R <sub>2</sub> = -CH <sub>2</sub> -CH <sub>2</sub> -N(Ph)-CH <sub>2</sub> -CH <sub>2</sub> -; Q = CH <sub>2</sub>	-0.46	-0.36
48	R = Cl; R <sub>1</sub> , R <sub>2</sub> = -CH <sub>2</sub> -CH <sub>2</sub> -N(2-pyridinyl)-CH <sub>2</sub> -CH <sub>2</sub> -; Q = CH <sub>2</sub>	0.19	0.23
49	R = Cl; R <sub>1</sub> , R <sub>2</sub> = -CH <sub>2</sub> -C(-C <sub>4</sub> H <sub>4</sub> -)C-CH <sub>2</sub> -CH <sub>2</sub> -; Q = CH <sub>2</sub>	0.77	1.03
50	R = Me; R <sub>1</sub> = allyl; R <sub>2</sub> = cyclopentyl; Q = CH <sub>2</sub>	2.00	1.56
51 <sup>b</sup>	R = Me; R <sub>1</sub> = Pr; R <sub>2</sub> = cyclobutylmethyl; Q = CH <sub>2</sub>	1.77	1.74
52	R = Me; R <sub>1</sub> = Et; R <sub>2</sub> = cyclohexyl; Q = CH <sub>2</sub>	2.22	2.04
53	R = Me; R <sub>1</sub> = 6-amino-6-oxohexyl; R <sub>2</sub> = cyclohexyl; Q = CH <sub>2</sub>	2.10	1.86
54	R = Me; R <sub>1</sub> = allyl; R <sub>2</sub> = cyclohexyl; Q = CH <sub>2</sub>	2.05	1.97
55	R = Me; R <sub>1</sub> = 3-ethoxy-3-oxopropyl; R <sub>2</sub> = cyclopropylmethyl; Q = CH <sub>2</sub>	1.19	1.09
56 <sup>b</sup>	R = Me; R <sub>1</sub> = Pr; R <sub>2</sub> = 3-furylmethyl; Q = CH <sub>2</sub>	2.05	1.54
57 <sup>b</sup>	R = Me; R <sub>1</sub> = Bu; R <sub>2</sub> = 3-furylmethyl; Q = CH <sub>2</sub>	2.10	1.57
58	R = Me; R <sub>1</sub> = iPr; R <sub>2</sub> = 3-furylmethyl; Q = CH <sub>2</sub>	2.16	1.99
59	R = Me; R <sub>1</sub> = cyclopropylmethyl; R <sub>2</sub> = 3-furylmethyl; Q = CH <sub>2</sub>	1.96	1.63
60	R = Me; R <sub>1</sub> = iPr; R <sub>2</sub> = 3-furylmethyl; Q = CH <sub>2</sub>	1.17	1.43
61	R = Me; R <sub>1</sub> = cyclobutyl; R <sub>2</sub> = 3-furylmethyl; Q = CH <sub>2</sub>	1.66	1.81
62	R = Me; R <sub>1</sub> = cyclohexylmethyl; R <sub>2</sub> = 3-furylmethyl; Q = CH <sub>2</sub>	1.68	1.80
63	R = Me; R <sub>1</sub> = 2-methoxyethyl; R <sub>2</sub> = 3-furylmethyl; Q = CH <sub>2</sub>	1.55	1.50
64	R = Me; R <sub>1</sub> = 3-amino-3-oxopropyl; R <sub>2</sub> = 3-furylmethyl; Q = CH <sub>2</sub>	0.96	1.18
65	R = Me; R <sub>1</sub> = 2-(ethylsulfanyl)ethyl; R <sub>2</sub> = 3-furylmethyl; Q = CH <sub>2</sub>	2.22	2.01
66	R = Me; R <sub>1</sub> = Pr; R <sub>2</sub> = 3-thienylmethyl; Q = CH <sub>2</sub>	1.32	1.43
67	R = Me; R <sub>1</sub> = Pr; R <sub>2</sub> = 3-pyridinylmethyl; Q = CH <sub>2</sub>	1.28	1.94
68 <sup>b</sup>	R = Me; R <sub>1</sub> = Pr; R <sub>2</sub> = 4-F benzyl; Q = CH <sub>2</sub>	1.68	1.37
69	R = Me; R <sub>1</sub> = cyclopropylmethyl; R <sub>2</sub> = 4-pyridinylmethyl; Q = CH <sub>2</sub>	1.62	1.35
70	R = Me; R <sub>1</sub> = Pr; R <sub>2</sub> = 1,3-thiazol-2-ylmethyl; Q = CH <sub>2</sub>	0.71	0.52
71	R = Me; R <sub>1</sub> = cyclopropylmethyl; R <sub>2</sub> = 3-thienylmethyl; Q = CH <sub>2</sub>	1.51	1.37
72 <sup>b</sup>	R = Me; R <sub>1</sub> = Pr; R <sub>2</sub> = 2-pyridinylmethyl; Q = CH <sub>2</sub>	1.03	2.29
73	R = Me; R <sub>1</sub> = cyclopropylmethyl; R <sub>2</sub> = 3-pyridinylmethyl; Q = CH <sub>2</sub>	1.47	1.50
74	R = Me; R <sub>1</sub> = Pr; R <sub>2</sub> = 4-pyridinylmethyl; Q = CH <sub>2</sub>	1.14	1.00
75 <sup>b</sup>	R = Me; R <sub>1</sub> = cyclopropylmethyl; R <sub>2</sub> = 2-pyridinylmethyl; Q = CH <sub>2</sub>	1.07	1.27
76	R = Me; R <sub>1</sub> = Pr; R <sub>2</sub> = 3-furylmethyl; Q = C(-C <sub>2</sub> H <sub>4</sub> -)	2.40	2.52
77 <sup>b</sup>	R = Me; R <sub>1</sub> = cyclopropylmethyl; R <sub>2</sub> = 3-furylmethyl; Q = C(-C <sub>2</sub> H <sub>4</sub> -)	2.40	0.96
<b>Compounds 78–94</b>			
compound	R, R <sub>1</sub> , Q, Q <sub>1</sub>	Exp. Log(10 <sup>3</sup> /Ki)	Calc. Log(10 <sup>3</sup> /Ki)
78	R = Me; R <sub>1</sub> = (2-Cl phenyl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	2.08	1.96
79	R = Me; R <sub>1</sub> = (2,3-dichlorophenyl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	2.03	2.19
80	R = Me; R <sub>1</sub> = (2-OMe phenyl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	1.96	1.73
81	R = Me; R <sub>1</sub> = (2-OCF <sub>3</sub> phenyl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	1.96	1.74
82 <sup>b</sup>	R = Me; R <sub>1</sub> = (3-Me phenyl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	2.22	1.83
83	R = Me; R <sub>1</sub> = (2-(MeSO <sub>2</sub> ) phenyl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	1.70	1.68
84	R = Me; R <sub>1</sub> = (2-CN phenyl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	2.04	2.02
85	R = Me; R <sub>1</sub> = (2-CF <sub>3</sub> phenyl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	2.36	2.23
86 <sup>b</sup>	R = Me; R <sub>1</sub> = (5-Cl thiophen-2-yl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	2.04	1.19
87 <sup>b</sup>	R = Me; R <sub>1</sub> = (naphthalen-1-yl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	1.36	1.07
88	R = Me; R <sub>1</sub> = (3-pyridinyl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	1.44	1.71
89	R = Me; R <sub>1</sub> = (quinolin-8-yl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	2.33	1.86
90 <sup>a</sup>	R = Me; R <sub>1</sub> = 2-CF <sub>3</sub> benzyl; Q = CH <sub>2</sub> ; Q <sub>1</sub> = O	0.38	-

(Continued)

Table 1. (Continued)

<b>91</b>	R = OMe; R <sub>1</sub> = (2-Cl phenyl)SO <sub>2</sub> ; Q = CH <sub>2</sub> ; Q <sub>1</sub> = CH <sub>2</sub>	0.00	-0.17
<b>92<sup>b</sup></b>	R = Me; R <sub>1</sub> = (2-OMe phenyl)SO <sub>2</sub> ; Q = bond; Q <sub>1</sub> = O	0.00	1.20
<b>93</b>	R = Me; R <sub>1</sub> = (2-OMe phenyl)SO <sub>2</sub> ; Q = (CH <sub>2</sub> ) <sub>2</sub> ; Q <sub>1</sub> = O	-0.11	0.24
<b>94<sup>b</sup></b>	R = Me; R <sub>1</sub> = (2-CN phenyl)SO <sub>2</sub> ; Q = C(-C <sub>2</sub> H <sub>4</sub> -); Q <sub>1</sub> = O	1.40	1.50
<b>Compounds 95–114</b>			
compound	R, R <sub>1</sub>	Exp. Log(10 <sup>3</sup> /Ki)	Calc. Log(10 <sup>3</sup> /Ki)
<b>95<sup>b</sup></b>	R = Me; R <sub>1</sub> = H	1.50	1.45
<b>96</b>	R = Me; R <sub>1</sub> = 2-CF <sub>3</sub>	1.85	1.40
<b>97</b>	R = Me; R <sub>1</sub> = 2-NO <sub>2</sub>	1.80	1.27
<b>98</b>	R = Me; R <sub>1</sub> = 2-NH <sub>2</sub>	1.28	1.41
<b>99</b>	R = Me; R <sub>1</sub> = 2-CO <sub>2</sub> Me	1.55	1.40
<b>100</b>	R = Me; R <sub>1</sub> = 3-Cl	1.41	1.17
<b>101<sup>b</sup></b>	R = Me; R <sub>1</sub> = 3-CF <sub>3</sub>	0.35	1.05
<b>102</b>	R = Me; R <sub>1</sub> = 3-Me	1.44	1.34
<b>103</b>	R = Me; R <sub>1</sub> = 3-NO <sub>2</sub>	0.94	0.95
<b>104</b>	R = Me; R <sub>1</sub> = 3-NH <sub>2</sub>	1.50	1.40
<b>105</b>	R = Me; R <sub>1</sub> = 2,3-diCl	1.66	1.14
<b>106</b>	R = Me; R <sub>1</sub> = 4-NO <sub>2</sub>	0.49	0.91
<b>107</b>	R = Me; R <sub>1</sub> = 4-NH <sub>2</sub>	0.40	0.45
<b>108<sup>b</sup></b>	R = Me; R <sub>1</sub> = 2,3 -CH = CH-CH = CH-	1.46	1.11
<b>109<sup>b</sup></b>	R = H; R <sub>1</sub> = 2-Cl	0.72	0.88
<b>110</b>	R = CO <sub>2</sub> Me; R <sub>1</sub> = 2-Cl	-0.57	-0.53
<b>111</b>	R = OMe; R <sub>1</sub> = 2-Cl	1.36	1.58
<b>112</b>	R = Et; R <sub>1</sub> = 2-Cl	1.41	1.48
<b>113</b>	R = CH <sub>2</sub> OH; R <sub>1</sub> = 2-Cl	0.31	0.40
<b>114</b>	R = Cl; R <sub>1</sub> = 2-Cl	1.52	1.05
<b>Compounds 115–126</b>			
compound	R	Exp. Log(10 <sup>3</sup> /Ki)	Calc. Log(10 <sup>3</sup> /Ki)
<b>115</b>	R = 3-amino-3-iminopropyl	0.34	0.73
<b>116</b>	R = 5-amino-5-iminopentyl	1.16	0.94
<b>117</b>	R = 4-[amino(imino)methyl]cyclohexyl	-0.32	0.27
<b>118<sup>b</sup></b>	R = 4-[amino(imino)methyl]phenyl	0.96	0.54
<b>119</b>	R = 3-[amino(imino)methyl]phenyl	1.00	0.77
<b>120</b>	R = 1-ethanimidoyl-4-piperidiny	-0.11	0.85
<b>121</b>	R = [methyl(4-pyridinyl)amino]methyl	1.96	2.27
<b>122<sup>b</sup></b>	R = 1-[amino(imino)methyl]-4-piperidiny	2.34	1.09
<b>123<sup>b</sup></b>	R = 1-[amino(imino)methyl]-3-piperidiny	0.50	0.03
<b>124</b>	R = 3-[[amino(imino)methyl]amino]ethyl	1.48	1.23
<b>125<sup>b</sup></b>	R = 3-[[amino(imino)methyl]amino]propyl	1.89	1.45
<b>126</b>	R = 3-[[amino(imino)methyl]amino]butyl	0.59	0.41
<b>Compounds 127–150</b>			
compound	R, R <sub>1</sub> , R <sub>2</sub>	Exp. Log(10 <sup>3</sup> /Ki)	Calc. Log(10 <sup>3</sup> /Ki)
<b>127</b>	R = phenyl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	2.64	3.10
<b>128</b>	R = pyridin-2-yl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	2.92	3.34
<b>129</b>	R = 3-methylpyridin-2-yl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	2.00	2.45
<b>130</b>	R = 4-methylpyridin-2-yl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	2.54	2.67
<b>131</b>	R = 5-methylpyridin-2-yl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	2.92	2.71
<b>132</b>	R = 6-methylpyridin-2-yl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	3.10	2.43

(Continued)

Table 1. (Continued)

133 <sup>c</sup>	R = 5-Cl-pyridin-2-yl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	3.24	3.35
134	R = 8-quinolyl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	3.42	2.96
135	R = 3-Cl-phenyl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	2.77	2.73
136	R = 3-F-phenyl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	2.50	2.46
137 <sup>b</sup>	R = 3,4-diF-phenyl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	2.75	3.43
138	R = 2-SO <sub>2</sub> Me-phenyl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	3.19	2.94
139 <sup>b</sup>	R = 1-oxy-2-pyridyl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	2.89	3.29
140 <sup>a</sup>	R = 5-Cl-1-oxy-2-pyridyl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-guanidinoxy-ethyl	3.18	-
141	R = pyridin-2-yl; R <sub>1</sub> = CN; R <sub>2</sub> = 2-methylguanidinoxy-ethyl	2.05	1.71
142	R = pyridin-2-yl; R <sub>1</sub> = CN; R <sub>2</sub> = (6-aminopyridin-3-yl)methyl	1.96	2.30
143	R = pyridin-2-yl; R <sub>1</sub> = CN; R <sub>2</sub> = (6-amino-2-methylpyridin-3-yl)methyl	3.11	2.87
144	R = pyridin-2-yl; R <sub>1</sub> = CN; R <sub>2</sub> = (3-fluoropyridin-2-yl)methyl	1.48	2.06
145	R = pyridin-2-yl; R <sub>1</sub> = CN; R <sub>2</sub> = (3-aminobenzo[d]isoxazol-6-yl)methyl	1.72	1.57
146	R = Ph; R <sub>1</sub> = Cl; R <sub>2</sub> = (6-NH <sub>2</sub> -2-Me-3-pyridyl)methyl	1.99	2.54
147 <sup>c</sup>	R = Ph; R <sub>1</sub> = Cl; R <sub>2</sub> = (6-NH <sub>2</sub> -2,4-dimethyl-3-pyridyl)methyl	2.48	2.45
148	R = 4-F-1-naphthyl; R <sub>1</sub> = Cl; R <sub>2</sub> = (6-NH <sub>2</sub> -2-Me-3-pyridyl)methyl	2.55	2.54
149 <sup>c</sup>	R = 5-Cl-2-pyridyl; R <sub>1</sub> = Cl; R <sub>2</sub> = (6-NH <sub>2</sub> -2-Me-3-pyridyl)methyl	3.16	2.69
150	R = 5-Cl-1-oxy-2-pyridyl; R <sub>1</sub> = Cl; R <sub>2</sub> = (6-NH <sub>2</sub> -2-Me-3-pyridyl)methyl	3.05	2.89
<b>Compounds 151–163</b>			
compound	R	Exp. Log(10 <sup>3</sup> /Ki)	Calc. Log(10 <sup>3</sup> /Ki)
151 <sup>b</sup>	R = 2,2-difluoro-2-phenylethyl	1.33	1.61
152	R = 2,2-difluoro-2-(pyridin-3-yl)ethyl	1.47	1.33
153	R = 2,2-difluoro-2-(quinolin-8-yl)ethyl	2.57	2.48
154	R = 2,2-difluoro-2-(isoquinolin-5-yl)ethyl	2.33	2.88
155 <sup>b</sup>	R = phenylsulfonyl	-0.11	0.19
156	R = 2-methyl-2-(pyridin-2-yl)propyl	1.24	1.55
157	R = 2-methyl-2-(pyridin-3-yl)propyl	0.89	0.42
158 <sup>b</sup>	R = 2,2-difluoro-2-(1-fluoronaphthalen-4-yl)ethyl	2.51	2.87
159	R = 2,2-difluoro-2-(pyridin-2-yl)ethyl	2.07	2.07
160	R = 2,2-difluoro-2-(quinolin-3-yl)ethyl	0.77	1.02
161	R = 2-(5-chloropyridin-2-yl)-2,2-difluoroethyl	2.75	2.55
162 <sup>b,c</sup>	R = (1-(pyridin-2-yl)cyclopropyl)methyl	1.24	0.95
163 <sup>b</sup>	R = (1-(pyridin-3-yl)cyclopropyl)methyl	0.85	1.53
<b>Compounds 164–177</b>			
compound	R, R <sub>1</sub> , R <sub>2</sub>	Exp. Log(10 <sup>3</sup> /Ki)	Calc. Log(10 <sup>3</sup> /Ki)
164	R = 2,2-difluoro-2-phenylethyl; R <sub>1</sub> = CN; R <sub>2</sub> = H	2.64	2.16
165	R = 2,2-difluoro-2-phenylethyl; R <sub>1</sub> = Me; R <sub>2</sub> = H	2.89	2.31
166	R = 2-(4-fluorophenyl)ethyl; R <sub>1</sub> = Me; R <sub>2</sub> = H	1.89	1.84
167	R = 2-(3,4-difluorophenyl)ethyl; R <sub>1</sub> = Me; R <sub>2</sub> = H	1.33	1.55
168	R = 2-(2,4-difluorophenyl)ethyl; R <sub>1</sub> = Me; R <sub>2</sub> = H	1.82	1.88
169	R = 2-(4-trifluoromethylphenyl)ethyl; R <sub>1</sub> = Me; R <sub>2</sub> = H	1.33	1.82
170 <sup>b</sup>	R = 2-(4-methoxyphenyl)ethyl; R <sub>1</sub> = Me; R <sub>2</sub> = H	1.96	1.63
171	R = 2-(3,4-dimethoxyphenyl)ethyl; R <sub>1</sub> = Me; R <sub>2</sub> = H	0.92	1.56
172 <sup>b</sup>	R = 2-(4-ethylphenyl)ethyl; R <sub>1</sub> = Me; R <sub>2</sub> = H	1.18	1.23
173	R = 2-(5-indanyl)ethyl; R <sub>1</sub> = Me; R <sub>2</sub> = H	1.36	1.64
174	R = 2-(1-naphthyl)ethyl; R <sub>1</sub> = Me; R <sub>2</sub> = H	1.92	1.76
175 <sup>b</sup>	R = 2,2-diphenylethyl; R <sub>1</sub> = Me; R <sub>2</sub> = H	2.17	1.67
176	R = 2,2-difluoro-2-phenylethyl; R <sub>1</sub> = Cl; R <sub>2</sub> = H	2.77	2.25

(Continued)

**Table 1.** (Continued)

<b>177<sup>c</sup></b>	R = 2-(4-methylphenyl)ethyl; R <sub>1</sub> = Me, R <sub>2</sub> = Me	1.43	1.53
------------------------	---	------	------

<sup>a</sup> QSAR outliers.

<sup>b</sup> test set compounds.

<sup>c</sup> compounds selected for MM-GBSA calculations

doi:10.1371/journal.pone.0142774.t001

occupancy, electrostatic interactions, local hydrophobicity, hydrogen bond (HB) donor, and HB acceptor properties, using a 0.3 attenuation factor in all CoMSIA applications.

### Molecular dynamics and MM-GBSA calculations

Complexes of the compounds **26**, **133**, **147**, **149**, **162**, and **177** inside thrombin active site were studied by MD simulations. All simulations were performed using the NAMD software package [33]. The initial coordinates for the MD calculations were taken from the docking experiments. To mimic the aqueous environment, an equilibrated water box with sides of 100 Å, centered on the mass center of each inhibitor, was used to solvate the thrombin-ligand system. Protein and inhibitors were described using the optimized potential for liquid simulations CHARMM force field [34] and the CGenFF force field [35], respectively. In turn, the water molecules belonging to the solvent box were described using the flexible TIP3P potential [36,37].

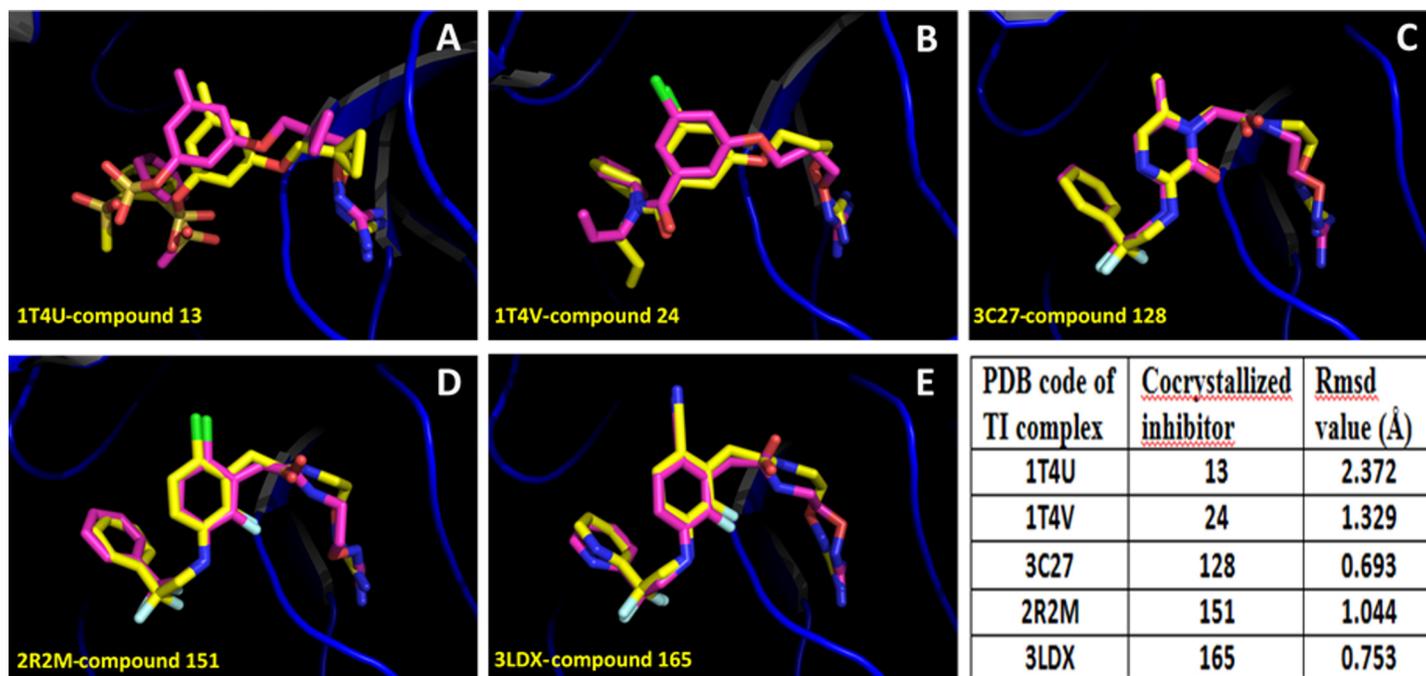
Energy minimization was performed on the models using conjugate gradient method (20,000 steps) to reduce any close contacts. Afterwards, the system was further equilibrated during 2.0 ns. The production run consisted of an MD simulation of 5.0 ns. In all cases, we applied a constraint to the backbone amino acids, we used the NPT ensemble at 300 K, we set a time step of 1 fs to solve the equations of motion, and we set a switched cutoff distance of 9.0 Å.

The free energy calculations were accomplished using the MM-GBSA method [38]. This method combines molecular mechanics energy and implicit solvation models at a reasonable computational cost, leading to outstanding results in several biological systems in recent years [39–41]. We applied this method to 500 snapshots extracted from the 5.0 ns production MD trajectories (explicit TIP3P water molecules and ions were removed for this).

Protein–ligand binding free energy using MM-GBSA was calculated as the difference between the energy of the bound complex and the energy of the unbound protein and inhibitor compound. The method allows for free energy decomposition into contributions originating from different types of physico-chemical interactions. Specifically, the energy is calculated for the protein–ligand complex, the ligand, and the protein, and their energies were computed using the CHARMM force field with the generalized Born implicit solvent model, in order to calculate the averaged binding free energy ( $\Delta G$ ) according to the following equation:

$$\Delta G_{binding} = \Delta E_{MM} + \Delta G_{solv} - T\Delta S \quad (1)$$

where  $\Delta E_{MM}$  includes  $\Delta E_{internal}$  (bond, angle, and dihedral energies),  $\Delta E_{elect}$  (electrostatic), and  $\Delta E_{vdw}$  (van der Waals) energies;  $\Delta G_{solv}$  is the electrostatic solvation energy (polar and non-polar contributions). The polar contribution is calculated using the Generalized Born (GB) model, while the non-polar energy is estimated using the solvent accessible surface area (SASA). The conformational entropy change  $-T\Delta S$  can be computed by normal-mode analysis on a set of conformational snapshots taken from MD simulations, but many authors have been reported that the lack of the evaluation of the entropy is not critical for calculating the



**Fig 2. Alignment of inhibitor docked structures on inhibitor X-ray reference structures, for the TI complexes.** (A) Compound 13 (PDB: 1T4U); (B) compound 24 (PDB: 1T4V); (C) compound 128 (PDB: 3C27); (D) compound 151 (PDB: 2R2M); (E) compound 165 (PDB: 3LDX). Crystal structures are represented in yellow, and docking results are represented in purple. Docking accuracy is reported by means of RMSD values.

doi:10.1371/journal.pone.0142774.g002

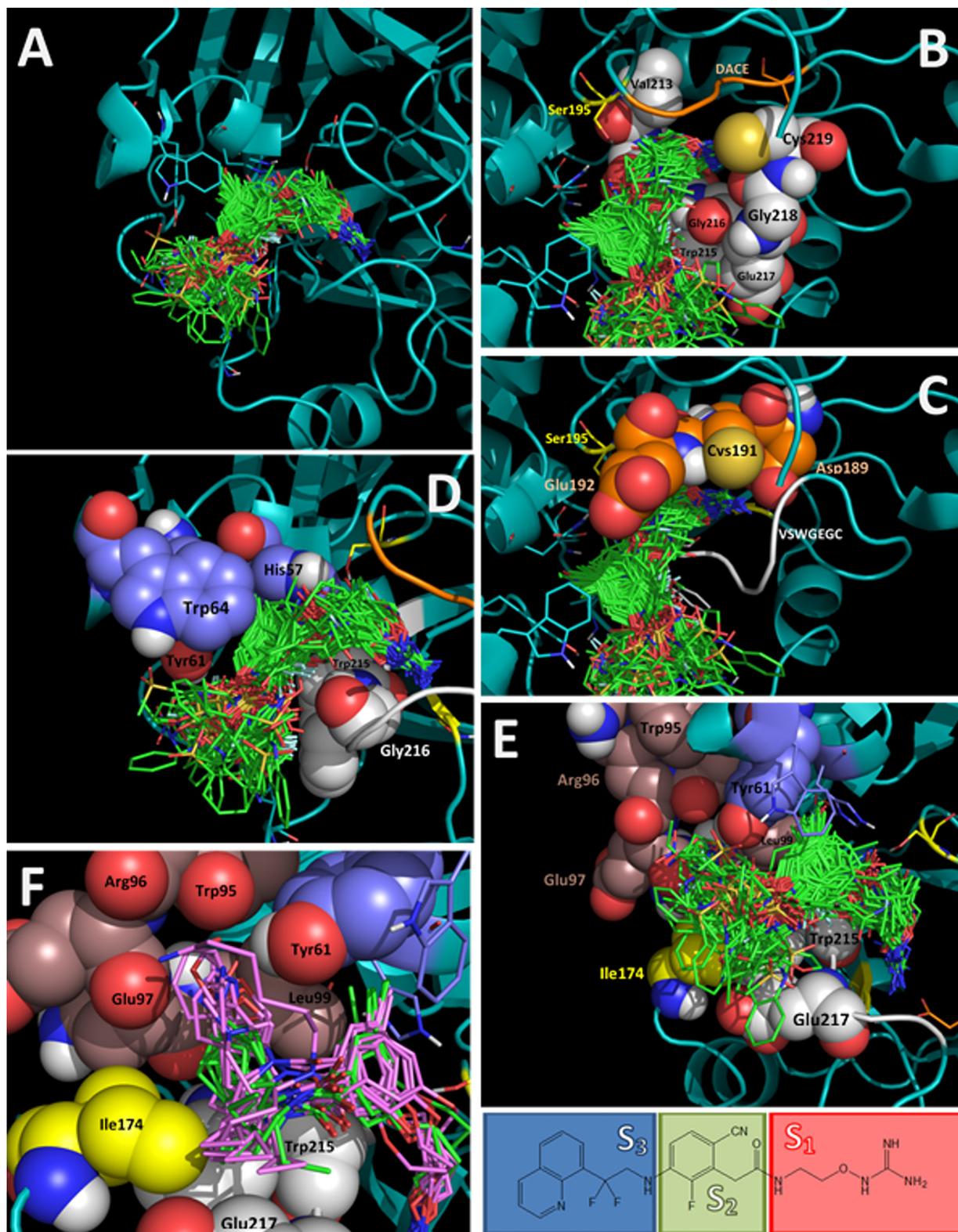
MM-GBSA (or MM-PBSA) free energies for similar systems [38,42,43]. In this work MM-GBSA calculations were also achieved in NAMD software [33]; the entropy term  $-T\Delta S$  was not calculated to reduce computational time.

## Results and Discussion

### Docking results

First, we applied docking methodology in order to reproduce the crystal structures of thrombin-ligand complexes for compounds 13, 24, 128, 151, and 165 (accession codes in PDB: 1T4U, 1T4V, 3C27, 2R2M, and 3LDX respectively). This initial test was used to assess the quality of the docking method to reproduce known structures. In Fig 2 is shown that the docked structures fitted in an acceptable way with available inhibitor X-ray crystal structures; all the inhibitors were adequately oriented. The values of the root mean square deviation (RMSD) for the docked structures with respect to the co-crystal inhibitor structures considering all heavy atoms were  $< 2.0 \text{ \AA}$  in almost all the cases analyzed. Considering that  $2.0 \text{ \AA}$  is the threshold value that differentiates between correct and incorrect docking solutions [44], we can state that Glide found the correct binding mode of the ligands in four of the five cases analyzed. Despite compound 13 had the RMSD value above  $2.0 \text{ \AA}$ , it had the expected orientation in the thrombin binding site. The high RMSD value for compound 13 was due to a small displacement of  $P_2$  group, and a different orientation of the  $P_3$  group of the docked conformation with respect to the conformation in the crystallographic structure.

The remaining compounds were docked following the same docking protocol. The analysis of the docking poses obtained for all 177 ligands shows that all compounds adopt the same binding mode (Fig 3; mol2 files are in S1 File). This similar binding mode was expected since



**Fig 3. Alignment of all docked structures within the thrombin binding site (docked ligands are in green stick representation in A-E).** (A) Full view of the binding site. (B) Subsite S<sub>1</sub>, residues of the VSWGEGC motif are represented with white spheres, DACE motif is represented as an orange loop, Ser195 is represented with yellow sticks. (C) Subsite S<sub>1</sub>, residues of the DACE motif are represented with orange spheres, VSWGEGC motif is represented as a

white loop, Ser195 is represented with yellow sticks. **(D)** Subsite  $S_2$ , residues His57, Tyr61, and Trp64 are represented with violet spheres, and residues Trp215 and Gly216 from the VSWGEGC motif are represented with white spheres. **(E, F)** Subsite  $S_3$ , residues of the WRENL motif are represented with brown spheres, residues Trp215 and Glu217 from the VSWGEGC motif are represented with white spheres, and Ile174 is represented with yellow spheres. In several compounds that have a  $P_1$  moiety with two branches are represented: compounds with two hydrophobic branches are represented with green sticks, and compounds with one of the branches containing a polar group are represented with pink sticks.

doi:10.1371/journal.pone.0142774.g003

all compounds contain groups with suitable features to occupy  $S_1$ ,  $S_2$ , and  $S_3$  subsites of the thrombin binding site.

The  $S_1$  subsite is a large cavity bounded in one part by the residues contained in the sequence VSWGEGC (Val213-Ser214-Trp215-Gly216-Glu217-Gly218-Cys219). The backbone atoms of these residues and the side chain atoms of Val213 form a semicircle controlling the shape of the  $S_1$  subsite (Fig 3B). The other part of the  $S_1$  subsite is bounded by the residues contained in the sequence DACE (Asp189-Ala190-Cys191-Glu192), where Asp189 contributes with a negative charge, Cys191 forms a disulfide bond with Cys219 from the VSWGEGC motif, and the residues Ala190, Cys191, and Glu192 contribute with their backbone atoms (Fig 3C). The hydroxyl of the residue Ser195 and the  $\text{CH}_2$  of Gly226 also face the subsite  $S_1$ . In general, all compounds kept the electrostatic interactions between a HB donor group in  $P_1$  and the side chain carboxylate of Asp189. The backbone carbonyl groups of Gly218 and Ala190 are close to the site where interactions with Asp189 are formed; therefore they contribute with negative density to attract the  $P_1$  group (according to crystal structures previously reported, Gly218 can form HB with the  $P_1$  group). The  $S_1$  subsite can bound linear or cyclic groups that can adopt several conformations. The residue Ser195 (serine from the catalytic triad) has the side chain hydroxyl facing the entrance of  $S_1$ ; polar groups in this zone can interact with this residue (for instance, NH group in compounds **128**, **151**, and **165**).

The  $S_2$  subsite is delimited on one side by the residues Tyr61 and Trp64 from the YPPW insertion loop, and the His57 from the catalytic triad. On the other side  $S_2$  is delimited by the residues Trp215 and Gly216 from the motif VSWGEGC (Fig 3D). The  $S_2$  subsite was occupied by central aromatic group in all the docked compounds. The residues His57, Tyr61, and Trp64 form a small hole that is occupied by short substituents such as  $\text{CH}_3$ , CN, or halogens. When bigger substituents (for instance  $\text{OCH}_3$  in the low active compound **91**) are present, the ligand central aromatic ring changes its orientation to place the bulky substituent out of the hole. The residue Trp215 exposes  $\text{C}\alpha$  and  $\text{C}\beta$  ( $\text{CH}$  and  $\text{CH}_2$  groups) to the  $S_2$  subsite; therefore it contributes to establish hydrophobic interactions in the portal between  $S_2$  and  $S_1$ . Finally, the residue Gly216 exposes backbone CO and NH groups, and several TIs form HBs with them (for instance, the highly active compounds **164–177** that contain the 3-aminopyrazin-2(1*H*)-one scaffold form HB with both Gly258 backbone groups).

The  $S_3$  subsite is an ample pocket exposed to the solvent bounded in one part by the residues contained in the sequence WRENL (Trp95-Arg96-Glu97-Asn98-Leu99). The backbone atoms of these residues and the side chain atoms of Leu99 form a wall where three CO groups (from Trp95, Arg96, and Glu97) are facing the  $S_3$  subsite (Fig 3E and 3F). Very close to this wall, the side chain hydroxyl group of Tyr61 faces to the  $S_3$  subsite. The other part of the  $S_3$  subsite is delimited by the hydrophobic chemical groups of the side chains of the residues Ile174, Trp215, and Glu217. Interestingly,  $S_3$  subsite has a hydrophilic wall (backbone CO groups of Trp95, Arg96, and Glu97, and side chain hydroxyl group of Tyr61) and a hydrophobic wall (side chains of the residues Leu99, Ile174, Trp215, and Glu217). Docking results show that the  $S_3$  subsite tolerates the presence of different aromatic groups. Very hydrophobic  $P_3$  groups were located close to the hydrophobic wall and groups with some polarity were located close to the hydrophilic wall. Very large  $P_3$  groups were also oriented towards the hydrophobic wall because a bigger available space in this area. For compounds that have a  $P_1$  moiety with

**Table 2. Stepwise development of CoMSIA models by using SAMPLS and different field combinations.**

	NC	R <sup>2</sup>	s	F	Q <sup>2</sup>	s <sub>cv</sub>	Fraction of fields included in the model					
							Steric	Electrostatic	Hydrophobic	HB donor	HB acceptor	
CoMSIA-S	4	0.577	0.576	45.30	0.182	0.800	1					
CoMSIA-E	4	0.582	0.572	46.37	0.219	0.782		1				
CoMSIA-H	3	0.598	0.559	66.38	0.277	0.749			1			
CoMSIA-D	7	0.499	0.634	18.50	0.261	0.769					1	
CoMSIA-A	7	0.695	0.494	42.32	0.304	0.747						1
CoMSIA-SE	5	0.716	0.473	66.63	0.259	0.764	0.252	0.748				
CoMSIA-SH	3	0.635	0.533	77.63	0.295	0.740	0.250		0.750			
CoMSIA-SD	8	0.743	0.455	46.73	0.291	0.756	0.356				0.644	
CoMSIA-SA	7	0.814	0.386	81.33	0.508	0.628	0.288					0.712
CoMSIA-EH	2	0.486	0.634	62.20	0.319	0.725		0.568	0.432			
CoMSIA-ED	5	0.658	0.520	50.72	0.276	0.756		0.600			0.400	
CoMSIA-EA	7	0.825	0.375	87.33	0.410	0.687		0.519				0.481
CoMSIA-HD	3	0.600	0.558	66.99	0.322	0.726			0.543		0.457	
CoMSIA-HA	6	0.806	0.393	90.43	0.328	0.731			0.447			0.553
CoMSIA-SHD	5	0.741	0.452	75.63	0.345	0.719	0.168		0.448		0.384	
CoMSIA-SHA	6	0.825	0.373	102.82	0.395	0.693	0.155		0.355			0.490
<b>CoMSIA-SDA</b>	<b>8</b>	<b>0.871</b>	<b>0.322</b>	<b>109.20</b>	<b>0.543</b>	<b>0.610</b>	<b>0.186</b>				<b>0.350</b>	<b>0.464</b>
CoMSIA-SHDA	7	0.865	0.329	118.66	0.424	0.679	0.117		0.270		0.273	0.340
CoMSIA-ALL	7	0.878	0.312	133.87	0.439	0.670	0.085	0.262	0.198		0.201	0.254

NC is the number of components from PLS analysis; R<sup>2</sup> is the square of the correlation coefficient; S is the standard deviation of the regression; F is the Fischer ratio; Q<sup>2</sup> and S<sub>cv</sub> are the correlation coefficient and standard deviation of the leave-one-out (LOO) cross-validation, respectively. The best model is indicated in boldface.

doi:10.1371/journal.pone.0142774.t002

two branches (compounds 24–77) we found that branches are distributed according to polarity. If the two branches are hydrophobic, both groups are located close to the hydrophobic face (compounds in green in Fig 3E), but if one of the branches contains a polar group, it is located close to the hydrophilic face and the other branch is located close to the hydrophobic face (compounds in pink colour in Fig 3E).

### CoMSIA results

CoMSIA models were carried out to provide information about the structural features affecting the thrombin inhibitory activity of the compounds under study. The use of conformations aligned using the binding site (obtained by docking) allows identifying the relevant pharmacophoric features required to best match in the binding site [45]. In this sense, the best model reported below accounts for the desired features that characterize the most potent TIs.

Firstly, the CoMSIA models were developed by including one field, and then, these fields were combined and the statistical quality of hybrid models was analyzed by considering Q<sup>2</sup> values [46]. According to this analysis, we detected that compounds 5, 18, 33, 90, and 140 presented large residuals in all models; in this sense, they were excluded as outliers. Outliers are those compounds that have unexpected biological activities and are unable to fit in a QSAR model [47].

The list of CoMSIA models using different fields, without considering outlier compounds, is presented in Table 2. All the combinations were tested but only the more relevant ones are reported in Table 2 (models including one field, models including two fields with Q<sup>2</sup> > 0.2,

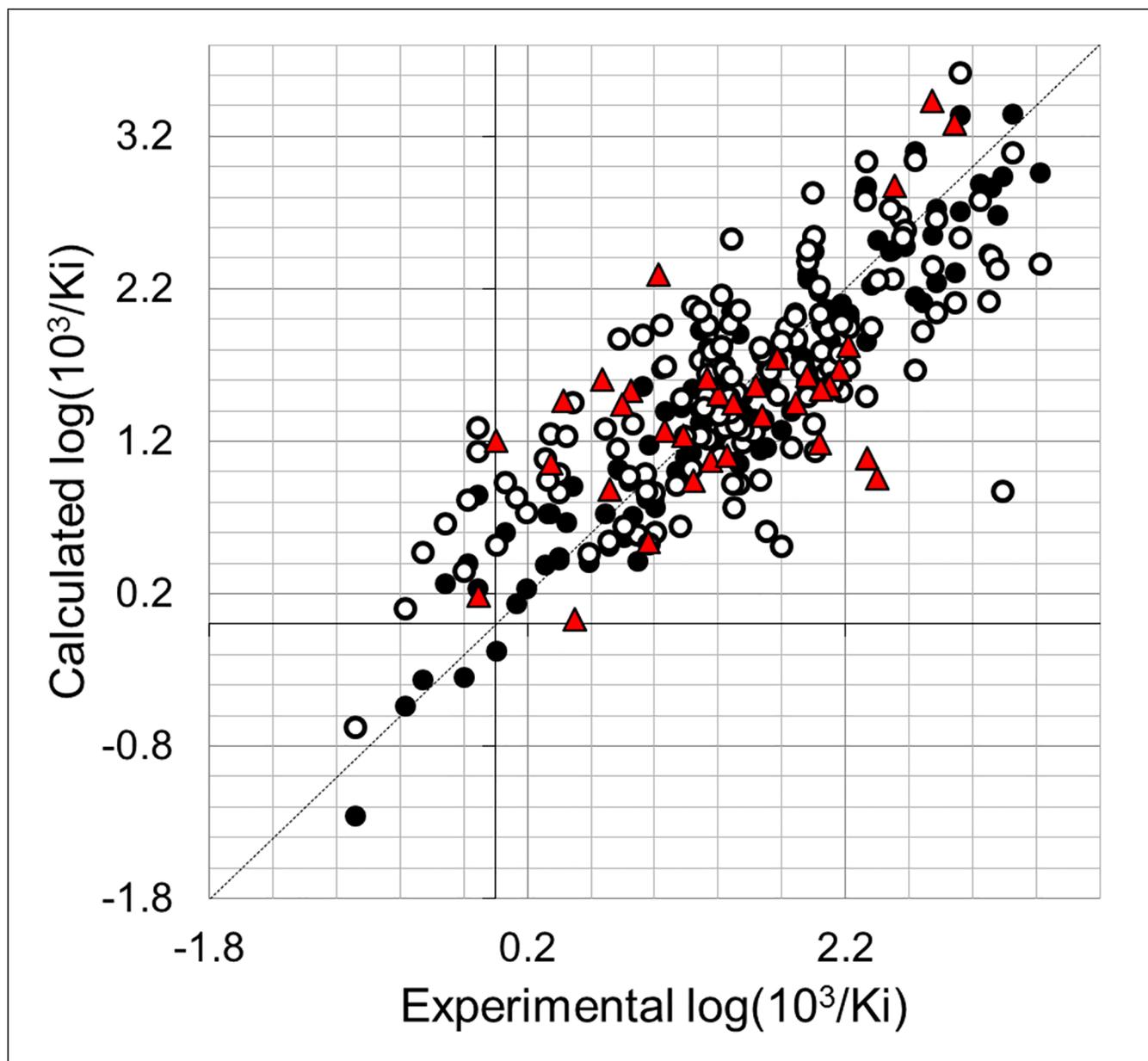
models including three or more fields with  $Q^2 > 0.3$ . The internal predictability of the models was the criterion that was used to select the best QSAR model. We observed that CoMSIA models using one field were statistically unacceptable ( $Q^2 < 0.5$ ). The analysis of the hybrid models yielded the model CoMSIA-SDA with the best  $Q^2$  value of 0.543. This model combines steric, HB donor, and HB acceptor fields. We included other fields and we observed that it does not produce an improvement in the internal validation of the model CoMSIA-SDA, since models including more fields had lower  $Q^2$  values (Table 2). The model CoMSIA-SDA was derived by using eight components and showed contributions of the steric field of 18.6%, HB donor field of 35.0%, and the HB acceptor field of 46.4%. In addition, it explains 87.1% of the variance, has a low standard deviation ( $s = 0.322$ ), and a high Fischer ratio ( $F = 109.20$ ). The predictions of  $\log(10^3/K_i)$  values for the 138 TIs from the training set using model CoMSIA-SDA are shown in Table 1. The correlation between the calculated and experimental values of  $\log(10^3/K_i)$  (from training and LOO cross-validation) is shown in Fig 4. Plots of the LOO cross-validation predictions reveal that the proposed model is able to discriminate between the most active and the less active compounds. Considering that the CoMSIA approach was applied to a big dataset, we consider that the value of  $Q^2 = 0.543$  reflects that there is no redundancy in the training set since each member of the training set is important for the model [48].

We also used model CoMSIA-SDA to predict the TI inhibitory activities of the test set compounds. The values are given in Table 1 and correlation between the calculated and experimental values are also represented in Fig 4. This analysis reveals that the proposed model also predicted adequately all the compounds in the test set.

The contour plots of the CoMSIA steric, HB donor, and HB acceptor fields are presented in Fig 5 for the best model CoMSIA-SDA. To aid in visualization, compound 134 is displayed in the maps with a superposition of the contour plots on active-site residues. The representation of contour plots and the thrombin active site in the same 3D space helps in delineating the relationship between the relevant chemical properties of the more active compounds and the chemical groups of the residues in the active site.

The colored isopleths in the map represent the 3D space where the structural properties changes are related to the changes in thrombin inhibitory potency. Green and yellow isopleths in Fig 5A indicate regions where bulky groups increase and decrease the inhibitory activity, respectively. A large region of yellow contour at  $S_1$  subsite suggests that bulky groups are not desired in this zone. This feature indicates that linear groups should be prioritized rather than branched groups in  $S_1$ . Other yellow isopleths at  $S_3$  subsite near the residues Ile174 and Tyr61 and green isopleths near the residues Glu217 and Leu99 indicate that bulky groups extended to the more hydrophilic part of  $S_3$  decrease the inhibitory activity, but groups extended to the hydrophobic part of  $S_3$  increase the inhibitory activity. The predictions of chemical features of substituents in  $S_3$  according to the analysis of steric fields included in CoMSIA-SDA model are difficult since it is difficult to foresee their occupation and interactions in this wide cavity.

Cyan and purple isopleths in Fig 5B are in regions where HB donor groups favor and disfavor the activity, respectively. On the other hand, in Fig 5C magenta isopleths indicate regions where HB acceptors enhance the activity, and red isopleths indicate regions where HB acceptors decrease the activity. Interestingly, all the cyan and purple isopleths are located at  $S_1$  subsite. The cyan and purple isopleths near Asp189, Gly218, Gly226, and Tyr228 represent 3D spaces that the model identified that are affected by the presence of HB donor groups in the deeper region of  $S_1$ . The purple isopleth near Ser214, the red isopleth near Gly216, and the magenta isopleth near Cys191 and Glu192 indicate that the entrance to  $S_1$  is affected by the presence of HB donor and acceptor groups. Two red isopleths in  $S_3$  indicate that HB acceptor groups are not necessary in this zone. The first one is located near Asn98 and the other is



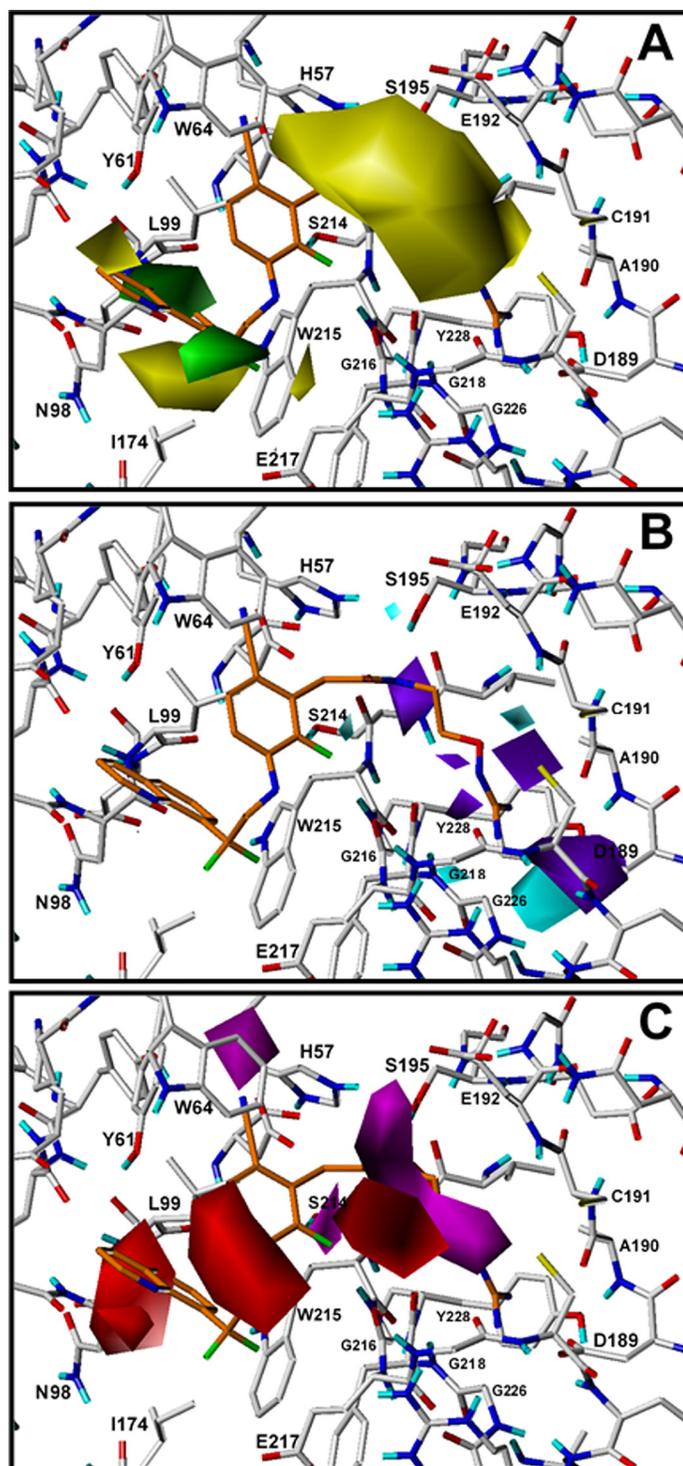
**Fig 4. Scatter plot of the experimental activities versus predicted activities for model CoMFA-SDA: (●) training set predictions (○) LOO cross-validated predictions (red ▲) test set predictions.**

doi:10.1371/journal.pone.0142774.g004

exposed to solvent between Glu217 and Trp64. The magenta isopleth in  $S_2$  near His57 indicates that HB acceptor groups such as CN are desired in this zone.

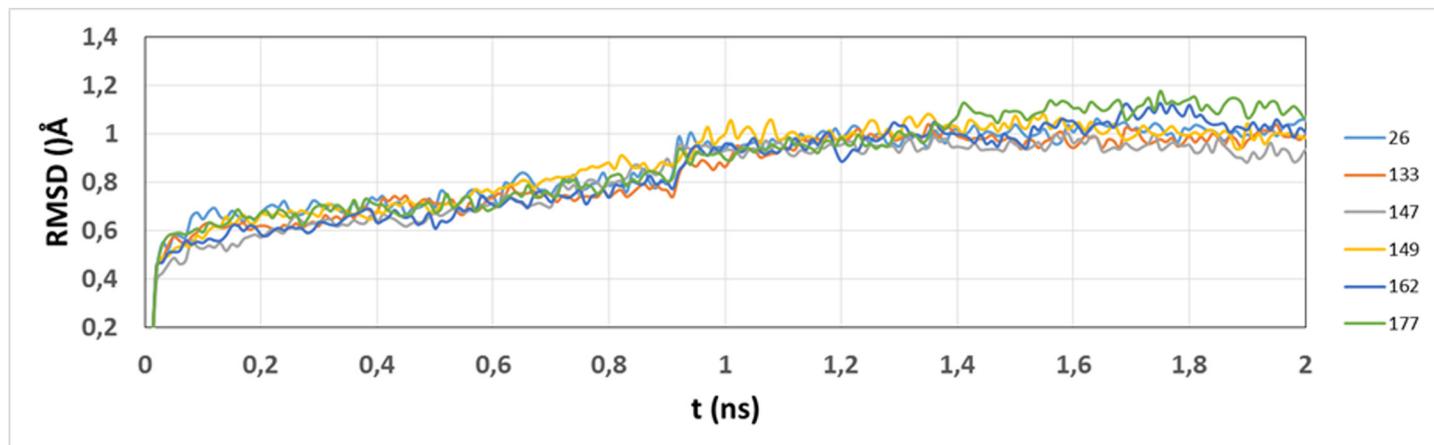
### Free energy calculations results

Compounds **26**, **133**, **147**, **149**, **162**, and **177** were selected as a sample for deriving free energy calculations using MM-GBSA method. The study includes a previous conformational sampling using MD simulations to consider averaged properties and to get more realistic conditions. Stability of the MD trajectories using the RMSD of the positions for all the protein atoms as a



**Fig 5. CoMSIA contour maps for TIs deriving from model CoMSIA-SDA.** The amino acid residues located close to the binding pocket of thrombin are represented for comparing their position with the position of isopleths derived from the model. Compound **134** is shown inside the field. **(A)** Steric field: green isopleths indicate regions where bulky groups enhance the activity and yellow isopleths indicate regions where bulky groups disfavor the activity. **(B)** HB donor field: cyan isopleths indicate regions where HB donors favor the activity, and purple isopleths indicate regions where HB donors disfavor the activity. **(C)** HB acceptor fields: magenta isopleths indicate regions where HB acceptors enhance the activity, and red isopleths indicate regions where HB acceptors decrease the activity.

doi:10.1371/journal.pone.0142774.g005



**Fig 6. Time dependence of the RMSD for protein atoms from starting structures during equilibration process.** RMSD for the studied systems are represented in colors indicated at the right.

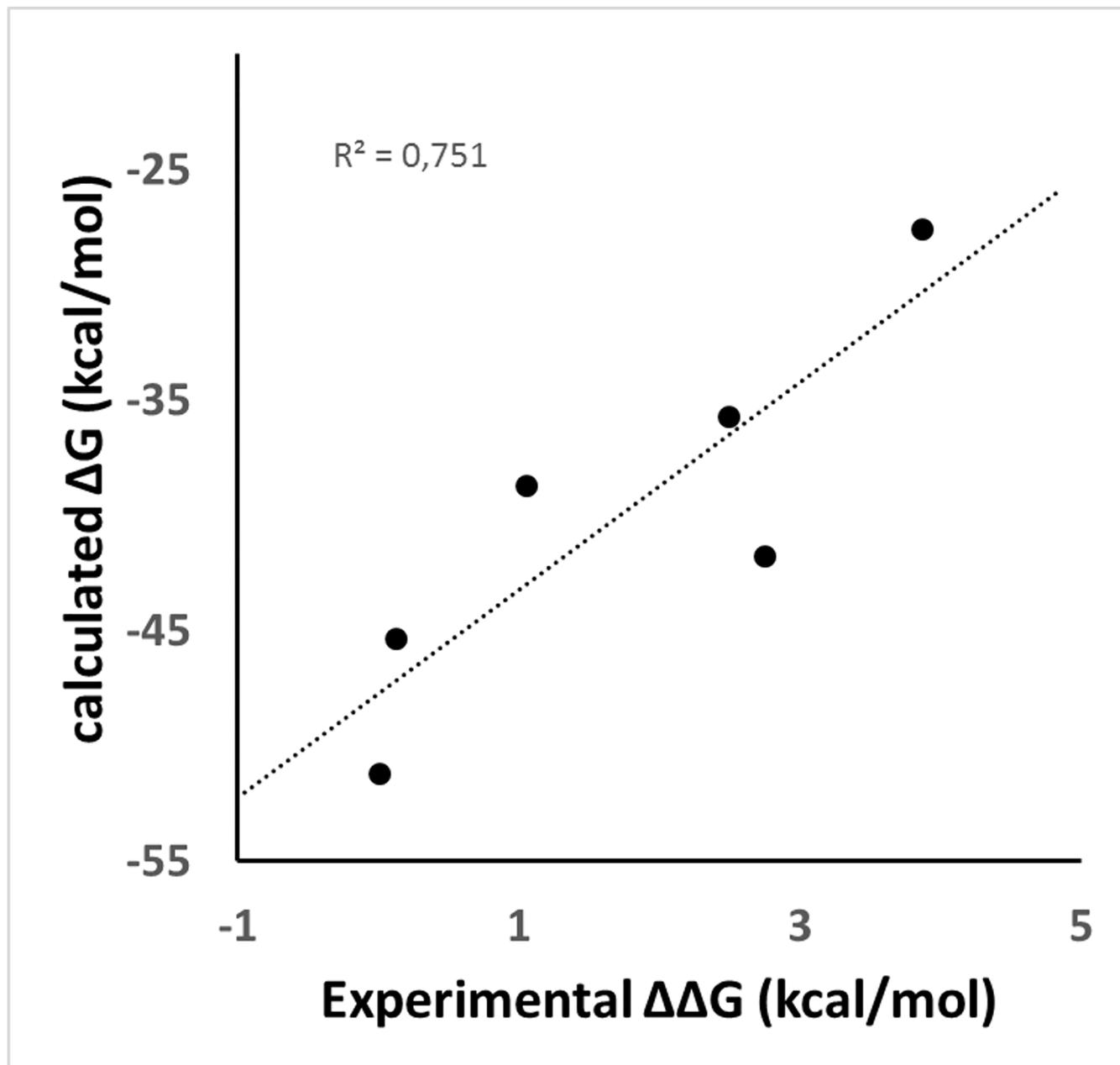
doi:10.1371/journal.pone.0142774.g006

function of simulation time was evaluated; RMSD was almost constant after the equilibration process (2.0 ns) for all the systems (Fig 6).

The above mentioned compounds were selected from data set having in mind the differences in their activities against thrombin and their differences in the substitution and structural characteristics of the groups located at  $S_1$ ,  $S_2$  and  $S_3$ . In this sense, MM-GBSA paradigm allowed finding the role of physico-chemical terms in the binding potency when different groups located at  $S_1$ ,  $S_2$ , and  $S_3$  are changed. Guanidinoxy groups as substituent of N-ethylacetamide or n-propanol, and 2-amino-6-methylpyridin-5-yl groups were considered in  $S_1$ . Methyl, Cl, F, CN, and oxo groups were considered in the aromatic ring in  $S_2$ . 2,2-Difluoro-2-arylethyl, (1-arylcyclopropyl)methyl, arylethyl, cyclopentyl groups were considered as chemical features in  $S_3$ .

MD allows analyzing the effect of the water media, the movement of the ligand inside the binding site, the movement of the residues in the binding site, and the stability of HBs. In general, from the analysis of all the simulations, it is possible to identify different water characteristics in the pockets  $S_1$ ,  $S_2$ , and  $S_3$ . There are many water molecules in  $S_3$  which is the more exposed site to the water media, most of these molecules move freely and the other ones have occasional interactions with polar groups forming HBs. On the other hand, there are no water molecules in  $S_2$  during MD simulations. Finally, there are some water molecules at  $S_1$  trapped in HB networks formed between the ligands and the enzyme; these water molecules have a very limited movement. Interestingly, it is easy to differentiate  $S_1$ ,  $S_2$ , and  $S_3$  pockets according to the distinct water dynamical behavior, and this is highly related to the TI characteristics in each pocket disclosed by docking analysis. The dynamic of the studied ligands using MD simulations revealed that groups at  $S_1$  form stable HB interactions, groups at  $S_2$  are anchored in a hydrophobic cage, and groups at  $S_3$  have a free movement.

The MM-GBSA free energy calculations were employed to get quantitative estimates for the binding free energies of the inhibitors inside the thrombin active site. We found a high correlation ( $R^2 = 0.751$ , Fig 7) between calculated  $\Delta G$  values and the experimental ones (the experimental values were expressed as  $\Delta\Delta G$  values with respect to the most active selected compound 133). However, only six compounds were considered; due to reduced data, we cannot conclude that MM-GBSA was able to explain the trend of the modeled TIs, but we can analyze the free energy components such as van der Waals (VDW), electrostatic, and solvation contributions to give detailed molecular information about the selected systems. The purpose is to evaluate



**Fig 7. Correlation between experimental  $\Delta\Delta G$  (calculated from  $K_i$  values of compounds) and calculated  $\Delta G$  values using MM-GBSA.** Experimental  $\Delta G$  was calculated from  $K_i$  values of compounds, and  $\Delta\Delta G$  was calculated using  $\Delta G$  of compound **133** as reference.

doi:10.1371/journal.pone.0142774.g007

the role of the physico-chemical features when different substituents are presented. The results of the predicted  $\Delta G$  values and the physico-chemical components  $\Delta E_{vdw}$ ,  $\Delta E_{elect}$ , and  $\Delta G_{solv}$  for the complexes were summarized in [Table 3](#). To get a better view on which energy terms have more impact on the inhibitory potency, these individual energy components were compared. From [Table 3](#), it can be seen that  $\Delta E_{vdw}$  has the major favorable contribution to the total free energy, but there is no big difference among this value for different complexes. In this sense, the VDW energy term is not primarily responsible for differentiating the binding affinity of the selected TIs. However, some of the most negative  $\Delta E_{vdw}$  values were obtained for the most

**Table 3. Predicted MM-GBSA free energies (kcal/mol) and individual energy terms of the thrombin-inhibitor complexes for selected compounds.**

compound	$\Delta G_{\text{calc}}$	$\Delta G_{\text{vdw}}$	$\Delta G_{\text{elect}}$	$\Delta G_{\text{solv}}$	$\Delta \Delta G_{\text{exp}}^a$
26	-28.06 ± 0.10	-47.05 ± 0.10	24.65 ± 0.08	-5.65 ± 0.28	3.86
133	-62.74 ± 0.20	-51.38 ± 0.19	-15.33 ± 0.19	3.97 ± 0.32	0.00
147	-41.58 ± 0.13	-46.91 ± 0.13	10.80 ± 0.14	-5.47 ± 0.41	1.05
149	-45.05 ± 0.14	-59.34 ± 0.11	19.91 ± 0.14	-5.63 ± 0.39	0.12
162	-41.90 ± 0.22	-40.90 ± 0.23	4.58 ± 0.25	-5.58 ± 0.69	2.74
177	-35.74 ± 0.22	-52.54 ± 0.14	23.02 ± 0.26	-6.22 ± 0.62	2.49

a Experimental  $\Delta \Delta G$  was calculated using  $\Delta G$  of compound **133** as reference.

doi:10.1371/journal.pone.0142774.t003

active compounds **133** and **149**. The solvation term also has the same effect in most of the complexes with a favorable contribution to the global free energy. Strangely, this term is positive for the most active compound **133**. On the contrary, the electrostatic term was only favorable for compound **133**, and the worst contribution of this term was for the less active compound **26**.

The analysis of the MM-GBSA terms indicates that not any physico-chemical property has a preponderant role in the potency of TIs. It is expected that this behavior could be maintained if we consider the remaining compounds considering the complexity of the thrombin binding site in  $S_1$ ,  $S_2$ , and  $S_3$  pockets. We previously observed that these pockets have very different characteristics; therefore, each of them could be modulated considering different physico-chemical features. However, the binding affinity can be the result of the combination of the particular effect of these features in each pocket.

## Conclusions

In this work, we report a theoretical study on drug design area for thrombin inhibitors (TIs). We selected a data set of 177 compounds with diverse groups at motifs  $P_1$ ,  $P_2$ , and  $P_3$ , and explored the 3D positioning of them inside the thrombin active site pockets  $S_1$ ,  $S_2$ , and  $S_3$  by docking experiments. Our approach reproduced the previously reported position of compounds **13**, **24**, **128**, **151**, and **165** inside thrombin's active site. The remaining compounds were oriented in a similar manner. The interactions established for all compounds within thrombin binding site were carefully described.

Additionally, predictive QSAR models were built by using CoMSIA method. We found that CoMSIA-SDA including 138 compounds in the training set was the best among the explored models. This model included steric, HB donor, and HB acceptor fields. It was validated by LOO cross-validation and then it was used in the prediction of an external set which contains 34 compounds that were not included during the training process. The CoMSIA model could be used to predict novel candidates; moreover, the interpretation of the CoMSIA fields makes it possible to draw conclusions concerning the most appropriate features for novel analogues.

Finally, six compounds were selected and were subjected to MD and free energy calculations using MM-GBSA method. The dynamical behavior of the systems were revealed by MD; it was interesting to note that water molecules have completely different dynamical behavior in pockets  $S_1$ ,  $S_2$ , and  $S_3$ , which is related to the desired characteristics of motifs  $P_1$ ,  $P_2$ , and  $P_3$  of TIs. The free energy calculations using MM-GBSA allowed studying the role of physico-chemical parameters such as van der Waals interactions, electrostatic interactions, and solvation. It was found that not any component was preponderant in the study of the differential potency of inhibitors.

## Supporting Information

**S1 File. Docking poses obtained for the TIs in mol2 format.**  
(ZIP)

## Author Contributions

Conceived and designed the experiments: JC. Performed the experiments: KMU. Analyzed the data: JC WT. Contributed reagents/materials/analysis tools: JC. Wrote the paper: JC.

## References

1. Qaseem A, Chou R, Humphrey LL, Starkey M, Shekelle P. Venous Thromboembolism Prophylaxis in Hospitalized Patients: A Clinical Practice Guideline From the American College of Physicians. *Ann Intern Med.* 2011; 155: 625–632. doi: [10.1059/0003-4819-155-9-2011111010-00011](https://doi.org/10.1059/0003-4819-155-9-2011111010-00011) PMID: [22041951](https://pubmed.ncbi.nlm.nih.gov/22041951/)
2. Melnikova I. The anticoagulants market. *Nat Rev Drug Discov.* 2009; 8: 353–353. doi: [10.1038/nrd2851](https://doi.org/10.1038/nrd2851) PMID: [19390569](https://pubmed.ncbi.nlm.nih.gov/19390569/)
3. Tanaka KA, Key NS, Levy JH. Blood coagulation: hemostasis and thrombin regulation. *Anesth Analg.* 2009; 108: 1433–1446. doi: [10.1213/ane.0b013e31819bcc9c](https://doi.org/10.1213/ane.0b013e31819bcc9c) PMID: [19372317](https://pubmed.ncbi.nlm.nih.gov/19372317/)
4. Coppens M, Eikelboom JW, Gustafsson D, Weitz JI, Hirsh J. Translational success stories: development of direct thrombin inhibitors. *Circ Res.* 2012; 111: 920–929. doi: [10.1161/CIRCRESAHA.112.264903](https://doi.org/10.1161/CIRCRESAHA.112.264903) PMID: [22982873](https://pubmed.ncbi.nlm.nih.gov/22982873/)
5. Tomczuk B, Lu T, Soll RM, Fedde C, Wang A, Murphy L, et al. Oxyguanidines: application to non-peptidic phenyl-based thrombin inhibitors. *Bioorg Med Chem Lett.* 2003; 13: 1495–1498. PMID: [12668020](https://pubmed.ncbi.nlm.nih.gov/12668020/)
6. Lu T, Markotan T, Coppo F, Tomczuk B, Crysler C, Eisennagel S, et al. Oxyguanidines. Part 2: Discovery of a novel orally active thrombin inhibitor through structure-based drug design and parallel synthesis. *Bioorg Med Chem Lett.* 2004; 14: 3727–3731. doi: [10.1016/j.bmcl.2004.05.002](https://doi.org/10.1016/j.bmcl.2004.05.002) PMID: [15203151](https://pubmed.ncbi.nlm.nih.gov/15203151/)
7. Soll RM, Lu T, Tomczuk B, Illig CR, Fedde C, Eisennagel S, et al. Amidinohydrazones as guanidine bioisosteres: application to a new class of potent, selective and orally bioavailable, non-amide-based small-molecule thrombin inhibitors. *Bioorg Med Chem Lett.* 2000; 10: 1–4. PMID: [10636229](https://pubmed.ncbi.nlm.nih.gov/10636229/)
8. Lu T, Soll RM, Illig CR, Bone R, Murphy L, Spurlino J, et al. Structure-activity and crystallographic analysis of a new class of non-amide-based thrombin inhibitor. *Bioorg Med Chem Lett.* 2000; 10: 79–82. PMID: [10636249](https://pubmed.ncbi.nlm.nih.gov/10636249/)
9. Lu T, Tomczuk B, Bone R, Murphy L, Salemme FR, Soll RM. Non-peptidic phenyl-based thrombin inhibitors: exploring structural requirements of the S1 specificity pocket with amidines. *Bioorg Med Chem Lett.* 2000; 10: 83–85. PMID: [10636250](https://pubmed.ncbi.nlm.nih.gov/10636250/)
10. Lu T, Tomczuk B, Illig CR, Bone R, Murphy L, Spurlino J, et al. In vitro evaluation and crystallographic analysis of a new class of selective, non-amide-based thrombin inhibitors. *Bioorg Med Chem Lett.* 1998; 8: 1595–1600. PMID: [9873397](https://pubmed.ncbi.nlm.nih.gov/9873397/)
11. Kreutter KD, Lu T, Lee L, Giardino EC, Patel S, Huang H, et al. Orally efficacious thrombin inhibitors with cyanofluorophenylacetamide as the P2 motif. *Bioorg Med Chem Lett.* 2008; 18: 2865–2870. doi: [10.1016/j.bmcl.2008.03.087](https://doi.org/10.1016/j.bmcl.2008.03.087) PMID: [18420408](https://pubmed.ncbi.nlm.nih.gov/18420408/)
12. Lee L, Kreutter KD, Pan W, Crysler C, Spurlino J, Player MR, et al. 2-(2-Chloro-6-fluorophenyl)acetamides as potent thrombin inhibitors. *Bioorg Med Chem Lett.* 2007; 17: 6266–6269. doi: [10.1016/j.bmcl.2007.09.013](https://doi.org/10.1016/j.bmcl.2007.09.013) PMID: [17889527](https://pubmed.ncbi.nlm.nih.gov/17889527/)
13. Lu T, Markotan T, Ballentine SK, Giardino EC, Spurlino J, Crysler CS, et al. Discovery and clinical evaluation of 1-[N-[2-(amidinoaminoxy)ethyl]amino]carbonylmethyl-6-methyl-3-[2,2-difluoro-2-phenylethylamino]pyrazinone (RWJ-671818), a thrombin inhibitor with an oxyguanidine P1 motif. *J Med Chem.* 2010; 53: 1843–1856. doi: [10.1021/jm901802n](https://doi.org/10.1021/jm901802n) PMID: [20102150](https://pubmed.ncbi.nlm.nih.gov/20102150/)
14. Kontogiorgis CA, Hadjipavlou-Litina D. Quantitative Structure—Activity Relationships (QSARs) of Thrombin Inhibitors: Review, Evaluation and Comparative Analysis. *Current Medicinal Chemistry.* 2003; 10: 525–577. doi: [10.2174/0929867033457935](https://doi.org/10.2174/0929867033457935) PMID: [12678789](https://pubmed.ncbi.nlm.nih.gov/12678789/)
15. Böhm M, Stürzebecher J, Klebe G. Three-Dimensional Quantitative Structure–Activity Relationship Analyses Using Comparative Molecular Field Analysis and Comparative Molecular Similarity Indices Analysis To Elucidate Selectivity Differences of Inhibitors Binding to Trypsin, Thrombin, and Factor Xa. *J Med Chem.* 1999; 42: 458–477. doi: [10.1021/jm981062r](https://doi.org/10.1021/jm981062r) PMID: [9986717](https://pubmed.ncbi.nlm.nih.gov/9986717/)
16. Li M-L, Ren Y-J, Dong M-H, Ren W-X. Design, synthesis and structural exploration of novel fluorinated dabigatran derivatives as direct thrombin inhibitors. *Eur J Med Chem.* 2015; 96: 122–138. doi: [10.1016/j.ejmech.2015.04.012](https://doi.org/10.1016/j.ejmech.2015.04.012) PMID: [25874337](https://pubmed.ncbi.nlm.nih.gov/25874337/)

17. Deswal S, Roy N. Quantitative structure activity relationship studies of aryl heterocycle-based thrombin inhibitors. *Eur J Med Chem.* 2006; 41: 1339–1346. doi: [10.1016/j.ejmech.2006.07.001](https://doi.org/10.1016/j.ejmech.2006.07.001) PMID: [16884829](https://pubmed.ncbi.nlm.nih.gov/16884829/)
18. Ramírez-Galicia G, Garduño-Juárez R, Correa-Basurto J, Deeb O. Exploring QSARs for inhibitory effect of a set of heterocyclic thrombin inhibitors by multilinear regression refined by artificial neural network and molecular docking simulations. *Journal of Enzyme Inhibition and Medicinal Chemistry.* 2012; 27: 174–186. doi: [10.3109/14756366.2011.582467](https://doi.org/10.3109/14756366.2011.582467) PMID: [21635208](https://pubmed.ncbi.nlm.nih.gov/21635208/)
19. Muley L, Baum B, Smolinski M, Freindorf M, Heine A, Klebe G, et al. Enhancement of hydrophobic interactions and hydrogen bond strength by cooperativity: synthesis, modeling, and molecular dynamics simulations of a congeneric series of thrombin inhibitors. *J Med Chem.* 2010; 53: 2126–2135. doi: [10.1021/jm9016416](https://doi.org/10.1021/jm9016416) PMID: [20148533](https://pubmed.ncbi.nlm.nih.gov/20148533/)
20. Silva V de A, Cargnelutti MT, Giesel GM, Palmieri LC, Monteiro RQ, Verli H, et al. Structure and Behavior of Human  $\alpha$ -Thrombin upon Ligand Recognition: Thermodynamic and Molecular Dynamics Studies. *PLoS ONE.* 2011; 6: e24735. doi: [10.1371/journal.pone.0024735](https://doi.org/10.1371/journal.pone.0024735) PMID: [21935446](https://pubmed.ncbi.nlm.nih.gov/21935446/)
21. de Candia M, Fiorella F, Lopopolo G, Carotti A, Romano MR, Lograno MD, et al. Synthesis and biological evaluation of direct thrombin inhibitors bearing 4-(piperidin-1-yl)pyridine at the P1 position with potent anticoagulant activity. *J Med Chem.* 2013; 56: 8696–8711. doi: [10.1021/jm401169a](https://doi.org/10.1021/jm401169a) PMID: [24102612](https://pubmed.ncbi.nlm.nih.gov/24102612/)
22. Nilsson M, Hämäläinen M, Ivarsson M, Gottfries J, Xue Y, Hansson S, et al. Compounds Binding to the S2–S3 Pockets of Thrombin. *J Med Chem.* 2009; 52: 2708–2715. doi: [10.1021/jm8011849](https://doi.org/10.1021/jm8011849) PMID: [19371038](https://pubmed.ncbi.nlm.nih.gov/19371038/)
23. Bhunia SS, Roy KK, Saxena AK. Profiling the Structural Determinants for the Selectivity of Representative Factor-Xa and Thrombin Inhibitors Using Combined Ligand-Based and Structure-Based Approaches. *J Chem Inf Model.* 2011; 51: 1966–1985. doi: [10.1021/ci200185q](https://doi.org/10.1021/ci200185q) PMID: [21761917](https://pubmed.ncbi.nlm.nih.gov/21761917/)
24. Maestro, version 9.0, 2007. New York, NY.: Schrödinger, LLC;
25. Caballero J, Vergara-Jaque A, Fernández M, Coll D. Docking and quantitative structure–activity relationship studies for sulfonyl hydrazides as inhibitors of cytosolic human branched-chain amino acid aminotransferase. *Mol Divers.* 2009; 13: 493–500. doi: [10.1007/s11030-009-9140-1](https://doi.org/10.1007/s11030-009-9140-1) PMID: [19350404](https://pubmed.ncbi.nlm.nih.gov/19350404/)
26. Caballero J, Quiliano M, Alzate-Morales JH, Zimic M, Deharo E. Docking and quantitative structure–activity relationship studies for 3-fluoro-4-(pyrrolo[2,1-f][1,2,4]triazin-4-yloxy)aniline, 3-fluoro-4-(1H-pyrrolo[2,3-b]pyridin-4-yloxy)aniline, and 4-(4-amino-2-fluorophenoxy)-2-pyridinylamine derivatives as c-Met kinase inhibitors. *J Comput-Aided Mol Des.* 2011; 25: 349–369. doi: [10.1007/s10822-011-9425-1](https://doi.org/10.1007/s10822-011-9425-1) PMID: [21487786](https://pubmed.ncbi.nlm.nih.gov/21487786/)
27. Fuentes E, Badimon L, Caballero J, Padró T, Vilahur G, Alarcón M, et al. Protective mechanisms of adenosine 5'-monophosphate in platelet activation and thrombus formation: *Thromb Haemost.* 2014; 111: 491–507. doi: [10.1160/TH13-05-0386](https://doi.org/10.1160/TH13-05-0386) PMID: [24306059](https://pubmed.ncbi.nlm.nih.gov/24306059/)
28. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, 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. doi: [10.1021/jm0306430](https://doi.org/10.1021/jm0306430) PMID: [15027865](https://pubmed.ncbi.nlm.nih.gov/15027865/)
29. Ligprep, version 2.5, 2011. New York, NY.: Schrödinger, LLC;
30. Munoz C, Adasme F, Alzate-Morales JH, Vergara-Jaque A, Kniess T, Caballero J. Study of differences in the VEGFR2 inhibitory activities between semaxanib and SU5205 using 3D-QSAR, docking, and molecular dynamics simulations. *J Mol Graph Model.* 2012; 32: 39–48. doi: [10.1016/j.jmgm.2011.10.005](https://doi.org/10.1016/j.jmgm.2011.10.005) PMID: [22070999](https://pubmed.ncbi.nlm.nih.gov/22070999/)
31. SYBYL, version 7.3. 1699 South Hanley Rd., St. Louis, MO 63144, USA.: Tripos Inc.; 2006.
32. Bush BL, Nachbar RB. Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA. *J Comput-Aided Mol Des.* 1993; 7: 587–619. doi: [10.1007/BF00124364](https://doi.org/10.1007/BF00124364) PMID: [8294948](https://pubmed.ncbi.nlm.nih.gov/8294948/)
33. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comp Chem.* 2005; 26: 1781–1802. doi: [10.1002/jcc.20289](https://doi.org/10.1002/jcc.20289)
34. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J Phys Chem B.* 1998; 102: 3586–3616. doi: [10.1021/jp973084f](https://doi.org/10.1021/jp973084f) PMID: [24889800](https://pubmed.ncbi.nlm.nih.gov/24889800/)
35. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem.* 2010; 31: 671–690. doi: [10.1002/jcc.21367](https://doi.org/10.1002/jcc.21367) PMID: [19575467](https://pubmed.ncbi.nlm.nih.gov/19575467/)
36. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* 1983; 79: 926–935. doi: [10.1063/1.445869](https://doi.org/10.1063/1.445869)

37. Steinbach PJ, Brooks BR. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J Comput Chem*. 1994; 15: 667–683. doi: [10.1002/jcc.540150702](https://doi.org/10.1002/jcc.540150702)
38. Massova I, Kollman PA. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspectives in Drug Discovery and Design*. 2000; 18: 113–135. doi: [10.1023/A:1008763014207](https://doi.org/10.1023/A:1008763014207)
39. Homeyer N, Gohlke H. Free Energy Calculations by the Molecular Mechanics Poisson–Boltzmann Surface Area Method. *Mol Inf*. 2012; 31: 114–122. doi: [10.1002/minf.201100135](https://doi.org/10.1002/minf.201100135)
40. Adasme-Carreño F, Muñoz-Gutierrez C, Caballero J, Alzate-Morales J. Performance of The MM/GBSA Scoring Using a Binding Site Hydrogen Bond Network-Based Frame Selection: The Protein Kinase Case. *Phys Chem Chem Phys*. 2014; 16: 14047–14058. doi: [10.1039/C4CP01378F](https://doi.org/10.1039/C4CP01378F) PMID: [24901037](https://pubmed.ncbi.nlm.nih.gov/24901037/)
41. Mena-Ulecia K, Vergara-Jaque A, Poblete H, Tiznado W, Caballero J. Study of the Affinity between the Protein Kinase PKA and Peptide Substrates Derived from Kempptide Using Molecular Dynamics Simulations and MM/GBSA. *PLoS ONE*. 2014; 9: e109639. doi: [10.1371/journal.pone.0109639](https://doi.org/10.1371/journal.pone.0109639) PMID: [25275314](https://pubmed.ncbi.nlm.nih.gov/25275314/)
42. Hou T, Yu R. Molecular Dynamics and Free Energy Studies on the Wild-type and Double Mutant HIV-1 Protease Complexed with Amprenavir and Two Amprenavir-Related Inhibitors: Mechanism for Binding and Drug Resistance. *J Med Chem*. 2007; 50: 1177–1188. doi: [10.1021/jm0609162](https://doi.org/10.1021/jm0609162) PMID: [17300185](https://pubmed.ncbi.nlm.nih.gov/17300185/)
43. Wang W, Kollman PA. Computational study of protein specificity: The molecular basis of HIV-1 protease drug resistance. *PNAS*. 2001; 98: 14937–14942. doi: [10.1073/pnas.251265598](https://doi.org/10.1073/pnas.251265598) PMID: [11752442](https://pubmed.ncbi.nlm.nih.gov/11752442/)
44. Gohlke H, Hendlich M, Klebe G. Knowledge-based scoring function to predict protein-ligand interactions. *J Mol Biol*. 2000; 295: 337–356. doi: [10.1006/jmbi.1999.3371](https://doi.org/10.1006/jmbi.1999.3371) PMID: [10623530](https://pubmed.ncbi.nlm.nih.gov/10623530/)
45. Caballero J, Zilocchi S, Tiznado W, Collina S, Rossi D. Binding studies and quantitative structure-activity relationship of 3-amino-1H-indazoles as inhibitors of GSK3 $\beta$ . *Chem Biol Drug Des*. 2011; 78: 631–641. doi: [10.1111/j.1747-0285.2011.01186.x](https://doi.org/10.1111/j.1747-0285.2011.01186.x) PMID: [21756288](https://pubmed.ncbi.nlm.nih.gov/21756288/)
46. Caballero J, Fernández M, González-Nilo FD. Structural requirements of pyrido[2,3-d]pyrimidin-7-one as CDK4/D inhibitors: 2D autocorrelation, CoMFA and CoMSIA analyses. *Bioorg Med Chem*. 2008; 16: 6103–6115. doi: [10.1016/j.bmc.2008.04.048](https://doi.org/10.1016/j.bmc.2008.04.048) PMID: [18468903](https://pubmed.ncbi.nlm.nih.gov/18468903/)
47. Verma RP, Hansch C. An approach toward the problem of outliers in QSAR. *Bioorg Med Chem*. 2005; 13: 4597–4621. doi: [10.1016/j.bmc.2005.05.002](https://doi.org/10.1016/j.bmc.2005.05.002) PMID: [15951183](https://pubmed.ncbi.nlm.nih.gov/15951183/)
48. Dowejko AM. 3D-QSAR illusions. *J Comput Aided Mol Des*. 2004; 18: 587–596. PMID: [15729857](https://pubmed.ncbi.nlm.nih.gov/15729857/)