



# Insights into the Interactions between Maleimide Derivates and GSK3 $\beta$ Combining Molecular Docking and QSAR

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

**1** Centro de Bioinformática y Simulación Molecular, Facultad de Ingeniería, Universidad de Talca, Talca, Chile, **2** Departamento de Ciencias Químicas, Facultad de Ciencias Exactas, Universidad Andres Bello, Santiago de Chile, Chile

## Abstract

Many protein kinase (PK) inhibitors have been reported in recent years, but only a few have been approved for clinical use. The understanding of the available molecular information using computational tools is an alternative to contribute to this process. With this in mind, we studied the binding modes of 77 maleimide derivates inside the PK glycogen synthase kinase 3 beta (GSK3 $\beta$ ) using docking experiments. We found that the orientations that these compounds adopt inside GSK3 $\beta$  binding site prioritize the formation of hydrogen bond (HB) interactions between the maleimide group and the residues at the hinge region (residues Val135 and Asp133), and adopt propeller-like conformations (where the maleimide is the propeller axis and the heterocyclic substituents are two slanted blades). In addition, quantitative structure–activity relationship (QSAR) models using CoMSIA methodology were constructed to explain the trend of the GSK3 $\beta$  inhibitory activities for the studied compounds. We found a model to explain the structure–activity relationship of non-cyclic maleimide (NCM) derivatives (54 compounds). The best CoMSIA model (training set included 44 compounds) included steric, hydrophobic, and HB donor fields and had a good  $Q^2$  value of 0.539. It also predicted adequately the most active compounds contained in the test set. Furthermore, the analysis of the plots of the steric CoMSIA field describes the elements involved in the differential potency of the inhibitors that can be considered for the selection of suitable inhibitors.

**Citation:** Quesada-Romero L, Mena-Ulecia K, Tiznado W, Caballero J (2014) Insights into the Interactions between Maleimide Derivates and GSK3 $\beta$  Combining Molecular Docking and QSAR. PLoS ONE 9(7): e102212. doi:10.1371/journal.pone.0102212

**Editor:** Zhe Zhang, Xuzhou Medical college, China

**Received:** April 20, 2014; **Accepted:** June 16, 2014; **Published:** July 10, 2014

**Copyright:** © 2014 Quesada-Romero et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability:** The authors confirm that all data underlying the findings are fully available without restriction. 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.

\* Email: jcaballero@utalca.cl

## Introduction

Glycogen synthase kinase-3 (GSK3) is a widely expressed and multifunctional serine/threonine protein kinase involved in a large number of cellular processes and diseases. GSK3 is regulated by several mechanisms including phosphorylation [1], intracellular localization [2], and protein complex formation [3]. On the other hand, GSK3 regulates many cellular processes such as cellular architecture and motility [4], and contributes to cell death and cell survival [5,6]. In the last decades, GSK3 has been extensively investigated because its dysregulation is associated to several diseases including Alzheimer's disease [7], diabetes [7,8], cancer [3], muscle hypertrophy [9], etc.

GSK3 is encoded by two isoforms in mammals named GSK3 $\alpha$  and GSK3 $\beta$  [10]. Both isoforms have almost identical catalytic domains, they are activated by tyrosine phosphorylation (Tyr279/216 in GSK3 $\alpha$ /GSK3 $\beta$ ) and are inhibited by phosphorylation in Ser21 in GSK3 $\alpha$  and Ser9 and Thr390 in GSK3 $\beta$  [11]. Different roles in diseases have been identified for each isoform: for instance, GSK3 $\beta$  is overexpressed in many types of cancer including ovarian cancer [12], pancreatic cancer [13], colon cancer [14], etc; meanwhile, there are few reports on the role of GSK3 $\alpha$  in cancer [15].

GSK3 $\beta$  has been proposed as a target for therapy in order to combat several diseases. Many small organic chemical compounds have been developed as ATP competitive GSK3 $\beta$  inhibitors [16]. Among them, a series of macrocyclic and non-cyclic maleimide derivatives (MCMs and NCMs) was reported, giving some candidates with high potency and selectivity [17–21]. In this work, we modeled the structure of the complexes between GSK3 $\beta$  and these compounds using docking. Active conformations are proposed and the interactions that contribute to form the complexes are described. We also develop quantitative structure–activity relationship (QSAR) models using CoMSIA method. The combined docking-CoMSIA protocol is used to provide information about the structural features of potent inhibitors. With this information, we speculated on the possible causes of differential biological activities.

## Materials and Methods

### Data set

The structures and GSK3 $\beta$  inhibitory activities of 23 MCMs and 54 NCMs were collected from the literature [17–20]. The tridimensional (3D) structures were sketched using Maestro's molecular editor (Maestro 9.0, Schrödinger LLC). Activities were

**Table 1.** Experimental and predicted GSK3β inhibitory activities ( $\log(10^3/IC_{50})$ ) of MCMs and NCMs using model CoMSIA-NCM-SHD.

Compounds <sup>a</sup>	R	X	Y	n	m	$\log(10^3/IC_{50})$	
						Exp.	Pred. <sup>b</sup>
1a	H	C	C	1	-	3.866	-
1b	H	C	C	2	-	4.657	-
1c	H	C	C	4	-	4.292	-
1d	H	C	N	1	-	3.860	-
1e	H	C	N	2	-	4.769	-
1f	H	N	N	1	-	3.207	-
1g	H	N	N	2	-	4.468	-
1h	H	N	N	3	-	4.318	-
2a	CH <sub>3</sub>	H	H	-	-	4.455	4.203
2b	CH <sub>3</sub>	5-Br	H	-	-	5.154	4.761
2c	CH <sub>3</sub>	6-OH	H	-	-	4.823	4.741
2d	CH <sub>3</sub>	6-OH	5'-F	-	-	4.853	4.810
2e <sup>c</sup>	CH <sub>3</sub>	5-C≡CH	H	-	-	5.017	3.995
2f	H	H	5'-F	-	-	3.173	3.521
2g <sup>c</sup>	CH <sub>3</sub>	H	5'-Br	-	-	3.259	5.481
2h	CH <sub>3</sub>	H	7-OCH <sub>3</sub>	-	-	3.744	4.277
2i	H	5-F	H	-	-	3.444	3.332
2j <sup>c</sup>	CH <sub>3</sub>	5-F	H	-	-	4.585	5.077
2k	CH <sub>3</sub>	5-Cl	5'-F	-	-	4.377	4.520
2l	CH <sub>3</sub>	5-OCH <sub>3</sub>	H	-	-	3.903	3.682
2m	H	5-OBn	H	-	-	2.783	2.900
2n	CH <sub>3</sub>	5-OBn	H	-	-	3.301	3.060
2o	(CH <sub>2</sub> ) <sub>3</sub> OH	5-OBn	H	-	-	3.657	3.703
2p	CH <sub>3</sub>	6-OBn	H	-	-	3.045	3.384
2q <sup>c</sup>	CH <sub>3</sub>	6-OBn	5'-F	-	-	3.796	2.948
2r	CH <sub>3</sub>	7-OBn	H	-	-	3.657	3.597
2s	CH <sub>3</sub>	5-cyclopropane	H	-	-	3.628	3.771
2t	CH <sub>3</sub>	5-OH	H	-	-	3.161	3.238
2u <sup>c</sup>	CH <sub>3</sub>	7-OH	H	-	-	4.260	5.977
2v	CH <sub>3</sub>	5-F	6'-CH <sub>2</sub> OH	-	-	6.455	6.596
2w	CH <sub>3</sub>	5-F	6'-CH <sub>2</sub> OCH <sub>3</sub>	-	-	4.623	4.491
2x	CH <sub>3</sub>	5-F	6'-OH	-	-	5.455	5.715
2y	CH <sub>3</sub>	5-F, 6-I	7'-OCH <sub>3</sub>	-	-	3.607	3.472
2z	CH <sub>3</sub>	5-F, 6-Cl	H	-	-	3.735	3.904

Table 1. Cont.

Compounds <sup>a</sup>	R	X	Y	n	m	log(10 <sup>3</sup> /IC <sub>50</sub> ) Exp.	Pred. <sup>b</sup>
2aa	CH <sub>3</sub>	5-F, 6-Cl	6'-CH <sub>2</sub> OH	-	-	6.022	6.124
2ab	CH <sub>3</sub>	5-F, 6-Cl	6'-OCH <sub>3</sub>	-	-	3.060	3.326
2ac	CH <sub>3</sub>	5-F, 6-Cl	6'-cyclopropylmethoxy	-	-	2.889	2.599
2ad	CH <sub>3</sub>	5-F, 6-Cl	6'-cyclobutylmethoxy	-	-	2.388	2.394
2ae <sup>c</sup>	CH <sub>3</sub>	5-F, 6-Cl	7'-OCH <sub>3</sub>	-	-	3.585	4.411
2af	CH <sub>3</sub>	5-F, 6-p-Cl-Ph	7'-OCH <sub>3</sub>	-	-	2.145	2.143
2ag	CH <sub>3</sub>	5-Br	7'-OCH <sub>3</sub>	-	-	5.124	5.262
2ah	CH <sub>3</sub>	5-Br	6'-CH <sub>2</sub> OH	-	-	6.292	6.333
2ai	CH <sub>3</sub>	5-Br	6'-prop-2'-ynyloxy	-	-	4.596	4.494
2aj	CH <sub>3</sub>	5-Br	6'-allyloxy	-	-	4.316	4.372
2ak	CH <sub>3</sub>	5-Br	6'-O-(p-CH <sub>3</sub> O)-Bn	-	-	3.474	3.430
2al	CH <sub>3</sub>	5,7-dibromo	7'-OCH <sub>3</sub>	-	-	4.052	4.034
2am <sup>c</sup>	CH <sub>3</sub>	5-I	H	-	-	4.462	4.531
2an <sup>c</sup>	CH <sub>3</sub>	5-I	5'-F	-	-	3.744	4.670
2ao	CH <sub>3</sub>	5-CN	6'-CH <sub>2</sub> OH	-	-	4.879	5.009
2ap	CH <sub>3</sub>	5-cyclopropylethynyl	5'-F	-	-	4.793	4.807
2aq	CH <sub>3</sub>	5,6-methylenedioxy	5'-F	-	-	3.148	3.511
2ar	CH <sub>3</sub>	5-OCH <sub>3</sub> , 6-Cl	H	-	-	3.356	3.077
2as	CH <sub>3</sub>	5-OCH <sub>3</sub> , 6-I	H	-	-	3.651	3.611
2at <sup>c</sup>	CH <sub>3</sub>	6-CF <sub>3</sub>	7'-OCH <sub>3</sub>	-	-	3.080	3.888
2au	CH <sub>3</sub>	7-CH <sub>2</sub> OH	H	-	-	5.267	4.997
2av	CH <sub>3</sub>	7-CH <sub>2</sub> OH	6'-CH <sub>2</sub> OH	-	-	5.292	5.441
2aw	CH <sub>3</sub>	7-CH <sub>2</sub> OMe	H	-	-	6.638	6.315
2ax <sup>c</sup>	CH <sub>3</sub>	7-CH <sub>2</sub> OMe	6'-CH <sub>2</sub> OH	-	-	6.136	7.371
2ay	CH <sub>3</sub>	7-CH <sub>2</sub> CH <sub>2</sub> COOEt	H	-	-	4.991	4.777
2az	CH <sub>3</sub>	7-CH <sub>2</sub> CH <sub>2</sub> COOH	H	-	-	5.920	5.748
2ba	CH <sub>3</sub>	7H-benzogl	5',6'-difluoro	-	-	3.503	3.513
3a	NEt	O	O	1	1	4.602	-
3b	NMe	O	O	1	1	4.149	-
3c	NMe	NMe	O	1	1	4.481	-
3d	NMe	NMe	NMe	1	1	4.367	-
3e	O	O	NMe	1	2	5.398	-
3f	O	O	NEt	1	1	5.398	-
3g	None	NMe	NMe	0	1	3.959	-

Table 1. Cont.

Compounds <sup>a</sup>	R	X	Y	n	m	log(10 <sup>3</sup> /IC <sub>50</sub> )	
						Exp.	Pred. <sup>b</sup>
3h	O	NMe	NMe	1	1	5.222	-
3i	O	NMe	NMe	2	1	5.097	-
4a	Me <sub>2</sub> N	1	1	-	meta	5.155	-
4b	(CH <sub>2</sub> ) <sub>4</sub> N	1	1	-	meta	4.959	-
4c	Me <sub>2</sub> N	1	1	-	para	4.432	-
4d	(CH <sub>2</sub> ) <sub>4</sub> N	1	1	-	para	4.523	-
4e	Me <sub>2</sub> N	1	2	-	para	5.523	-
4f	(CH <sub>2</sub> ) <sub>4</sub> N	2	2	-	para	4.620	-
5	-	-	-	-	-	4.301	4.155

<sup>a</sup>Compounds **1** from reference [17], compounds **2a–2u** from reference [19], compounds **2v–2ba** and **5** from reference [20], compounds **3** from reference [21], and compounds **4** from reference [18].

<sup>b</sup>Predictions using model CoMSIA-NCM-SHD.

<sup>c</sup>Compounds predicted in the test set.  
doi:10.1371/journal.pone.0102212.t001

collected and transformed into log(10<sup>3</sup>/IC<sub>50</sub>) values, where IC<sub>50</sub> values represent the compound μM concentrations that inhibit the GSK3β activity by 50%. The compounds under study and their inhibitory biological activities are summarized in Figure 1 and Table 1.

## Docking

Generally, docking algorithms reproduce the bound form of ligands inside the active site of proteins [22–30]. In our current work, docking was performed using Glide [31]. Protein coordinates were extracted from the X-ray crystal structure of the GSK3β-inhibitor complex with the code 2OW3 in Protein Data Bank [18]. A grid box of 30Å×30Å×30Å was centered on the center of mass of the inhibitor in this crystal structure covering the ATP-binding site of the enzyme. The module LigPrep 2.5 [32] was used to assign ionization states, stereochemistries, and ring conformations of the sketched ligands. Docking parameters were used as in previous works [29], and Glide standard (SP) and extra-precision (XP) modes were explored during the search. The better docking poses for each ligand were analyzed by examining their relative total energy scores. Among docking poses, the more energetically favorable conformation was selected by considering the total energy value.

## QSAR modeling

CoMSIA was applied to selected training set, and then external set was predicted. Firstly, CoMSIA was performed on the whole dataset including 64 compounds in the training set and 13 compounds in the external set, but we did not find a proper result. Therefore, we constructed separated CoMSIA models describing the structure-activity relationships of MCMs and NCMs. MCMs (23 compounds) were modeled without splitting of the dataset, and NCMs (54 compounds) were modeled including 44 compounds in the training set and 10 compounds in the external set.

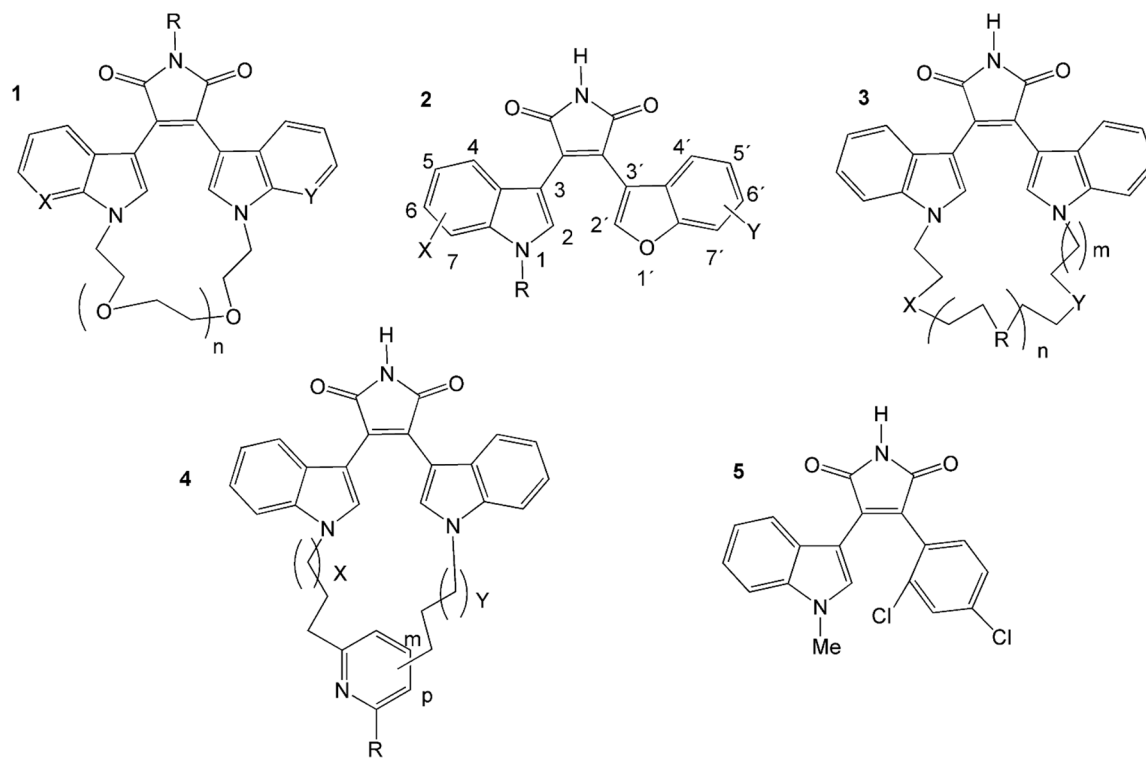
CoMSIA was performed using the Sybyl 7.3 software of Tripos [33]. Field descriptors were calculated on the 3D conformations obtained by the docking approach; this guaranteed that all compounds were aligned in the GSK3β active site. The maleimide derivatives were placed in a rectangular grid extended beyond 4 Å in each direction from the coordinates of each molecule. The interaction energies between a probe atom (an sp<sup>3</sup>-hybridized carbon atom with +1 charge) and all compounds were computed at the surrounding points, using a volume-dependent lattice with 2.0-Å grid spacing. Similarity is expressed in terms of steric occupancy, electrostatic interactions, local hydrophobicity, and hydrogen bond (HB) donor and acceptor properties, using a 0.3 attenuation factor. The number of components in PLS was optimized using the Q<sup>2</sup> value, obtained from the leave-one-out (LOO) crossvalidation procedure, with the SAMPLS [34] sampling method.

## Results and Discussion

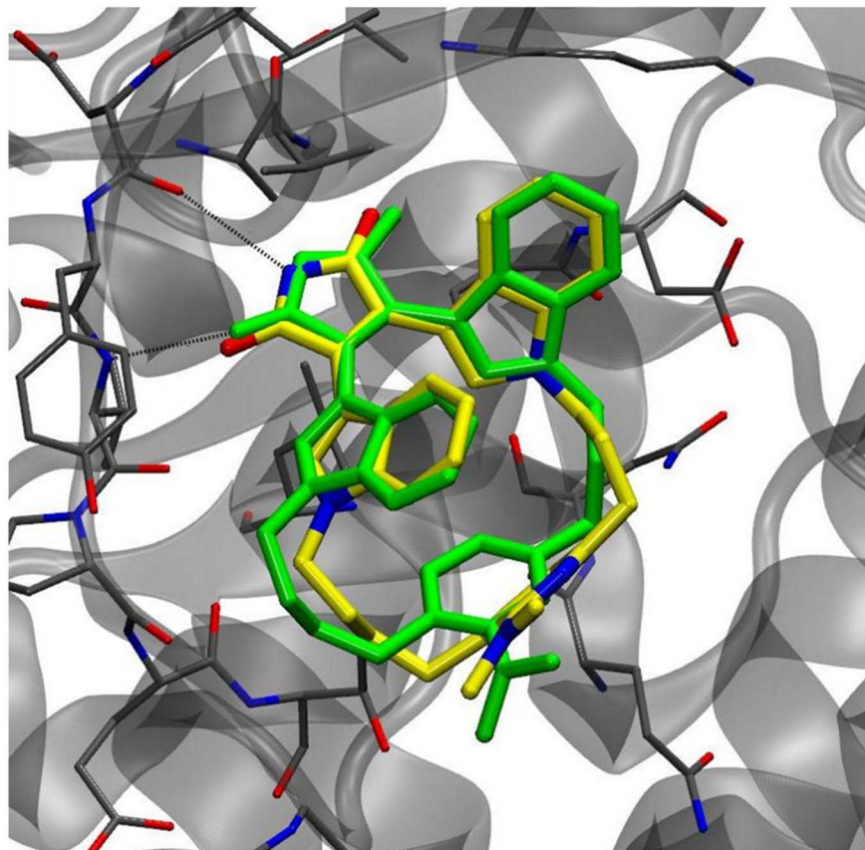
### Docking results

We tested the ability of the docking method to reproduce the pose of compound **4e** close to that found in an X-ray complex reported in reference [18]. As it can be observed in Figure 2, the docked structure closely corresponds to the inhibitor in the X-ray structure 2OW3 (root mean square deviation, RMSD, of all heavy atoms: 0.910 Å). Therefore, we can state that Glide software found a correct binding mode of the studied ligand.

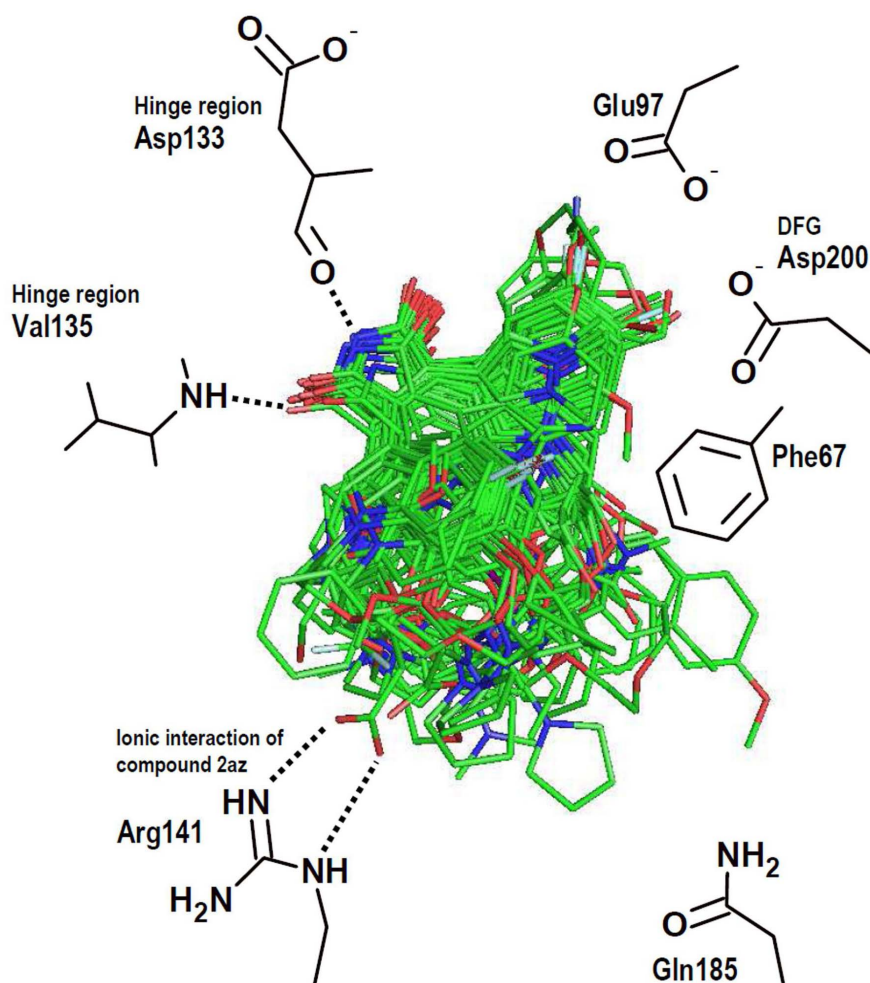
The alignment of the docked structures inside the GSK3β binding site for the remaining studied compounds is shown in Figure 3. The docking protocol was successful since all the ligands



**Figure 1. Structures of MCMs (1, 3, and 4) and NCMs (2 and 5).**  
doi:10.1371/journal.pone.0102212.g001



**Figure 2. Conformational comparison of compound 4e from the crystal structures in GSK3 $\beta$ -inhibitor complex (inhibitor in yellow in online version, light gray in printed version) and from docking results (inhibitor in green in online version, dark gray in printed version).**  
doi:10.1371/journal.pone.0102212.g002



**Figure 3. Predicted binding conformations of all the investigated maleimide derivatives and their interactions with the residues in the GSK3 $\beta$  active site.**

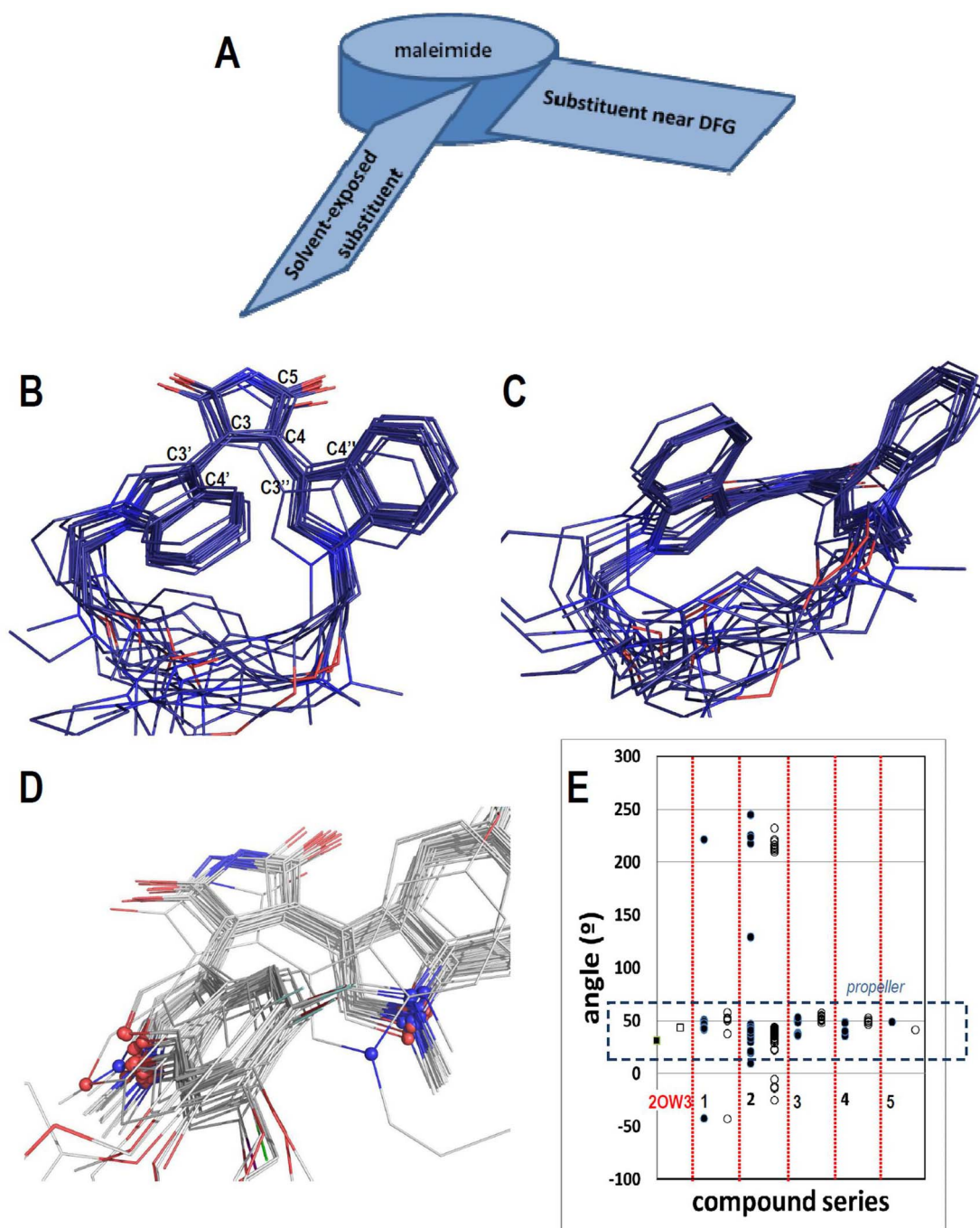
doi:10.1371/journal.pone.0102212.g003

were docked in the expected orientations. In general, our docking results show that all the maleimide derivatives adopt the same mode of binding, characterized by interactions of the maleimide carbonyl and NH groups with residues in the GSK3 $\beta$  hinge region (Figure 3): one of the carbonyl groups forms an HB interaction with the backbone NH of the residue Val135, and the maleimide NH group forms an HB interaction with the backbone carbonyl group of Asp133. In general, the studied compounds contain heterocyclic substituents (benzofuran-3-yl or indol-3-yl) at positions 3 and 4 of maleimide. One of the heterocyclic substituents is located close to the residues Asp200 (DFG motif), Lys85 (catalytic lysine), and Cys199; the other is oriented towards the surface of the pocket (in the region between the residues Ile62 and Leu188). All the MCMs (series **1**, **3**, and **4**) contain two indol-3-yl (or other nitrogen containing heterocycles) substituents. The groups are oriented so as to accommodate the macrocyclic linker in the region between the residues Phe67, Thr138, and Gln185. In this zone the macrocyclic linker is largely solvent-exposed.

Compounds of series **1** contain a polyoxygenated macrocyclic linker and have  $\log(10^3/IC_{50})$  between 3.2 and 4.8; while compounds of series **3** contain multiple heteroatoms on the macrocyclic linker, and have  $\log(10^3/IC_{50})$  between 3.9 and 5.4. The orientations of the poses obtained for these series suggest that HB donor groups at the macrocyclic linkers could interact with the

side chain NH<sub>2</sub> group of the residue Asn186. On the other hand, compounds of series **4** are macrocyclic pyridinophanes and have  $\log(10^3/IC_{50})$  between 4.4 and 5.6. The groups that differentiate the activity in this series are exposed to the solvent close to the residues Gln185 and Thr138. The obtained docking poses do not reveal the source of the differential activity among compounds of the series **4**, since the pyridinophane groups are not involved in important contacts with the protein.

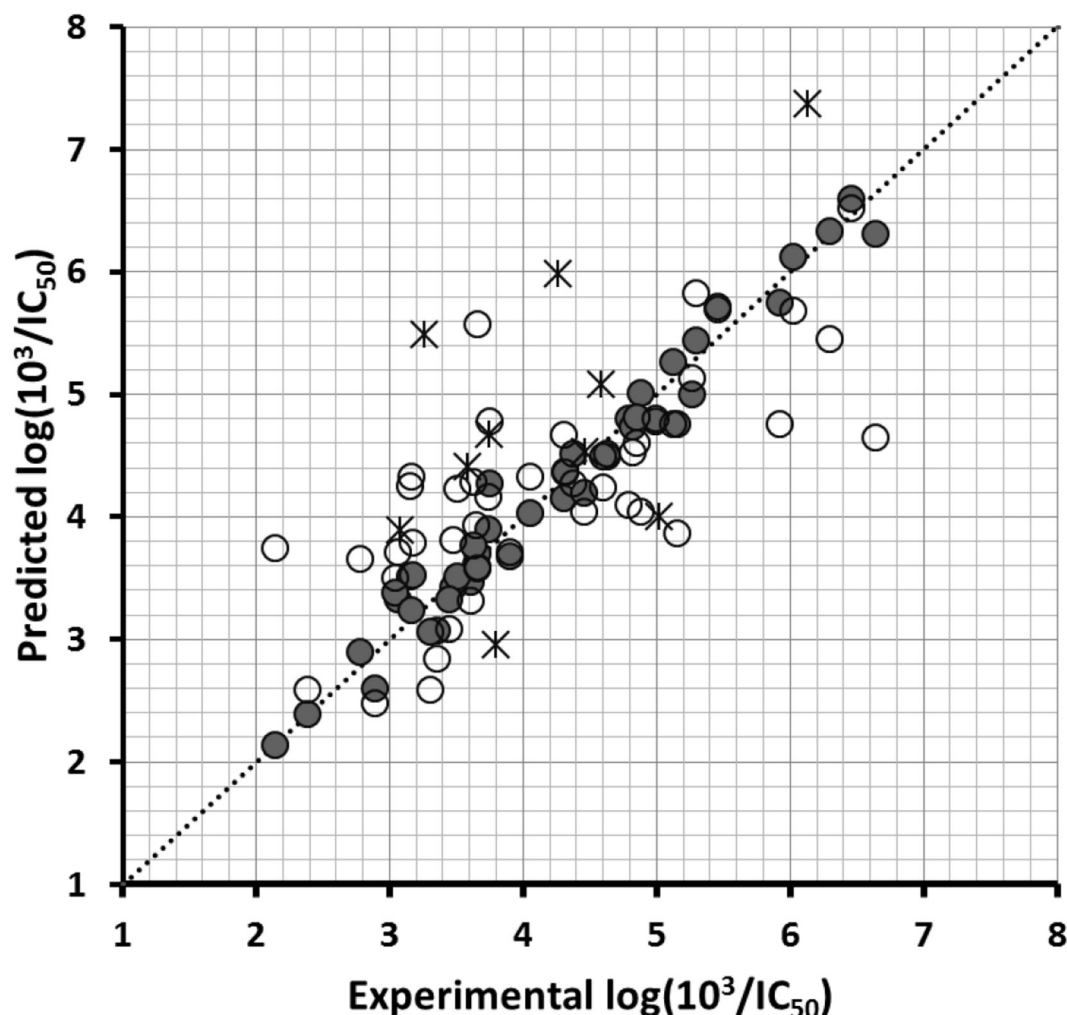
A remarkable feature of the obtained docking poses is that 3, 4 disubstituted maleimides form a propeller-like conformation inside the GSK3 $\beta$  binding site: the propeller axis (the maleimide) with two slanted blades (the heterocyclic substituents), as represented in the Figure 4A. Figures 4B and 4C show two views of the propeller conformations for MCMs; where the indol-3-yl groups at the left are solvent-exposed and the indol-3-yl groups at the right are near the DFG motif (considering the positions of the inhibitors inside the GSK3 $\beta$  binding site). The figure 4D shows the propeller conformations for NCMs. In general, the benzofuran-3-yl groups are solvent-exposed and indol-3-yl groups are near the DFG motif, but several compounds exchange these common orientations, which is reasonable considering the symmetry of the NCMs. The inclination of the blades (heterocyclic substituents) was analyzed for each series by measuring dihedral angles C4'-C3'-C3-C4 (Angle1) and C4''-C3''-C4-C5 (Angle2) defined in Figure 4B. The



**Figure 4. Analysis of the propeller conformations of the docked maleimides.** (A) Scheme of the propeller conformations. (B, C) Two views of the propeller conformations of MCMs. (D) Solvent-exposed substituents (at the left) and substituents near the DFG Asp200 (at the right) of NCMs showing the presence of indol-3-yl and benzofuran-3-yl groups at each position. (E) Dihedral angles Angle1 C4'-C3'-C3-C4 (close circles to the left) and Angle2 C4'-C3'-C4-C5 (open circles to the right) for each compound series (consider labels of the atoms as represented in Figure 3A). The same dihedral angles are represented for the crystallographic structure of **4e** (close and open squares, PDB code: 2OW3). The dashed lines box encloses the propeller conformations.  
doi:10.1371/journal.pone.0102212.g004

results are plotted in Figure 4E, where it is possible to see that the inclinations that lead to propeller conformations are characterized by Angle1 and Angle2 values between 30 and 56°. All the compounds from series **3** and **4** have perfect propeller conformations (Angle1 = 44.8° ± 7.1 and 42.8° ± 6.2; Angle2 = 52.5° ± 2.8 and 49.4° ± 2.1, for series **3** and **4**, respectively). The majority of compounds from series **1** also have propeller conformations

(Angle1 = 46.0° ± 3.8; Angle2 = 51.8° ± 0.9). In general, the analysis of the docking poses for MCMs (compounds from series **1**, **3**, and **4**) shows that the heterocyclic substituents near the DFG motif lean at higher angle values with respect to solvent-exposed ones. We consider that this is an effect of the rigid macrocyclic linkers. It is noteworthy that the known co-crystallized structure of compound **4e** in GSK3 $\beta$  (pdb: 3OW3) has slightly lesser values



**Figure 5. Scatter plot of the experimental activities versus predicted activities for model CoMSIA-NCM-SHD: (●) training-set predictions, (○) LOO cross-validated predictions, (×) test-set predictions.**  
doi:10.1371/journal.pone.0102212.g005

for Angle1 (31.2°) and Angle2 (43.7°), but Angle2 is higher than Angle1 in accordance with docking results. The majority of compounds from series **2** also have propeller conformations, but Angle1 and Angle2 have similar values for these compounds (Angle1 = 40.4° ± 2.5; Angle2 = 39.6° ± 3.3). Finally, Angle1 and Angle2 are 48.8 and 41.2 respectively for the docking pose of compound **5**. According to this analysis, NCMs (compounds of series **2** and compound **5**) also have propeller conformations with less inclined blades inside GSK3 $\beta$ .

In some cases, the presence or absence of propeller conformations are due to the formation of HBs with polar groups in the active site. For instance, compounds that contain CH<sub>2</sub>OH substituent at position 6 of the indol-3-yl near the DFG motif, such as **2ah**, **2au**, and **2av**, form the propeller conformation with an additional HB between the CH<sub>2</sub>OH and the catalytic lysine (Lys85) or DFG aspartate (Asp200). On the other hand, compounds that contain CH<sub>2</sub>OH substituent at position 6 of the solvent-exposed indol-3-yl, such as **2v** and **2aa**, form the propeller conformation with an additional HB between the CH<sub>2</sub>OH and the backbone carbonyl of the residue Ile62. It is noteworthy that the majority of compounds with CH<sub>2</sub>OH substituent at position 6 of the indol-3-yl have log(10<sup>3</sup>/IC<sub>50</sub>) above 6. A striking case is the compound **2az** that contains a carboxylate group and a high

log(10<sup>3</sup>/IC<sub>50</sub>) value of 5.920. This compound has the propeller conformation and establishes an interesting ionic interaction between the carboxylate group and the residue Arg141 of GSK3 $\beta$  (it is indicated in the Figure 3). During the analysis of docking poses, we identified that compounds that have bulky substituents, such as **2m**, **2n**, **2q**, **2s**, **2ap**, and **2ay**, do not have a propeller conformation due to steric problems. In general, these compounds are less active.

#### QSAR models

We also constructed CoMSIA models to identify the structural features of the maleimide derivatives that affect their inhibitory activity against GSK3 $\beta$ . The models were developed by using the docking aligned conformations, allowing the CoMSIA contours display into the GSK3 $\beta$  active site. Models were derived from different combinations of up to five fields. The best model was selected by the analysis of the statistical quality of the internal LOO cross-validation for each model, taking into account  $Q^2$  values. We developed models for describing the whole dataset (77 compounds), the MCMs (23 compounds), and NCMs (54 compounds), after performing the splittings mentioned above in the Materials and methods section.



**Table 2.** CoMSIA results using the best field combinations.<sup>a</sup>

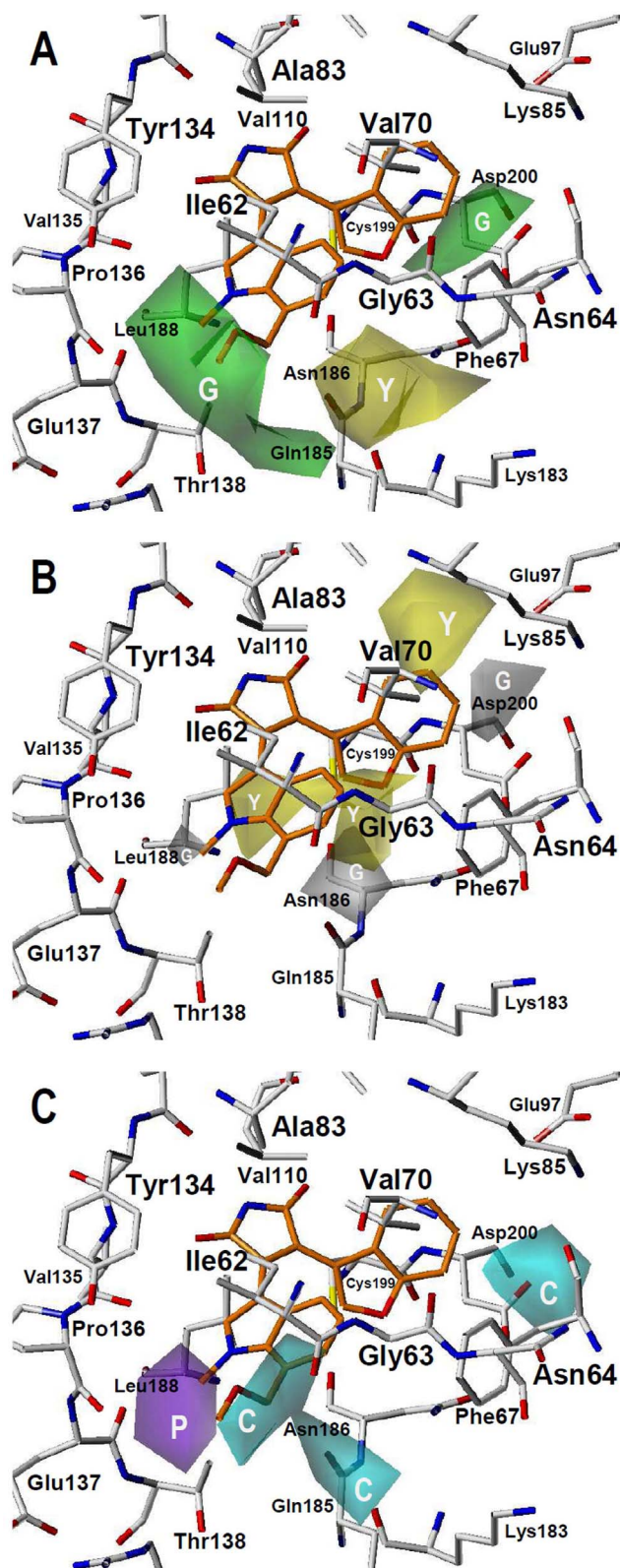
Models	NC	R <sup>2</sup>	S	F	Q <sup>2</sup>	S <sub>cv</sub>	Fraction				
							S	E	H	D	A
CoMSIA-SHD	5	0.843	0.410	62.17	0.304	0.863	0.208	0.549	0.243		
CoMSIA-SEHD	5	0.901	0.325	106.02	0.251	0.895	0.112	0.380	0.170		
CoMSIA-SH	5	0.759	0.508	36.52	0.236	0.904	0.291	0.709			
CoMSIA-MCM-SA	1	0.492	0.416	20.36	0.217	0.516	0.171		0.829		
CoMSIA-MCM-A	1	0.464	0.427	18.18	0.206	0.520				1	
CoMSIA-MCM-SDA	1	0.547	0.393	25.37	0.184	0.527	0.142		0.168	0.690	
<b>CoMSIA-NCM-SHD</b>	<b>5</b>	<b>0.921</b>	<b>0.329</b>	<b>89.13</b>	<b>0.539</b>	<b>0.802</b>	<b>0.171</b>	<b>0.562</b>	<b>0.267</b>		
CoMSIA-NCM-HD	5	0.908	0.355	75.39	0.520	0.813		0.697	0.303		
CoMSIA-NCM-EHD	5	0.945	0.275	131.22	0.459	0.863		0.380	0.418	0.202	

<sup>a</sup>NC is the number of components from the 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; and Q<sup>2</sup> and S<sub>cv</sub> are the correlation coefficient and standard deviation, respectively, of the leave-one-out (LOO) cross-validation.  
doi:10.1371/journal.pone.0102212.t002

The results of the search are included in Table 2. We could not find predictive models for describing the structure-activity relationship of the whole dataset and the subset of MCMs. On the other hand, we found the model CoMSIA-NCM-SHD ( $Q^2 = 0.539$ ) for describing the differential GSK3β inhibitory activities of NCMs. This model uses five components and combines steric, hydrophobic, and HB donor fields with contributions of 17.1%, 56.2%, and 26.7% respectively. In addition, it explains 92.1 of the variance and has a low standard deviation ( $S = 0.329$ ) and a high Fischer ratio ( $F = 89.13$ ). The predictions of  $\log(10^3/IC_{50})$  values for the 44 NCMs from the training set using model CoMSIA-NCM-SHD are shown in Table 1. The correlations between the calculated and experimental  $\log(10^3/IC_{50})$  values (from training and LOO cross-validation) are shown in Figure 5. According to these plots, the model properly discriminates between the most and less active compounds. We also predicted the GSK3β inhibitory activities of the test set compounds using model CoMSIA-NCM-SHD. The obtained prediction values are given in Table 1, and correlation between the calculated and experimental values is represented in Figure 5. This analysis reveals that the proposed model identified the most active compounds in the test set, but certainly not a good correlation was obtained when  $R^2$  of the test set was analyzed. Therefore, the predictive evaluation of this model using rigorous external validation testing (calculation of  $R_m^2$  [35]) lead to unfavorable results.

The contour plots of the CoMSIA steric, hydrophobic, and HB donor fields are presented in Figure 6 for the best model CoMSIA-NCM-SHD. The highly active compound **2aw** is displayed in the maps to aid in visualization, and the superposition of CoMSIA contour plots on active-site residues is also shown. The colored isopleths in the map represent the 3D locations where the structural properties changes are related to the changes in biological potency. Green and yellow isopleths in Figure 6A indicate regions where bulky groups increase and decrease the inhibitory activity, respectively. A large region of green contour near the solvent-exposed indol-3-yl moiety of compound **2aw** suggests that there is a favorable steric region near the residues Leu188, Thr138, and Gln185. In fact, compound **2aw** has methyl and methoxymethyl substituents at positions 1 and 7 respectively of the indol-3-yl moiety that occupy this region. Other green isopleth near the residues Asp200 and Phe67 indicates that bulky groups at positions 6 and 7 on the indol-3-yl or benzofuran-3-yl in the neighborhood of the DFG motif increase the inhibitory activity. This is the case of active compounds that contain hydroxymethyl group; for instance, compound **2av** ( $\log(10^3/IC_{50}) = 5.292$ ) contains this group at position 7 of the indol-3-yl substituent, and compound **2ah** ( $\log(10^3/IC_{50}) = 6.292$ ) contains this group at position 6 of the benzofuran-3-yl substituent. Finally, the yellow isopleth near the residue Asn186 suggests that large groups are not desired in this zone. It must be remembered that bulky groups that impair the formation of the propeller conformation occupy this region; for instance, the less active compound **2ad** ( $\log(10^3/IC_{50}) = 2.388$ ) contains a cyclobutyl in this region.

Yellow isopleths in Figure 6B indicate regions where hydrophobic groups are favorable for activity or hydrophilic groups diminish the activity; meanwhile, gray isopleths represent areas where hydrophobic groups are not favorable for activity or hydrophilic groups increase the activity. According to the analysis of the yellow isopleths, hydrophobic groups are tolerated in the region in front of the positions 2 and 3 of the maleimide. Several active compounds contain a methyl group at position 1 of the indol-3-yl near the DFG motif in this region. There is other yellow



**Figure 6. CoMSIA contour maps for GSK3 $\beta$  inhibitors deriving from model CoMSIA-NCM-SHD.** The amino acid residues located close to the binding pocket of GSK3 $\beta$  are represented for comparing their position with the position of isopleths derived from the model. Compound **2aw** is shown inside the fields. (A) Steric field: green isopleths indicate regions where bulky groups favor the activity, and yellow isopleths indicate regions where bulky groups disfavor the

activity. (B) Hydrophobic field: yellow isopleths indicate regions where hydrophobic groups favor the activity, and gray isopleths indicate regions where hydrophilic groups favor the activity. (C) HB donor field: cyan isopleths indicate regions where HB donors favor the activity, and purple isopleths indicate regions where HB donors disfavor the activity. doi:10.1371/journal.pone.0102212.g006

isopleth near the catalytic lysine (Lys85) that is occupied by halogen substituents at position 5 of the indol-3-yl group near the DFG motif (for instance: compound **2v**,  $\log(10^3/IC_{50}) = 6.455$ ). Very close, there is a gray isopleth that is occupied by halogen substituents at position 6 of the same group in several non-active compounds (for instance: compound **2ab**,  $\log(10^3/IC_{50}) = 3.060$ ). Those features together indicate that hydrophobic groups are desired at position 6 of the indol-3-yl group near the DFG motif, but not at position 5. There is other gray isopleth near Asn186 that is occupied by groups at position 6 of the solvent-exposed indol-3-yl or benzofuran-3-yl. The analysis of this isopleth indicates that hydrophobic groups are not desired in this region. For instance, compound **2y** contains an iodide at this position and has  $\log(10^3/IC_{50}) = 3.607$ ; meanwhile, compound **2aa** contains a hydroxyl in this region and has  $\log(10^3/IC_{50}) = 6.022$ . A very small gray isopleth near Leu188 is occupied by 1-methyl substituents at solvent-exposed indol-3-yl, suggesting that benzofuran-3-yl groups are preferred in this region (instead of the 1-methyl-1*H*-indol-3-yl groups).

Cyan and purple isopleths in Figure 6C are in regions where HB donor groups favor and disfavor the activity, respectively. The cyan isopleth near Asp200 indicates that substituents containing HB donor groups on the indol-3-yl or benzofuran-3-yl near the DFG motif, such as 6-hydroxyl (compound **2x**,  $\log(10^3/IC_{50}) = 5.455$ ), 6-hydroxymethyl (compound **2ah**,  $\log(10^3/IC_{50}) = 6.292$ ), and 7-hydroxymethyl (compound **2av**,  $\log(10^3/IC_{50}) = 5.292$ ), increase the inhibitory activity. Other cyan isopleth is near the backbone carbonyl of the residue Ile62, which is occupied by hydroxymethyl substituents at position 6 of the solvent-exposed benzofuran-3-yl of several active compounds (for instance, compound **2v**,  $\log(10^3/IC_{50}) = 6.596$ ). Finally, a purple isopleth is located near the backbone carbonyl of the residue Pro136, suggesting that HB donor groups are not desired in this zone. In fact, compounds containing a 1-methyl-indol-3-yl have a better inhibitory activity than [36] **2f** and **2i** ( $\log(10^3/IC_{50}) = 3.133$  and 3.444, respectively).

Given the model CoMSIA-NCM-SHD reported here, it is easier to explain the trend of the potency of the NCMs against GSK3 $\beta$ . This information can orient in chemical synthesis of new candidates. Some maleimide derivatives studied in this work were previously studied by Fang et al [36]. These authors extracted the compounds (only NCMs) from the reference [21] and constructed CoMFA and CoMSIA models using 30 compounds in the training set and 8 compounds in the test set. They did not determine the active poses (using docking) before QSAR calculations; therefore, their approach only considers ligand-based alignment. The goal of the work of Fang et al. was the construction of predictive models (of diverse GSK3 $\beta$  inhibitors) for identifying new hits with topologically diverse scaffolds. Instead, our approach has the purpose of reporting the structural characteristics of the active conformations of the maleimide derivatives (NCMs and MCMs) inside GSK3 $\beta$ , and developing CoMSIA models (with atom fit alignment rules based on receptor-ligand bound conformations) that describe the structure-activity relationship.

## Conclusions

The structure–activity relationship of maleimides derivatives as GSK3 $\beta$  inhibitors was studied by using docking and QSAR methods. Some characteristics of these compounds that explain their activities were described such as their orientation and the interactions that they establish with the residues located in the binding site. The most interesting finding is that the most active maleimides adopt a propeller-like conformation. A CoMSIA model was derived including steric, hydrophobic, and HB donor

fields. A map of the characteristics that are desired for highly active compounds, and their relation with the residues of the active site, are presented.

## Author Contributions

Conceived and designed the experiments: JC WT. Performed the experiments: LQR KMU. Analyzed the data: JC WT. Contributed reagents/materials/analysis tools: JC WT LQR KMU. Contributed to the writing of the manuscript: LQR JC.

## References

- Grimes CA, Jope RS (2001) The multifaceted roles of glycogen synthase kinase 3 $\beta$  in cellular signaling. *Prog Neurobiol* 65: 391–426. doi:10.1016/S0301-0082(01)00011-9.
- Diehl JA, Cheng M, Roussel MF, Sherr CJ (1998) Glycogen synthase kinase-3 $\beta$  regulates cyclin D1 proteolysis and subcellular localization. *Genes Dev* 12: 3499–3511.
- Manoukian AS, Woodgett JR (2002) Role of glycogen synthase kinase-3 in cancer: Regulation by Wnts and other signaling pathways. *Adv Cancer Res* 84: 203–229.
- Wakefield JG, Stephens DJ, Tavaré JM (2003) A role for glycogen synthase kinase-3 in mitotic spindle dynamics and chromosome alignment. *J Cell Sci* 116: 637–646.
- Pérez M, Rojo AI, Wandosell F, Díaz-Nido J, Avila J (2003) Prion peptide induces neuronal cell death through a pathway involving glycogen synthase kinase 3. *Biochem J* 372: 129–136. doi:10.1042/BJ20021596.
- Pap M, Cooper GM (1998) Role of Glycogen Synthase Kinase-3 in the Phosphatidylinositol 3-Kinase/Akt Cell Survival Pathway. *J Biol Chem* 273: 19929–19932. doi:10.1074/jbc.273.32.19929.
- Gao C, Hölscher C, Liu Y, Li L (2012) GSK3: a key target for the development of novel treatments for type 2 diabetes mellitus and Alzheimer disease. *Rev Neurosci* 23: 1–11. doi:10.1515/rns.2011.061.
- Rayasam GV, Tulasi VK, Sodhi R, Davis JA, Ray A (2009) Glycogen synthase kinase 3: more than a namesake. *Br J Pharmacol* 156: 885–898. doi:10.1111/j.1476-5381.2008.00085.x.
- Sugden PH, Fuller SJ, Weiss SC, Clerk A (2008) Glycogen synthase kinase 3 (GSK3) in the heart: a point of integration in hypertrophic signalling and a therapeutic target? A critical analysis. *Br J Pharmacol* 153 Suppl 1: S137–153. doi:10.1038/sj.bjpp.0707659.
- Darrington RS, Campa VM, Walker MM, Bengoa-Vergniory N, Gorrone-Etxebarria I, et al. (2012) Distinct expression and activity of GSK-3 $\alpha$  and GSK-3 $\beta$  in prostate cancer. *Int J Cancer J Int Cancer* 131: E872–883. doi:10.1002/ijc.27620.
- Kaidanovich-Beilin O, Woodgett JR (2011) GSK-3: Functional Insights from Cell Biology and Animal Models. *Front Mol Neurosci* 4: 40. doi:10.3389/fnmol.2011.00040.
- Miyashita K, Kawakami K, Nakada M, Mai W, Shakoori A, et al. (2009) Potential therapeutic effect of glycogen synthase kinase 3 $\beta$  inhibition against human glioblastoma. *Clin Cancer Res Off J Am Assoc Cancer Res* 15: 887–897. doi:10.1158/1078-0432.CCR-08-0760.
- Ougolkov AV, Fernandez-Zapico ME, Bilim VN, Smyrk TC, Chari ST, et al. (2006) Aberrant nuclear accumulation of glycogen synthase kinase-3 $\beta$  in human pancreatic cancer: association with kinase activity and tumor dedifferentiation. *Clin Cancer Res Off J Am Assoc Cancer Res* 12: 5074–5081. doi:10.1158/1078-0432.CCR-06-0196.
- Shakoori A, Ougolkov A, Yu ZW, Zhang B, Modarressi MH, et al. (2005) Deregulated GSK3 $\beta$  activity in colorectal cancer: Its association with tumor cell survival and proliferation. *Biochem Biophys Res Commun* 334: 1365–1373. doi:10.1016/j.bbrc.2005.07.041.
- Wilson W 3rd, Baldwin AS (2008) Maintenance of constitutive I $\kappa$ B $\alpha$  kinase activity by glycogen synthase kinase-3 $\alpha$ /3 $\beta$  in pancreatic cancer. *Cancer Res* 68: 8156–8163. doi:10.1158/0008-5472.CAN-08-1061.
- Bertrand JA, Thieffine S, Vulpetti A, Cristiani C, Valsasina B, et al. (2003) Structural Characterization of the GSK-3 $\beta$  Active Site Using Selective and Non-selective ATP-mimetic Inhibitors. *J Mol Biol* 333: 393–407. doi:10.1016/j.jmb.2003.08.031.
- Kuo G-H, Prouty C, DeAngelis A, Shen L, O'Neill DJ, et al. (2003) Synthesis and discovery of macrocyclic polyoxygenated bis-7-azaindoly-maleimides as a novel series of potent and highly selective glycogen synthase kinase-3 $\beta$  inhibitors. *J Med Chem* 46: 4021–4031. doi:10.1021/jm030115o.
- Zhang H-C, Boñaga LVR, Ye H, Derian CK, Damiano BP, et al. (2007) Novel bis(indolyl)maleimide pyridinophanes that are potent, selective inhibitors of glycogen synthase kinase-3. *Bioorg Med Chem Lett* 17: 2863–2868. doi:10.1016/j.bmcl.2007.02.059.
- Kozikowski AP, Gaisina IN, Yuan H, Petukhov P, Blond SY, et al. (2007) Structure-based design leads to the identification of lithium mimetics that block mania-like effects in rodents. possible new GSK-3 $\beta$  therapies for bipolar disorders. *J Am Chem Soc* 129: 8328–8332. doi:10.1021/ja068969w.
- Zhang H-C, White KB, Ye H, McComsey DF, Derian CK, et al. (2003) Macrocyclic bisindolylmaleimides as inhibitors of protein kinase C and glycogen synthase kinase-3. *Bioorg Med Chem Lett* 13: 3049–3053. doi:10.1016/S0960-894X(03)00644-9.
- Gaisina IN, Gallier F, Ougolkov AV, Kim KH, Kurome T, et al. (2009) From a natural product lead to the identification of potent and selective benzofuran-3-yl-(indol-3-yl)maleimides as glycogen synthase kinase 3 $\beta$  inhibitors that suppress proliferation and survival of pancreatic cancer cells. *J Med Chem* 52: 1853–1863. doi:10.1021/jm801317h.
- Alzate-Morales JH, Vergara-Jaque A, Caballero J (2010) Computational Study on the Interaction of N1 Substituted Pyrazole Derivatives with B-Raf Kinase: An Unusual Water Wire Hydrogen-Bond Network and Novel Interactions at the Entrance of the Active Site. *J Chem Inf Model* 50: 1101–1112. doi:10.1021/ci100049h.
- Caballero J, Muñoz C, Alzate-Morales JH, Cunha S, Gano L, et al. (2012) Synthesis, in silico, in vitro, and in vivo investigation of 5-[<sup>14</sup>C]methoxy-substituted sunitinib, a tyrosine kinase inhibitor of VEGFR-2. *Eur J Med Chem* 58: 272–280. doi:10.1016/j.ejmech.2012.10.020.
- Caballero J, Zilocchi S, Tiznado W, Collina S (2012) Docking and quantitative structure–activity relationship studies for imidazo[1,2-a]pyrazines as inhibitors of checkpoint kinase-1. *Med Chem Res* 21: 1912–1920. doi:10.1007/s00044-011-9714-1.
- Caballero J, Alzate-Morales JH (2012) Molecular dynamics of protein kinase-inhibitor complexes: a valid structural information. *Curr Pharm Des* 18: 2946–2963. doi:10.2174/138161212800672705.
- Caballero J, Quiliano M, Alzate-Morales JH, Zimic M, Deharo E (2011) 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* 25: 349–369. doi:10.1007/s10822-011-9425-1.
- Alzate-Morales J, Caballero J (2010) Computational Study of the Interactions between Guanine Derivatives and Cyclin-Dependent Kinase 2 (CDK2) by CoMFA and QM/MM. *J Chem Inf Model* 50: 110–122. doi:10.1021/ci900302z.
- Caballero J, Alzate-Morales JH, Vergara-Jaque A (2011) Investigation of the differences in activity between hydroxycycloalkyl N1 substituted pyrazole derivatives as inhibitors of B-Raf kinase by using docking, molecular dynamics, QM/MM, and fragment-based de novo design: study of binding mode of diastereomer compounds. *J Chem Inf Model* 51: 2920–2931. doi:10.1021/ci200306w.
- Munoz C, Adasme F, Alzate-Morales JH, Vergara-Jaque A, Knies T, et al. (2012) Study of differences in the VEGFR2 inhibitory activities between semaxanib and SU5205 using 3D-QSAR, docking, and molecular dynamics simulations. *J Mol Graph Model* 32: 39–48. doi:10.1016/j.jmgm.2011.10.005.
- Caballero J, Zilocchi S, Tiznado W, Collina S, Rossi D (2011) Binding studies and quantitative structure–activity relationship of 3-amino-1H-indazoles as inhibitors of GSK3 $\beta$ . *Chem Biol Drug Des* 78: 631–641. doi:10.1111/j.1747-0285.2011.01186.x.
- Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, et al. (2004) Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J Med Chem* 47: 1739–1749. doi:10.1021/jm030643o.
- Suite 2011: LigPrep, version 2.5, Schrödinger, LLC, New York, NY., (2011).
- SYBYL, version 7.3, Tripos Inc, St. Louis, MO. (2006).
- Bruce L, Bush RBNJ (1993) Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA. *J Comput Aided Mol Des* 7: 587–619.
- Ojha PK, Mitra I, Das RN, Roy K (2011) Further exploring rm2 metrics for validation of QSPR models. *Chemom Intell Lab Syst* 107: 194–205. doi:10.1016/j.chemolab.2011.03.011.
- Fang J, Huang D, Zhao W, Ge H, Luo H-B, et al. (2011) A New Protocol for Predicting Novel GSK-3 $\beta$  ATP Competitive Inhibitors. *J Chem Inf Model* 51: 1431–1438. doi:10.1021/ci2001154.