Journal of Computer-Aided Molecular Design
https://doi.org/10.1007/s10822-018-0129-7

Theoretical models to predict the inhibitory effect of ligands of sphingosine kinase 1 using QTAIM calculations and hydrogen bond dynamic propensity analysis

Marcela Vettorazzi[1,2] · Cintia Menéndez[3] · Lucas Gutiérrez[1,2] · Sebastián Andujar[1,2] · Gustavo Appignanesi[3] · Ricardo D. Enriz[1,2]

Received: 19 March 2018 / Accepted: 2 July 2018

Abstract
We report here the results of two theoretical models to predict the inhibitory effect of inhibitors of sphingosine kinase 1 that stand on different computational basis. The active site of SphK1 is a complex system and the ligands under the study possess a significant conformational flexibility; therefore for our study we performed extended simulations and proper clusterization process. The two theoretical approaches used here, hydrogen bond dynamics propensity analysis and Quantum Theory of Atoms in Molecules (QTAIM) calculations, exhibit excellent correlations with the experimental data. In the case of the hydrogen bond dynamics propensity analysis, it is remarkable that a rather simple methodology with low computational requirements yields results in excellent accord with experimental data. In turn QTAIM calculations are much more computational demanding and are also more complex and tedious for data analysis than the hydrogen bond dynamic propensity analysis. However, this greater computational effort is justified because the QTAIM study, in addition to giving an excellent correlation with the experimental data, also gives us valuable information about which parts or functional groups of the different ligands are those that should be replaced in order to improve the interactions and thereby to increase the affinity for SphK1. Our results indicate that both approaches can be very useful in order to predict the inhibiting effect of new compounds before they are synthesized.

Keywords Sphingosine kinase inhibitors · QTAIM calculations · Hydrogen bond dynamic propensity analysis · Theoretical approaches

Introduction

Sphingosine-1-phosphate is a potent sphingolipid mediator [1–4], and the kinase that produces it, Sphingosine Kinase 1 (SphK1) has been implicated in cancer progression, inflammation and cardiovascular diseases [5, 6].

The first crystal structure of SphK1 was reported in 2013 [7], such report presents the atomic structure of SphK1 in complexes that expanded the mechanistic understanding of this important pharmacological target. This information has provided a crucial baseline from which to generate predictions and targeted modifications that further probe the many nuances of functional elements in SphK1 activity and regulation. More recently two SphK1 co-crystal structures (4L02 and 4V24) with potent inhibitory compounds have also been published [8, 9]. These structures provide useful structural information on the interaction of ligands at the active site of SphK1. The structure of SphK1 provides a structural framework to understand how many unrelated residues are involved. Within the cavity, sphingosine is proposed to be guided into position by a tunneling mechanism that is driven by energetically favorable interactions between nonpolar cavity residues and aliphatic carbons long the sphingosine tail group [10]. The pocket itself is revealed not as a

✉ Ricardo D. Enriz
denriz@unsl.edu.ar

[1] Facultad de Química, Bioquímica y Farmacia, Universidad Nacional de San Luis, Chacabuco, 915, 5700 San Luis, Argentina

[2] IMIBIO-CONICET, UNSL, Chacabuco 915, 5700 San Luis, Argentina

[3] INQUISUR, Departamento de Química, Universidad Nacional del Sur (UNS)-CONICET, Av. Alem 1253, 8000 Bahía Blanca, Argentina

Published online: 03 July 2018

simple tunnel but is rather J-shaped, allowing bulky groups like those possessed by the inhibitors **SKI-II** [11], **1v2** [8] and **PF-543** [12] nestle within the cavity and competitively occlude substrate binding [10].

Although the crystal structure of SphK1 is a stepping key to new perspectives in our quest to resolve the cellular function and regulation of this key enzyme, however, as is often the case, from this information new questions and doubts are arising. Further structural studies might explain the seemingly contradictory observation that some SphK1 inhibitors [13, 14] markedly affect cell growth and survival, whereas a potent and selective inhibitor did not [12].

To further improve SphK inhibitors, particularly regarding potency and selectivity, it is necessary to continue to probe the binding pockets of SphK1 with structurally diverse sets of molecules. In this line we have recently reported two new series of compounds possessing different structural scaffolds [15]. Such compounds were obtained from a virtual screening and among these new compounds; molecules **1–3** (Fig. 1) displayed the strongest inhibitory effects in this series. Although the inhibitory activity of these compounds is only moderated in comparison to **PF-543**, it should be noted that such compounds were obtained from a primary virtual screening and therefore the activity obtained is more than acceptable. It is clear that one of our main objectives is to increase the potency of these new compounds through the introduction of structural changes. However the question that arises is: what are the changes that must be made on these structures, and how we can know before their synthesis if these compounds are going to be as active as we expect. This task demands the availability of accurate theoretical approaches to assess the inhibitory activity of potential inhibitors. A first step in this regard would be to test the efficiency of some existing techniques we have already applied in other contexts in the evaluation of the inhibitory activity of our new compounds on SphK1 along with other already existing inhibitors. If we are successful in correlating their theoretical predictions on SphK1 with experimental data, these techniques might be useful for selecting new compounds generated from such scaffolds with potentially improved inhibitory effect prior to their synthesis. Moreover, we also expect that the knowledge gained from the application of such techniques might provide us with additional clues to re-engineer this new family of compounds.

Computational methods

Receptor preparation

The X-ray crystal structure of SphK1 was taken from the Protein Data Bank with the accession code 3VZD. The C chain was used. The SKI-II inhibitor and the water molecules were removed; only the ADP, Mg^{2+} and a water molecule (that makes hydrogen-bond interactions with Asp178 and Ser168) is left since it is considered structural part of the

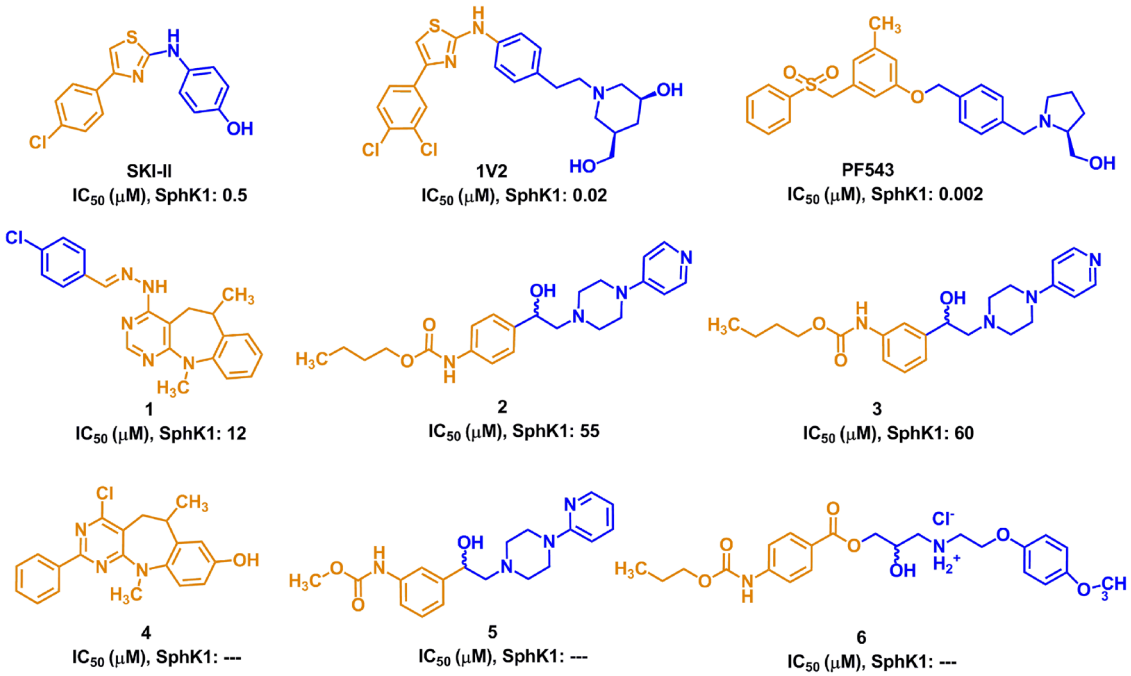

Fig. 1 Structural features of the compounds studied here. The polar head of each molecule has been marked in blue and their respective hydrophobic tails are shown in orange color (online version to see the colors)

protein [7]. Ionizable groups were assumed as its ionization state at pH 7.0.

Docking procedure

The docking simulations were carried out using of AutoDock 4.2 [16]. In all experiments the following parameters were used: the initial population of trial ligands was constituted by 200 individuals; the maximum number of generations was set to 270,000. The maximum number of energy evaluations was 10.0×10^6. All other run parameters were maintained at their default setting. The 3D affinity map was a cube with $46 \times 46 \times 48$ points separated by 0.375 Å and centered at the active site of SphK1. The resulting docked conformations were clustered into families by the backbone Root-Mean-Square Deviation (RMSD). The lowest docking-energy conformation of each family was considered the most favorable orientation.

Refinement of the anchoring

After the docking calculations, leading lowest energy structures were refined by performing molecular dynamics simulations, using the Amber14 packages [17]. The molecular dynamics simulations (MD) were performing using the all-atom force field ff99SB [18] to describe the receptor whereas the general Amber force field (GAFF) [19] was used to handle small organic molecules and the force field parameters of the inhibitors were produced by the antechamber program in Amber. Each model was soaked in a truncated octahedral periodic box of TIP3P water molecules [20]. The distance between the edges of the water box and the closest atom of the solutes was at least 10 Å. Sodium ions were added to neutralize the charge of the system. The entire system was subjected to energy minimization.

To remove possible bumps, the geometry of the system went through an energy minimization process with 10,000 steps of a conjugate gradient method: (i) In the first 5000 steps, only the backbone atoms of the complex were constrained with 10 kcal/(mol Å^2) force constants. (ii) In the last 5000 steps, the solute and solvent atoms were allowed to move without any constraint. The relaxed geometry resulted in a backbone RMSD < 0.5 Å from the reference crystal structure.

In the next place each system was then heated in the NVT ensemble from 0 to 300 K in 500 ps and equilibrated at an isothermal isobaric (NPT) ensemble for another 2 ns. A Langevin thermostat [21] was used for temperature coupling with a collision frequency of 1.0 ps^{-1}. The particle mesh Ewald (PME) method was employed to treat the long-range electrostatic interactions in a periodic boundary condition [22]. The SHAKE method was used to constrain hydrogen atoms. The time step for all MD is 2 fs, with a direct-space, non-bonded cutoff of 8 Å. Three MD simulations of 50 ns were conducted for each system under different starting velocity distribution functions; thus, in total 150 ns were simulated for each complex.

It should be noted that compounds **2, 3, 5** and **6** possess one chiral center, and are therefore enantiomeric with the possibility of two isomers (R and S). However, we did not perform an enantiomeric resolution for the previously reported biological assays [15]; thus, the racemic mix was used in each case. For the molecular modeling study, only one isomer of each compound was evaluated in our MD simulations and later computations. To choose the isomeric forms of each compound, we considered preliminary and specially performed exploratory simulations determining the spatially preferred form for these compounds. Our preliminary and exploratory docking and short MD simulations (three runs of 5 ns each) indicate that the spatial ordering adopted by the R-forms gives a more adequate orientation of the molecules to interact in the active site of SphK1. Thus, on the basis of such results the R-forms were chosen for the simulations.

Binding energy calculations

The MM-PBSA and MM-GBSA protocol was applied to each MD trajectory in order to calculate the relative binding energies of the SphK1–INH complexes. The MM-PBSA and MM-GBSA method was used in a hierarchical strategy, and the details of this method have been presented elsewhere [23]. This protocol was applied to 27,000 equidistant snapshots extracted from the last 45.0 ns of the dynamics in triplicate and was used within the one-trajectory approximation. Briefly, the binding free energy (ΔG_{bind}) resulting from the formation of a RL complex between a ligand (L) and a receptor (R) is calculated as:

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

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{electrostatic} + \Delta E_{vdW} \quad (2)$$

$$\Delta G_{sol} = \Delta G_{PB} + \Delta G_{SA} \quad (3)$$

where ΔE_{MM}, ΔG_{sol}, and $-T\Delta S$ are the changes in the gas-phase MM energy, the solvation free energy, and the conformational entropy upon binding, respectively. ΔE_{MM} includes $\Delta E_{internal}$ (bond, angle, and dihedral energies), $\Delta E_{electrostatic}$ (electrostatic), and ΔE_{vdw} (van der Waals) energies. ΔG_{solv} is the sum of electrostatic solvation energy (polar contribution), ΔG_{PB}, and the non-electrostatic solvation component (nonpolar contribution), ΔG_{SA}. Polar contribution is calculated using the PB model, while the nonpolar energy is estimated by solvent accessible surface area. The conformational entropy change $-T\Delta S$ is usually computed by normal-mode analysis, but in this study the entropy contributions

were not calculated due to the computational cost involved in such calculations.

Cluster

To perform clustering we used the Cpptraj program, part of the AmberTools package. 135 ns was analyzed for each complex, discarding the initial 5 ns of each dynamic. A total of 27,000 frames was evaluated. Ten representative pdb structures was obtained for each complex from a separation by using RMSD, which was ordered from highest to lowest population. For each complex, several representative pdb structures were analyzed from QTAIM methods; thus, we calculate the necessary number of structures to reach 50% of the population.

Atoms in molecules theory

The wave functions of the inhibitors bound to the binding site residues residues that have at least one heavy atom within 5 Å from the ligand molecule (first shell residues), generated at the M062X/6-31G(d) level of theory, were subjected to a Quantum Theory Atoms In Molecules (QTAIM) analysis [24] using Multiwfn software [25]. This type of calculations have been used in recent works because it ensures a reasonable compromise between the wave function quality required to obtain reliable values of the derivatives of $\Sigma\rho_{(r)}$ and the computer power available, due to the extension of the systems in study [26–29].

For each complex, several pdb structures corresponding to 50% of the population were analyzed. The $\Sigma\rho(r)$ obtained for each interaction between Receptor and Ligand (R–L) was averaged and then all the interactions were added to obtain a final value of $\Sigma\rho(r)$.

Hydrogen bond dynamic propensity analysis

Dynamic propensities (fraction of time formed) of BHBs belonging to the binding site of SphK1 were calculated over a total of 27,000 MD configurations (three replicas of 45 ns), in all complexes studied. At each evaluation time (configuration), if a pair of residues *i* and *j* of a certain BHBs satisfy a hydrogen bonding criterion (N–O cutoff distance, $r < 3.5°$ Å; N–H–O cutoff angle, $\theta > 140°$), the corresponding interaction becomes 1, while it is 0 otherwise. Then, we averaged the results for all configurations evaluated and so we obtained the fractions of time that each BHBs remain formed in each case. Finally, for each BHB we calculated D-values how the difference between dynamic formation propensity and its corresponding state-value (formed or not formed according to the same criterion of distance and angle) in the PDB structure (in absolute value). For more details see "Results and discussion" section.

Results and discussion

Nine compounds were selected for this study: three well-known inhibitors (**PF-543, 1v2** and **SKI-II**) for which experimental structural data are available (Protein Data Bank (PDB) code 4V24 [9], 4L02 [8] and 3VZD [7], respectively) and six compounds recently reported by us [15]. Among these, three active compounds (**1**–**3**) as positive controls and three inactive compounds (**4**–**6**) were selected to provide negative controls (Fig. 1).

We have previously performed molecular dynamics simulations in different biological systems with different degrees of structural complexity like for example D2 dopamine receptor (D2DR) [30–34], dihydrofolate reductase (DHFR) [35, 36], Acetylcholinesterase (AchE) [37, 38], beta secretase (BACE 1) [39, 40]. In the case of simulations at the D2 dopamine receptor, three simulations of 10 ns each have been shown to be sufficient to obtain satisfactory results [30–34]. However, the degree of structural complexity and other factors that have a direct influence on this type of simulations are different for the active site of Sphingosine Kinase 1 with respect to the active site of D2DR. A comparative study between the active sites of these two receptors is summarized in Table 1; such study was carried out analyzing the following aspects: (a) size and depth of the

Table 1 Parameters considered analyzing both the active sites of the molecular targets as well the R–L complexes

	D2	SphK1
Solvent accessible surface area (SASA) (Å^2)	489	721
Buried area (Å^2)	428	675
Number of mainly interactions stabilizing or destabilizing the complexes	From 3 to 6	More than 10
Flexibility of the active site	Markedly reduced (Tight)	Large
Flexibility of the ligand	Partially restricted (two rotations)	Very flexible (more than 10 rotations)
Types of interactions involved in the complex formation	Ionic, hydrogen bonds and hydrophobic	Ionic, pi stacking and hydrophobic
Structural variations on the ligands	From scarce to significant	High variability

active site; (b) number of the main interactions stabilizing or destabilizing the complexes; (c) flexibility of the active site; (d) flexibility of the ligand; (e) types of interactions involved in the complex formation and (f) structural variability of the different ligands.

A numerical study on the Solvent Access Surface Area (SASA) [41] was used to evaluate the size and depth (buried) of the active site. As can be seen in Table 1, the SASA and buried obtained for Sphingosine Kinase 1 is almost twice to those obtained for D2DR. This information by itself is significant to see that the complexity of the active site of Sphingosine Kinase 1 is much greater than that of D2DR. However, all the aspects analyzed here show that the Receptor–Ligand (R–L) complexes in Sphingosine Kinase are more intricate and complicated to study than those R–L complexes of D2DR. Those that show greater differences are the number of interactions involved as well as the flexibility and structural variability of the ligands (see Table 1). On the basis of this information, we decided to extend the simulations for this system performing three simulations of 50 ns each (150 ns for the three full simulations).

Our first attempt was try to find a correlation between the binding energies obtained through the Generalized Born surface area (GBSA) and Poisson–Boltzmann surface area (PBSA) calculations with the experimental data (IC_{50}) (Fig. 2). Unfortunately the results by using GBSA were very poor. Not only because the value of r is very low (r = 0.53), but also because these results do not allow to clearly differentiate between the most active compounds with respect to those less active. Note, for example, that the SKI-II compound that is very active has a very high energy value. In turn PBSA calculations give a better result with respect to GBSA giving an r value of 0.69; however such result is still insufficient. It should be noted that the bad fitting when using GBSA and PBSA methods might partly due to the lack of inclusion of entropy. As previously indicated, the active site of SphK1 is large and has some flexibility; furthermore, the ligands by themselves are also quite flexible and can be moved sufficiently before to fit into the binding pocket. Thus, the entropy contribution could be significant for this particular enzyme and therefore at least in part the unsatisfying performance of these calculations might because entropy is turned off and not because of the method itself. However it is clear that these theoretical techniques without consideration of entropy fail to show a good performance in predicting the activity of new compounds (not yet synthesized) for

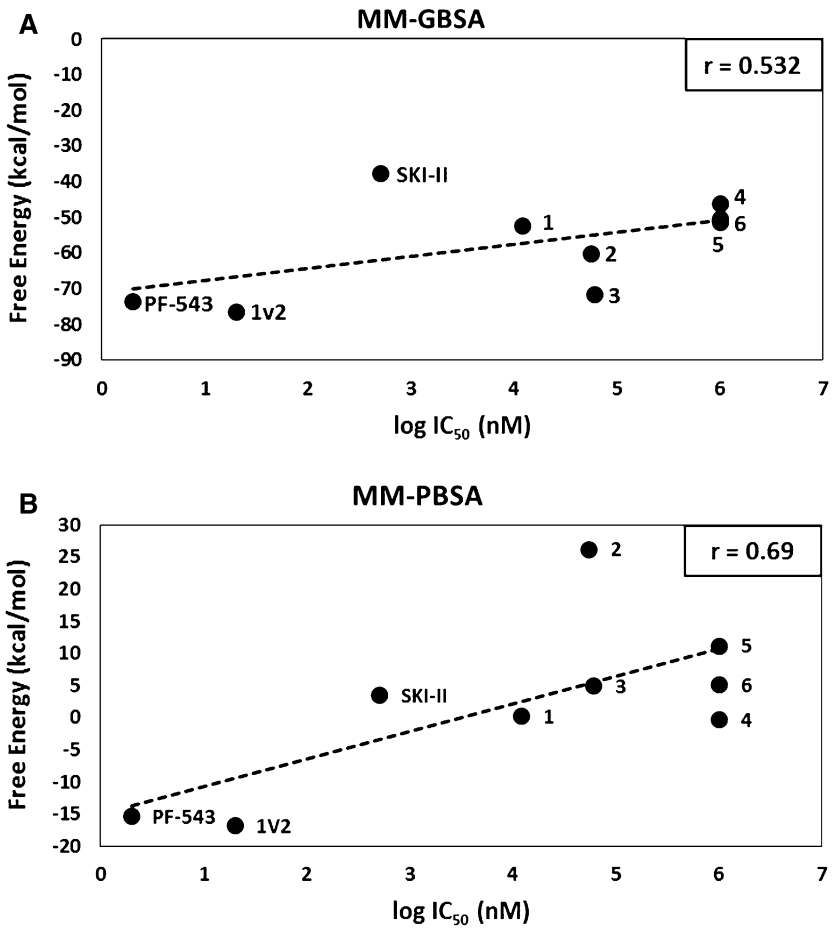

Fig. 2 Correlation obtained for the different compounds between the experimental data (IC_{50}) and the binding energy obtained from **a** MM-GBSA and **b** MM-PBSA calculations

this molecular target. Thus, we decided to try two different techniques that have given satisfactory results in other biological systems. Thus, we used Quantum Theory of Atoms In Molecules (QTAIM) [24] calculations and also performed a study of the dynamic propensity of backbone hydrogen bonds [42, 43].

QTAIM calculations

Studies using QTAIM calculations have been carried out by our research group with great success in various biological systems [31–36, 39, 40]. In fact QTAIM calculations have already been used previously in complexes of Sphingosine Kinase 1, but in that case they were used for another purpose [15]. In reference 15 the QTAIM calculations were used only to analyze the different interactions that stabilized the complexes of the most representative compounds of the series and compare them with the interactions of the complex SphK1/**PF-543**.

It is important to note that QTAIM calculations are performed on a determinate geometry, while the R–L binding come from an ensemble of conformations. Therefore, to understand the activity of these compounds it would be convenient to analyze several conformations of each R–L complex. Thus, the first doubt that arose was how many different structures for each complex would be necessary consider. To solve this problem, we performed a clusterization process on the trajectories of MD. As preliminary and exploratory analysis, we made different tests with the three complexes of which we had crystallographic data (**SKI-II, 1v2** and **PF-543**). The results obtained for these exploratory calculations are shown in Table 2.

Although all the correlations obtained between the experimental data and the QTAIM analysis for these three compounds were acceptable, significant differences were found depending on how many structures were considered. If we only consider the three structures reported experimentally, the result obtained is just acceptable (r = 0.74). This result improves if we take into account the more populated file of the cluster (r = 0.86). In turn, the correlation improves significantly if we consider 50% of the files of the clusterization process, obtaining an r value of 0.95. It is interesting to note that this correlation is good enough to consider that it would have good predictive power. If we consider 100% of the cluster, the value improves even more (r = 0.98). It should be noted, however, that performing these calculations considering 100% of the cluster represents highly demanding task in terms of calculation time and the analysis of results. Therefore we consider that using 50% of the cluster is a good compromise between calculation accuracy degree and computational demand. Based on the above results, the QTAIM study was carried out considering 50% of the clustering process (for more details see methods section).

Figure 3 shows that an excellent correlation was obtained (r = −0.95) for the complete series by using QTAIM calculations. This study allows to differentiate perfectly between the most active compounds (**SKI-II, 1v2** and **PF-543**), compounds with moderate activity (**1, 2** and **3**) and the inactive compounds (**4, 5** and **6**). On the other hand, the obtained correlation is good enough to expect good predictive power for this type of ligands. It is important to note that the QTAIM study also allows us to determine which portions of the molecules should be modified to increase their affinity for SphK1. This can be observed from Fig. 4. In this figure the interactions produced by the polar part of the molecules

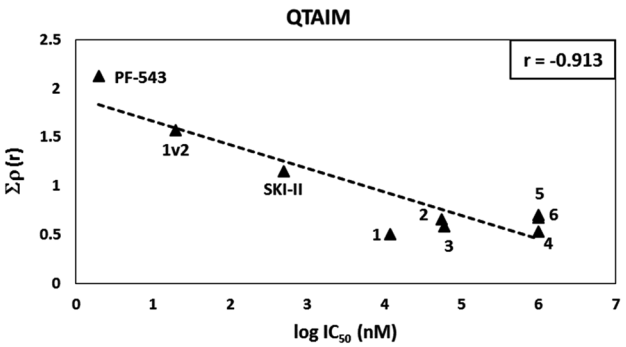

Fig. 3 Correlation obtained by using QTAIM calculations for the whole series

Table 2 Different correlations between experimental data (IC_{50}) and Σρ(r)(a.u.) values (QTAIM calculations) obtained for compounds SKI-II, 1v2 and PF-543

Ligand	IC_{50} (nM)	log IC_{50}	Crystallographic structure	More populated[a]	50% CLUSTER[b]	100% CLUSTER[c]
SKI-II	500	2.699	0.227	0.178	0.425	1.146
1v2	20	1.301	0.495	0.478	0.716	1.567
PF-543	2	0.301	0.413	0.441	1.412	2.123
r			−0.743	−0.858	−0.947	−0.985

[a]More populated corresponds to the structure of the leader of the most populated cluster

[b]50% cluster corresponds to the leading structures that represent 50% of the population

[c]100% cluster corresponds to the 10 structures obtained through the clustering process

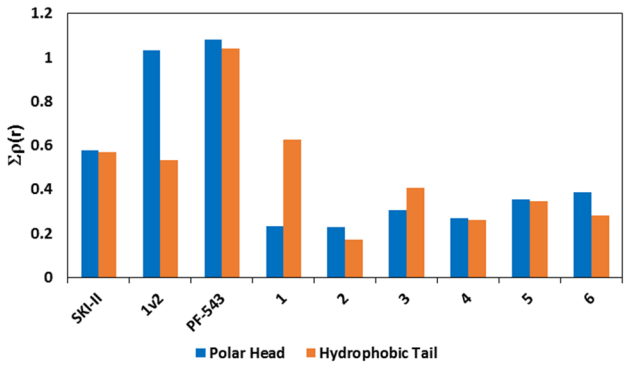

Fig. 4 Charge density values for the total interactions of the polar head (blue stacked bars) and the hydrophobic portion (orange stacked bars) for **SKI-II, 1v2, PF-543** (more active compounds), compounds **1**–**3** (less active molecules) and compounds **4**–**6** (inactive molecules). The repulsive short C–H●●●H–C contacts were not included (Online version to see the colors)

(blue color) have been differentiated from those produced by the hydrophobic tail of the compounds (orange color). Observing this figure it is evident that the active compounds have stronger interactions than the less active compounds. It is interesting to note that the compound **PF-543** has very strong interactions both in the area of the cationic head and in the hydrophobic tail. This clearly explains its high affinity for the active site of Sphingosine Kinase 1 and therefore its inhibitory activity (it is the strongest inhibitor reported to date). Comparing the Σρ(r) values obtained for **1v2** and **PF-543** it can be seen that the interactions in the zones of the cationic heads of these compounds are comparable; however there is a noticeable difference in the interactions of the hydrophobic tail. In this case, the interactions observed for **PF-543** are significantly stronger than those of **1v2** and therefore they are explaining the stronger activity of **PF-543**.

If we compare the Σρ(r) values obtained for **SKI-II** and **PF-543**, it is evident that both the cationic head zone and the hydrophobic tail of **PF-543** establish more interactions and stronger ones than those obtained for **SKI-II**, which is in total agreement with the experimental data. On the other hand, the compounds previously reported by us (**1**–**3**) present significantly weaker interactions in both portions of the receptor, indicating that it is necessary to introduce structural changes in both parts of the ligand in order to obtain a greater inhibitory activity. Once again these results are supported by the experimental data.

Considering the most active compound reported by our group (compound **1**) it can be seen that the interactions in the hydrophobic tail of this compound are more than acceptable. In fact, these interactions are greater than those found for **1v2** and **SKI-II**. However, the interactions obtained for the polar head of compound **1** are markedly weaker than those found for **1v2** and **SKI-II** and this would explain the lower inhibitory effect of compound **1** with respect to **1v2** and **SKI-II**. In contrast, the results obtained for compounds **2** and **3** suggest that the interactions that should be increased are those of the hydrophobic portion (orange zone) (Figs. 4, 5).

Regarding the results obtained for the inactive compounds (**4**–**6**), it is possible to observe that, in general, the interactions of these compounds are the weakest. The case of compound **4** is very particular, because it adopts a different spatial arrangement from the rest of the compounds and can only establish interactions with the hydrophobic zone of the active site (orange zone in Fig. 5). This different behavior could explain its lack of activity.

If we compare the interactions obtained for compounds **5** and **6** with the more active compounds (**SKI-II, 1v2** and **PF-543**), the differences are very marked. However, this difference is only subtle when the interactions are compared with those of compound **3**. In fact, although the interactions of compound **3** are slightly larger, this difference is very small. There are at least two possible explanations for this situation: (a) it should be noted that although compound **3** is considered active, its activity could be considered almost marginal ($IC_{50} = 60$ μM) (b) these compounds are chiral. Whereas racemic mixtures were used for the bioassays, only the R enantiomer, which is the preferred form (see methods section), was used for the simulations. This situation could introduce some "noise" that is not possible to evaluate so satisfactorily through these simulations. Despite this, it is evident that the QTAIM analysis gives an excellent correlation with the experimental data. This information is extremely useful, since it indicates which part of the molecule should be modified in order to increase its inhibitory activity against SphK1.

Dynamic propensity study of backbone hydrogen bonds

As a supplementary approach, we now study the dynamic propensity of the backbone hydrogen bonds of the protein SphK1 in order to determine its binding properties. This approach has been fostered by the observation that certain backbone hydrogen bonds (BHB), mainly those belonging to binding sites in several apo proteins, exhibit a dynamical propensity in simulations that differs markedly form their state-value (that is, formed/not formed) in the reported PDB structure. This fact makes them appealing as targets for protein local stabilization upon ligand binding [42, 43]. The dynamic propensity study of BHBs represents a simple method that has been accounted for in detail in previous works [42, 43], but for the sake of completeness we briefly describe it in what follows: Specifically, for each BHB of the apo protein we calculate the difference (D value) between its dynamics formation propensity (fraction of time formed

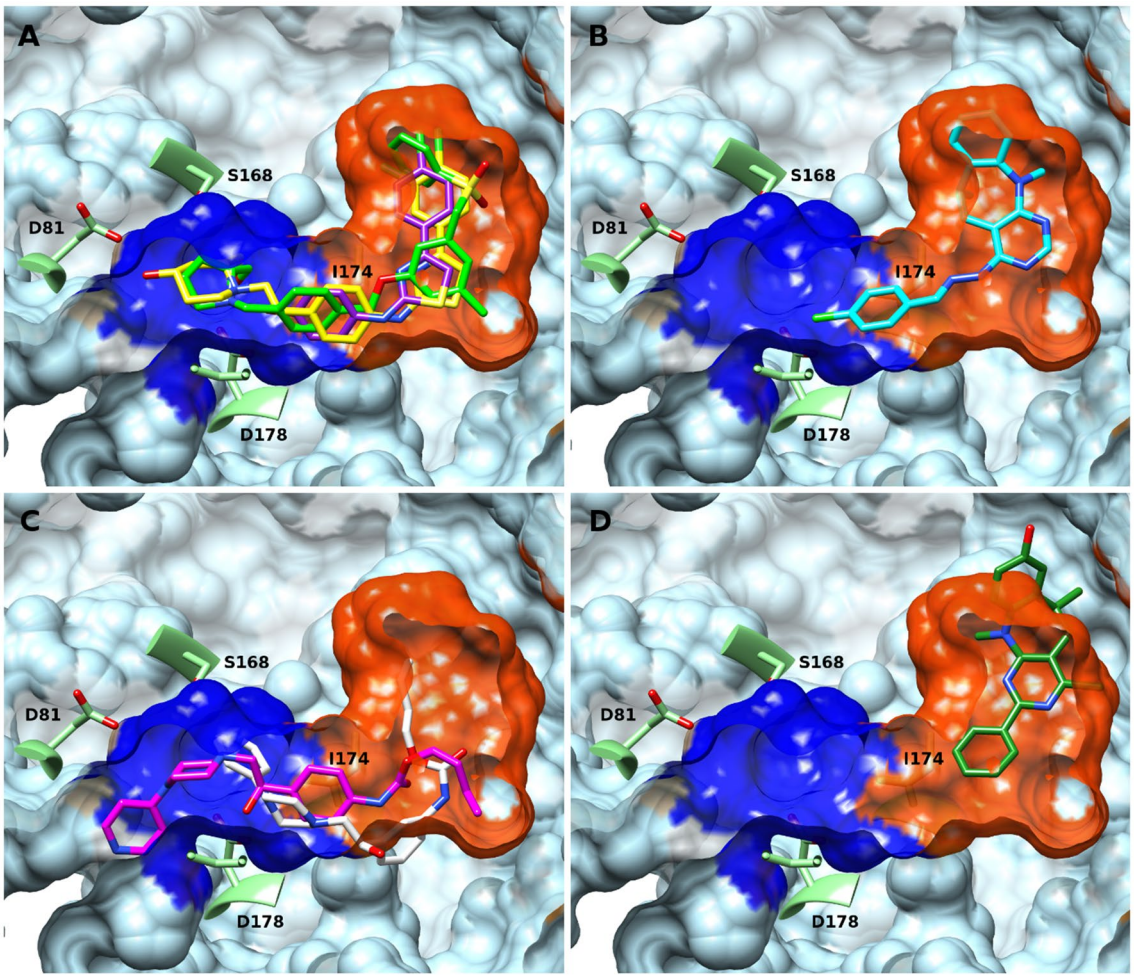

Fig. 5 Spatial view of the different ligands bonded in the binding pocket of SphK1. The blue and orange zones represent the cationic and hydrophobic portions of the active site, respectively. **a PF-543** (in green), **1v2** (in yellow) and **SKI-II** (in purple) superimposed. **b** Compound **1** (in cyan). **c** Compound **2** (magenta) and **3** (white) superimposed. **d** Compound **4** (dark green). Some of the main amino acids involved in the formation of the complexes receptor-ligand are also shown in this figure. The structures were taken from a clustering process in which the last 45 ns of each of the three simulations were considered. (online version to see the colors)

during a molecular dynamics trajectory) and its corresponding PDB-structure state-value. By analyzing the D values it is usually found that most BHBs of the PDB are stable during the dynamics (low distance value, D). However, certain BHBs do exhibit large D-values, corresponding either to interactions that are present in the PDB and tend to be disrupted during (part of) the dynamics or, less frequently, to interactions that while absent in the PDB are nonetheless persistently formed during the dynamics. Then, if the absolute value of D for a given BHB is above ½, we call it a C-HB reflecting the fact that such BHB tends to "change it state" from formed to mostly disrupted or from not formed to mostly formed between the PDB structure and the dynamics [42, 43]. Additionally, studies on the location of such C-HBs for a series of proteins have detected a clear enrichment of them in binding sites. In turn, previous studies on a series of protein–protein or protein–ligand complexes showed that most C-HBs are strongly stabilized in the complex form compared to the apo form, thus pointing to the fact that stabilization of C-HBs provides a driving force for binding [42]. In this sense, C-HBs embody motifs or regions of the protein whose stabilization depends on the establishment of a proper local context (usually involving water removal) provided by the ligand upon binding. Indeed, C-HBs stabilization has been proven in a quantitative way for a series of protein–protein and protein-drug complexes, obtaining good correlations with biding affinity experimental values [42]. It is interesting to point out that this method provides a complementary approach since it does not make emphasis on *direct* protein–ligand interactions, but rather on the local stabilization of the target protein binding site upon ligand binding.

We should note that in the present study it has not been possible to determine C-HBs exactly in the same way as in previous ones [42, 43] since we lack a comparable SphK1 apo structure. Indeed, there exist no experimentally obtained (NMR or X-ray) structure of apo SphK1 at exactly the same conditions as that for the different complexes reported and used in the present study (crystallized with ADP and Mg^{2+}). Thus, we generated an apo structure by removing the inhibitor SKI-II from the complex and we determine BHBs belonging to the binding site as that whose distance (measured form the N amide or the carbonyl O) to any heavy atom of the SKI-II in the complex (PDB: 3VZD) is less than 6 Å. Thereby, dynamic propensities of such BHBs were calculated over a total of 27,000 MD configurations (three replicas of 45 ns). Then all those BHBs that are found to be formed in less than half of the configurations analyzed (that is, D-value > 0.5) will be C-HBs. Table 3 shows the C-HBs considered, together with their corresponding D-values. It is important to mention that the BHB 303–299 was included although it does not meet strictly the established cut-off value since in two of the replicas its dynamic propensity is less than 0.5, while in the other one it is higher.

Subsequently, dynamic propensities for C-HBs in the complexes between SphK1 and the most active compounds were evaluated: **SK-II, 1v2** and **PF-543**. Table 4 shows DM values (average values of D), that is for each C-HBs we calculate the D-values and, then by averaging such quantities, we provide the DM-value. This table displays an excellent correlation between the experimental binding affinity data and DM-values for these three compounds. Dynamic propensities were evaluated in the same period of time that QTAIM calculations.

The Fig. 6 show also an excellent correlation for the complete series of compounds. This study, like the QTAIM study, enables us to differentiate the most active compounds (**SKI-II, 1v2** and **PF-543**) from compounds with moderate activity (**1, 2** and **3**) and inactive compounds (**4, 5** and **6**). However it is important to remark that the times demanded for the computer calculations as well as for the analysis of the results are significantly inferior in comparison to those required for the QTAIM analysis and therefore this technique could be very adequate for an exploratory analysis in this type of compounds.

Table 3 D-values for C-HBs belonging to binding site of SphK1

C-HBs		SphK1: APO form
nRES	nRES	D-values
178	174	0.841
181	177	0.561
272	269	0.762
303	299	0.391
302	298	0.546

Table 4 Correlation between experimental data (IC_{50}) and DM values obtained for compounds SKI-II, 1v2 and PF-543

Compound	IC_{50} (nM)	log (IC_{50})	DM (5–50 ns)
SKI-II	500	2.69897	0.532
1v2	20	1.30103	0.470
PF-543	2	0.30103	0.370
r			0.9736

Conclusions

Being able to predict the inhibitory activity of new compounds before being synthesized is certainly a highly desirable goal. For the specific case of SphK1 complexes, we here present two theoretical methods that stand on different computational basis that, in both cases yield excellent correlations with the experimental data. From these results, it is reasonable to expect that such methods might also present predictive ability for other structurally related compounds that have not been synthesized yet. In the case of the hydrogen bond dynamics propensity analysis, it is remarkable that a rather simple methodology with low computational requirements yields results in excellent accord with experimental data. In this respect, the excellent performance of QTAIM would be a priori more expected in terms of its more accurate description of the molecular interactions.

Comparing the two techniques used here, it is evident that QTAIM calculations are much more computational demanding (time consuming). With regard to CPU time, the relationship between the hydrogen bond dynamic propensity analysis and QTAIM calculations is 1 h versus 10 hs. But also QTAIM analysis requires a longer time in the preparation of the model system to study (prior to

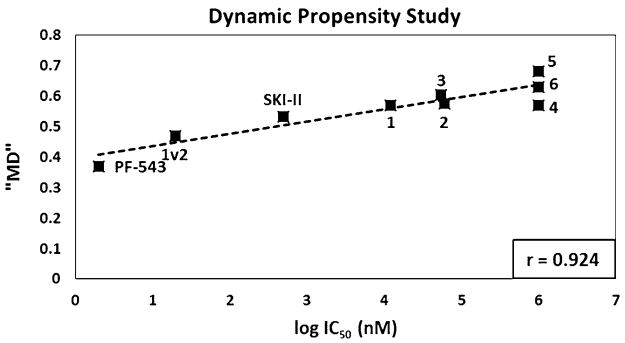

Fig. 6 Correlation between DM-value and experimental data (IC_{50}) for the whole series

the calculation) and especially in the final analysis of the interactions (after the calculation). Both processes are much more complex and tedious than those required for the hydrogen bond dynamic propensity analysis. However, this greater computational effort is justified because the QTAIM study, in addition to giving an excellent correlation with the experimental data, also gives us valuable information about which parts or functional groups of the different ligands are those that should be replaced in order to improve the interactions and thereby to increase the affinity for SphK1. Possibly, the main limitation in the use of QTAIM calculations for the analysis of interactions R–L, is that it is necessary to have the correct geometries for these complexes. It is good to keep in mind that the QTAIM study involves static calculations that are used to evaluate a dynamic process. Therefore, having many geometries (or at least the most representative) of each complex is fundamental, in order to obtain satisfactory results. In our study in order to guarantee this information, extensive DM simulations were carried out and a clusterization process was performed for the different complexes studied. It should be noted that the active site of SphK1 can be considered a relatively complex system; in turn the conformational flexibility of the ligands and the structural variety of the compounds impose a high degree of difficulty for molecular modeling of the systems under study. This requires performing extended simulations and proper clusterizations if we pretend to obtain satisfactory results.

In short, we report here two different theoretical approaches that exhibit excellent correlations with the experimental data and that we believe can be very useful in order to predict the inhibiting effect of new compounds before they are synthesized. At this moment we are analyzing and synthesizing new structurally related compounds by using the theoretical approaches reported here.

Acknowledgements Grants from Universidad Nacional de San Luis (UNSL-Argentina) partially supported this work. This work was supported in part by a grant from MinCyt (PICT-2015/1769). GAA and CAM acknowledge financial support from CONICET, UNS and MinCyt (PICT-2015/1893).

References

1. Maceyka M, Payne SG, Milstien S, Spiegel S (2002) Sphingosine kinase, sphingosine-1-phosphate, and apoptosis. Biochim Biophys Acta 1585:193–201
2. Hait NC, Oskeritzian CA, Paugh SW et al (2006) Sphingosine kinases, sphingosine 1-phosphate, apoptosis and diseases. Biochim Biophys Acta 1758:2016–2026
3. Taha TA, Mullen TD, Obeid LM (2006) A house divided: ceramide, sphingosine, and sphingosine-1-phosphate in programmed cell death. Biochim Biophys Acta 1758:2027–2036. https://doi.org/10.1016/j.bbamem.2006.10.018
4. Wattenberg BW (2010) Role of sphingosine kinase localization in sphingolipid signaling. World J Biol Chem 1:362–368. https://doi.org/10.4331/wjbc.v1.i12.362
5. Ricci C, Onida F, Ghidoni R (2006) Sphingolipid players in the leukemia arena. Biochim Biophys Acta 1758:2121–2132. https://doi.org/10.1016/j.bbamem.2006.06.016
6. Cuvillier O (2002) Sphingosine in apoptosis signaling. Biochim Biophys Acta 1585:153–162
7. Wang Z, Min X, Xiao S-H et al (2013) Molecular basis of sphingosine kinase 1 substrate recognition and catalysis. Structure 21:798–809. https://doi.org/10.1016/j.str.2013.02.025
8. Gustin DJ, Li Y, Brown ML et al (2013) Structure guided design of a series of sphingosine kinase (SphK) inhibitors. Bioorg Med Chem Lett 23:4608–4616. https://doi.org/10.1016/j.bmcl.2013.06.030
9. Wang J, Knapp S, Pyne NJ et al (2014) Crystal structure of sphingosine kinase 1 with PF-543. ACS Med Chem Lett 5:1329–1333. https://doi.org/10.1021/ml5004074
10. Lima S, Spiegel S (2013) Sphingosine kinase: a closer look at last. Structure 21:690–692. https://doi.org/10.1016/j.str.2013.04.006
11. French KJ, Schrecengost RS, Lee BD et al (2003) Discovery and evaluation of inhibitors of human sphingosine kinase. Cancer Res 63:5962–5969
12. Schnute ME, McReynolds MD, Kasten T et al (2012) Modulation of cellular S1P levels with a novel, potent and specific inhibitor of sphingosine kinase-1. Biochem J 444:79–88. https://doi.org/10.1042/BJ20111929
13. Paugh SW, Paugh BS, Rahmani M et al (2008) A selective sphingosine kinase 1 inhibitor integrates multiple molecular therapeutic targets in human leukemia. Blood 112:1382–1391. https://doi.org/10.1182/blood-2008-02-138958
14. Orr Gandy KA, Obeid LM (2013) Targeting the sphingosine kinase/sphingosine 1-phosphate pathway in disease: review of sphingosine kinase inhibitors. Biochim Biophys Acta 1831:157–166. https://doi.org/10.1016/j.bbalip.2012.07.002
15. Vettorazzi M, Angelina E, Lima S et al (2017) An integrative study to identify novel scaffolds for sphingosine kinase 1 inhibitors. Eur J Med Chem 139:461–481. https://doi.org/10.1016/j.ejmech.2017.08.017
16. Morris GM, Goodsell DS, Huey R, Olson AJ (1996) Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. J Comput Aided Mol Des 10:293–304
17. Case DA, Babin V, Berryman JT et al (2014) {Amber 14} OR, University of California, San Francisco
18. Lindorff-Larsen K, Piana S, Palmo K et al (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 78:1950–1958. https://doi.org/10.1002/prot.22711
19. Wang J, Wolf RM, Caldwell JW et al (2004) Development and testing of a general amber force field. J Comput Chem 25:1157–1174. https://doi.org/10.1002/jcc.20035
20. Mark P, Nilsson L (2001) Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. J Phys Chem A 105:9954–9960. https://doi.org/10.1021/jp003020w
21. Izaguirre JA, Catarello DP, Wozniak JM, Skeel RD (2001) Langevin stabilization of molecular dynamics. J Chem Phys 114:2090–2098. https://doi.org/10.1063/1.1332996
22. Essmann U, Perera L, Berkowitz ML et al (1995) A smooth particle mesh Ewald method. J Chem Phys 103:31–34
23. Kollman PA, Massova I, Reyes C et al (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33:889–897. https://doi.org/10.1021/ar000033j
24. Bader RFW (1985) Atoms in molecules. Acc Chem Res 18:9–15. https://doi.org/10.1021/ar00109a003

25. Lu T, Chen F (2012) Multiwfn: a multifunctional wavefunction analyzer. J Comput Chem 33:580–592. https://doi.org/10.1002/jcc.22885
26. Gutierrez LJ, Angelina E, Gyebrovszki A et al (2016) New small-size peptides modulators of the exosite of BACE1 obtained from a structure-based design. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2016.1145143
27. Gutierrez LJ, Barrera Guisasola EE, Peruchena N, Enriz RD (2016) A QM/MM study of the molecular recognition site of bapineuzumab toward the amyloid-β peptide isoforms. Mol Simul. https://doi.org/10.1080/08927022.2015.1032276
28. Vega-Hissi EG, Tosso R, Enriz RD, Gutierrez LJ (2015) Molecular insight into the interaction mechanisms of amino-2H-imidazole derivatives with BACE1 protease: a QM/MM and QTAIM study. Int J Quantum Chem. https://doi.org/10.1002/qua.24854
29. Barrera Guisasola EE, Gutiérrez LJ, Salcedo RE et al (2016) Conformational transition of Aβ42 inhibited by a mimetic peptide. A molecular modeling study using QM/MM calculations and QTAIM analysis. Comput Theor Chem. https://doi.org/10.1016/j.comptc.2016.02.002
30. Luchi AM, Angelina EL, Andujar SA et al (2016) Halogen bonding in biological context: a computational study of D2 dopamine receptor. J Phys Org Chem 29:645–655. https://doi.org/10.1002/poc.3586
31. Parraga J, Cabedo N, Andujar S et al (2013) 2,3,9- and 2,3,11-trisubstituted tetrahydroprotoberberines as D2 dopaminergic ligands. Eur J Med Chem 68:150–166. https://doi.org/10.1016/j.ejmech.2013.07.036
32. Parraga J, Andujar SA, Rojas S et al (2016) Dopaminergic isoquinolines with hexahydrocyclopenta[ij]-isoquinolines as D2-like selective ligands. Eur J Med Chem 122:27–42. https://doi.org/10.1016/j.ejmech.2016.06.009
33. Andujar S, Tosso RD, Suvire FD et al (2012) Searching the "biologically relevant"conformation of dopamine: a computational approach. J Chem Inf Model 52:99–112. https://doi.org/10.1021/ci2004225
34. Angelina EL, Andujar SA, Tosso RD et al (2014) Non-covalent interactions in receptor–ligand complexes. A study based on the electron charge density. J Phys Org Chem 27:128–134. https://doi.org/10.1002/poc.3250
35. Tosso R, Vettorazzi M, Andujar A S, et al (2016) The electronic density obtained from a QTAIM analysis used as molecular descriptor. A study performed in a new series of DHFR inhibitors. J Mol Struct 1134:464–474
36. Tosso RD, Andujar SA, Gutierrez L et al (2013) Molecular modeling study of dihydrofolate reductase inhibitors. Molecular dynamics simulations, quantum mechanical calculations, and experimental corroboration. J Chem Inf Model 53:2018–2032. https://doi.org/10.1021/ci400178h
37. Ortiz JE, Garro A, Pigni NB et al (2018) Cholinesterase-inhibitory effect and in silico analysis of alkaloids from bulbs of Hieronymiella species. Phytomedicine 39:66–74. https://doi.org/10.1016/j.phymed.2017.12.020
38. Siano A, Garibotto FF, Andujar SA et al (2017) Molecular design and synthesis of novel peptides from amphibians skin acting as inhibitors of cholinesterase enzymes. J Pept Sci 23:236–244. https://doi.org/10.1002/psc.2974
39. Gutierrez LJ, Parravicini O, Sanchez E et al (2018) New substituted aminopyrimidine derivatives as BACE1 inhibitors: in silico design, synthesis and biological assays. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2018.1424036
40. Gutierrez LJ, Angelina E, Gyebrovszki A et al (2017) New small-size peptides modulators of the exosite of BACE1 obtained from a structure-based design. J Biomol Struct Dyn 35:413–426. https://doi.org/10.1080/07391102.2016.1145143
41. Durham E, Dorr B, Woetzel N et al (2009) Solvent accessible surface area approximations for rapid and accurate protein structure prediction. J Mol Model 15:1093–1108. https://doi.org/10.1007/s00894-009-0454-9
42. Menendez CA, Accordino SR, Gerbino DC, Appignanesi GA (2016) Hydrogen bond dynamic propensity studies for protein binding and drug design. PLoS ONE 11:e0165767. https://doi.org/10.1371/journal.pone.0165767
43. Menendez CA, Accordino SR, Gerbino DC, Appignanesi GA (2015) "Chameleonic" backbone hydrogen bonds in protein binding and as drug targets. Eur Phys J E 38:107. https://doi.org/10.1140/epje/i2015-15107-3