

Structure-Based Drug Design Studies on a Series of Aldolase Inhibitors

Leonardo G. Ferreira, Ricardo N. dos Santos and Adriano D. Andricopulo*

Laboratório de Química Medicinal e Computacional, Instituto de Física de São Carlos,
Universidade de São Paulo, Av. Trabalhador São-Carlense 400, 13560-970 São Carlos-SP, Brazil

A tripanossomíase africana, também conhecida como doença do sono, é responsável por um grande número de mortes na África. Até o presente, não há tratamento seguro e eficaz disponível. A enzima aldolase do parasita *Trypanosoma brucei* é um alvo atrativo e validado para o desenvolvimento de novos fármacos. Uma série de ésteres fosfóricos foi estudada com uma combinação de métodos de planejamento de fármacos. Modelos de relações quantitativas tridimensionais entre estrutura e atividade (3D QSAR) foram gerados empregando-se a análise comparativa de campos moleculares (CoMFA). Resultados significativos foram obtidos ($r^2 = 0,95$, coeficiente de correlação de validação não-cruzada, e $q^2 = 0,80$, coeficiente de correlação de validação cruzada), indicando a capacidade de predição do melhor modelo para novas moléculas. O modelo foi então utilizado para prever valores das variáveis dependentes (pK_i) de um conjunto teste de compostos, e os valores obtidos apresentaram boa concordância com os resultados experimentais. A integração de estudos de QSAR 3D, docagem molecular e dinâmica molecular forneceram informações úteis em relação às bases estruturais para a inibição seletiva da enzima alvo.

Human African trypanosomiasis, also known as sleeping sickness, is a major cause of death in Africa, and for which there are no safe and effective treatments available. The enzyme aldolase from *Trypanosoma brucei* is an attractive, validated target for drug development. A series of alkyl-glycolamido and alkyl-monoglycolate derivatives was studied employing a combination of drug design approaches. Three-dimensional quantitative structure-activity relationships (3D QSAR) models were generated using the comparative molecular field analysis (CoMFA). Significant results were obtained for the best QSAR model ($r^2 = 0.95$, non-cross-validated correlation coefficient, and $q^2 = 0.80$, cross-validated correlation coefficient), indicating its predictive ability for untested compounds. The model was then used to predict values of the dependent variables (pK_i) of an external test set, and the predicted values were in good agreement with the experimental results. The integration of 3D QSAR, molecular docking and molecular dynamics simulations provided further insight into the structural basis for selective inhibition of the target enzyme.

Keywords: African trypanosomiasis, drug, design, inhibitors, QSAR, molecular modeling

Introduction

The African continent has been devastated by the enormous impact of the neglected tropical diseases (NTDs), as they disproportionately affect the poorest and most vulnerable populations. For decades, NTDs represent an insurmountable obstacle to the socioeconomic development of communities in the infested areas. Among the most important diseases is human African trypanosomiasis (HAT, or sleeping sickness), which is caused by two sub-species of protozoa, *Trypanosoma brucei gambiense* and

Trypanosoma brucei rhodesiense. HAT threatens approximately 60 million people in the sub-Saharan Africa, causing long-term disability and death. The disease, transmitted by the bite of tsetse flies, is fatal in untreated cases.^{1,2}

The limited existing chemotherapy, consisting of old drugs such as suramin (discovered in 1920), pentamidine (1921), melarsoprol (1949) and eflornithine (1991), suffers from a variety of drawbacks including poor efficacy, resistance and serious side effects.³ Moreover, due to these limitations and growing drug resistance, combination therapy has been employed as an important clinical alternative.^{4,5} According to the World Health Organization

*e-mail: aandrico@ifsc.usp.br

(WHO), there is an urgent need for new drugs that are safe, affordable and effective for use in humans.¹

The identification of novel molecular targets is of great importance in drug discovery.⁶ Fructose-1,6-bisphosphate aldolase from *T. brucei* (EC 4.1.2.13) has been considered an attractive and validated drug target for HAT therapy.⁷ The enzyme, a homotetramer of 164 KDa, belongs to the type I class of aldolases, which in contrast to type II, does not require a metal ion to catalyze the cleavage of the substrate.^{8,9} *Trypanosoma brucei* aldolase plays a pivotal role in the glycolytic flux, the sole source of ATP (adenosine triphosphate) in the parasite, as it cleaves the hexose fructose-1,6-bisphosphate into two trioses, glyceraldehyde-3-phosphate and dihydroxyacetone-phosphate.^{10,11} Therefore, aldolase has been exploited in medicinal chemistry efforts toward new lead compounds against HAT.¹²⁻¹⁴

Cheminformatic tools have been employed to successfully analyze biological and chemical data, creating useful information for drug design. Among them, quantitative structure-activity relationships (QSAR) have been widely used for the design of a variety of biologically active compounds.¹⁵⁻¹⁷ Comparative molecular field analysis (CoMFA), one of the most popular 3D QSAR methods, is based on the assumption that the activity of a series of compounds is directly related to ligand-receptor interactions, requiring the 3D structural alignment of the data set molecules in order to calculate the interaction molecular fields.¹⁸ In the present work, the CoMFA approach was employed to investigate a series of aldolase inhibitors. Molecular docking and molecular dynamics (MD) studies were also used to explore the molecular basis underlying the inhibition of *T. brucei* aldolase, considering the stability and evolution of the main intermolecular interactions in such a dynamic system.¹⁹⁻²²

Methodology

Computational approach

The QSAR modeling analyses, calculations and visualizations for CoMFA were performed using the SYBYL 8.0 software suite (Tripos Inc., St. Louis, USA), running on CentOS Linux workstations. The 3D structures of the molecules were built employing the Powell method and the Tripos force field, implemented in SYBYL 8.0.^{23,24} Each conformation was further energetically minimized and submitted to charge calculation applying the semi-empirical quantum chemistry program MOPAC 6.0. The Austin method 1 (AM1) was applied and charges were calculated by electrostatic potential (ESP), with a 1.4 scaling factor for van der Waals distances and a 0.4 increment between layers

in ESP.^{25,26} Molecular docking was performed with GOLD 5.0 software (Cambridge Crystallographic Data Centre, Cambridge, UK). The molecular dynamics simulations were carried out with NAMD 2.7 at a Dell PowerEdge 1950 II cluster and visualized with VMD 1.9 program (University of Illinois, Urbana, USA).

Data set

The data set employed for the QSAR studies was selected from the literature and consists of a series of 38 alkyl-glycolamido and alkyl monoglycolate phosphoric esters along with K_i values (the dissociation constant for the enzyme-inhibitor complex), determined for the aldolase enzyme from rabbit muscle.²⁷⁻²⁹ Table 1 includes the structures and corresponding pK_i (dependent variables) values for the complete set of inhibitors. The complete data set was divided into training (compounds **2**, **4**, **5**, **7**, **8**, **10-13**, **15**, **16**, **18-21**, **23-30**, **32-38**) and test sets (compounds **1**, **3**, **6**, **9**, **14**, **17**, **22** and **31**) in the ratio of approximately 80 and 20%, respectively. In order to select appropriate training and test sets, a statistical cluster analysis was carried out using the software Tsar version 3.3 (Accelrys, San Diego, USA), employing the complete linkage clustering method.³⁰ The structurally diverse molecules having a significant coverage of property values were included in both sets. Thus, the data set is suitable for QSAR model development. The predictive ability of the final models was assessed by the full cross-validated (q^2), partial least squares (PLS), leave-one-out (LOO) and leave-many-out (LMO) methods.³¹ Progressive scrambling was applied to determine the sensitivity of the QSAR models to chance correlations. External model validation was performed with the test set (Table 1), whose compounds were not considered for QSAR model generation. After generation of the PLS training set models, values of the dependent variables (pK_i) were predicted for the test set compounds, allowing the determination of predictive- r^2 (non-cross-validated correlation coefficient) values.

Structural alignment and molecular docking

The structural alignment of the data set molecules was carried out using the X-ray coordinates of the *T. brucei* aldolase enzyme (PDB ID 1F2J, resolution of 1.9 Å).⁸ The preparation of the structure and the docking procedure were accomplished with the GOLD 5.0 program.³² After removal of the water molecules, hydrogen atoms were added in standard geometry. The binding site was defined as a sphere with a radius of 10 Å around the side chain nitrogen atom of Lys239. The cavity detection function was applied to restrict

Table 1. Chemical structures and corresponding pK_i values of the aldolase inhibitors

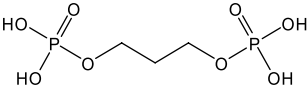
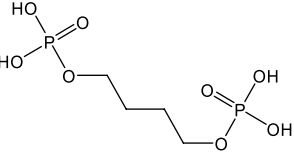
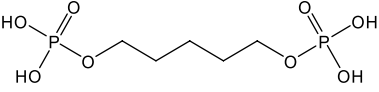
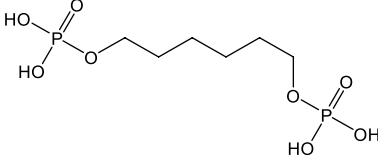
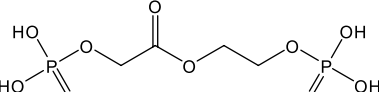
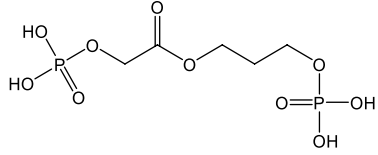
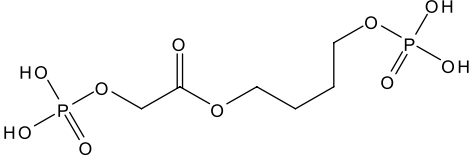
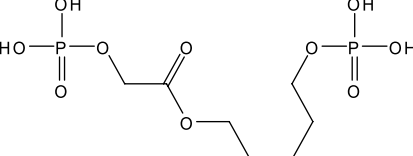
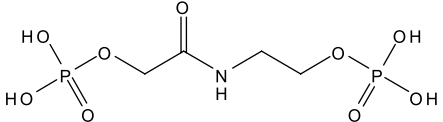
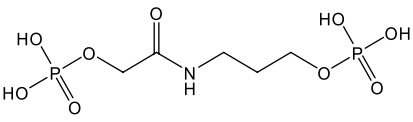
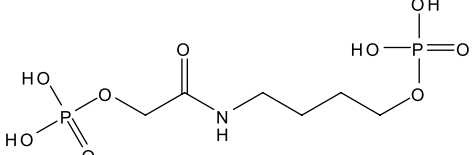
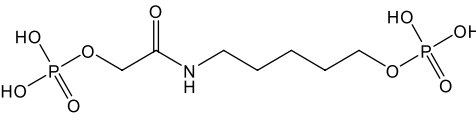
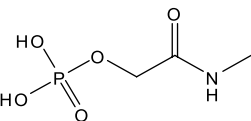
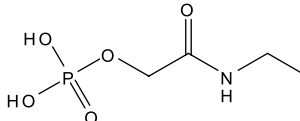
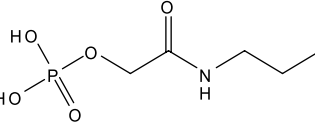
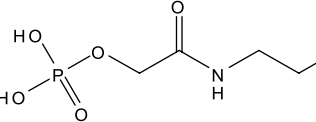
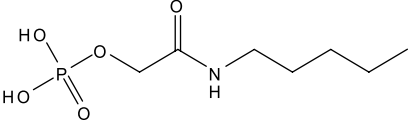
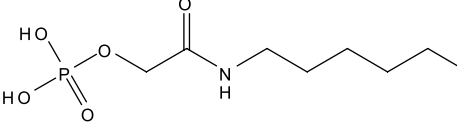
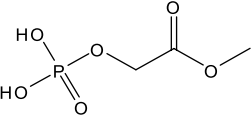
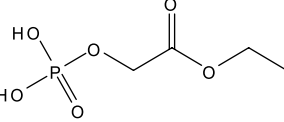
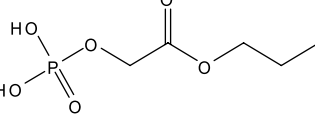
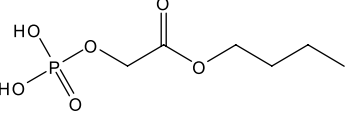
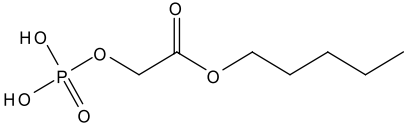
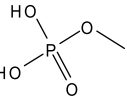
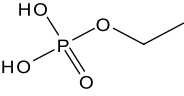
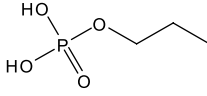
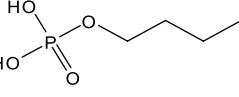
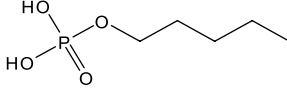
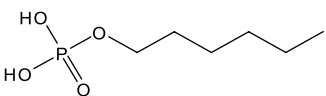
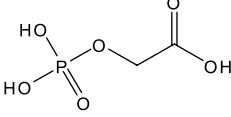
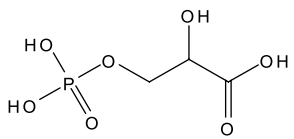
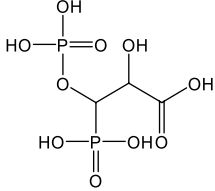
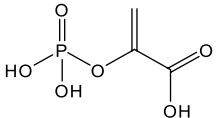
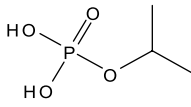
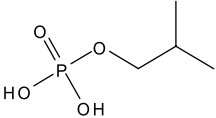
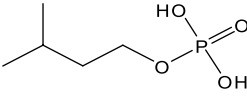
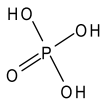
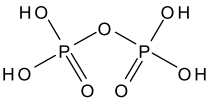
Cpd	Structure	pK_i / M	Cpd	Structure	pK_i / M
1		2.44	2		2.58
3		2.85	4		3.27
5		3.14	6		3.48
7		3.28	8		3.00
9		3.48	10		4.26
11		3.60	12		3.24
13		1.67	14		2.10
15		2.14	16		1.53
17		1.31	18		1.92
19		2.26	20		2.06
21		1.64	22		1.72

Table 1. continuation

Cpd	Structure	pK _i / M	Cpd	Structure	pK _i / M
23		1.88	24		2.22
25		1.82	26		1.88
27		1.79	28		1.88
29		2.30	30		2.30
31		1.96	32		2.52
33		2.18	34		1.52
35		1.30	36		1.77
37		2.30	38		2.11

the atom selection to solvent accessible surface. It was attributed to the residues of Arg52, Arg158, Arg313, Lys116, Lys156 and Lys239, a rotational freedom of 10 degrees. GoldScore function was employed to evaluate and score the binding solutions. Each molecule of the data set was docked 20 times and the top scoring conformations were further considered for the CoMFA studies. The complete aligned data set is depicted in Figure 1.

3D QSAR CoMFA

To engender useful QSAR models and access the contributions of electrostatic and steric fields to the affinity of the data set inhibitors, CoMFA studies were performed upon the 3D molecular alignment illustrated in Figure 1.

The aligned training set molecules were posed in a 3D grid box, with a grid spacing of 2.0 Å in *x*, *y* and *z* directions,

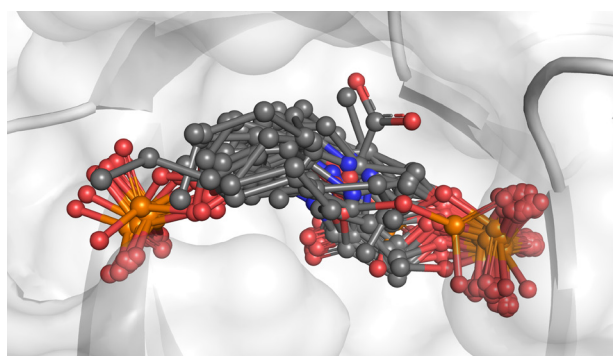


Figure 1. Structural alignment of the inhibitors into the active site of *T. brucei* aldolase.

with an extension of 4.0 Å in each direction beyond the molecules. The molecular descriptors (CoMFA steric and electrostatic fields) were generated at each grid point with Tripos force field and a sp³ carbon atom possessing

a +1 net charge, as a probe. CoMFA region focusing was applied to refine the models by increasing the weight for those lattice points which are most pertinent to the model. This proceeding enhances the resolution and predictive power of the subsequent PLS analyses. The default value of 30 kcal mol⁻¹ was stipulated as the maximum steric and electrostatic energy cutoff. The statistical analyses of the models were performed with the q^2_{LOO} and q^2_{LMO} PLS methods and the CoMFA standard for variable scaling.

Molecular dynamics

In order to explore key molecular interactions responsible for the binding affinity of the inhibitors, molecular dynamics simulations were performed with NAMD and CHARMM force field for the aldolase enzyme in complex with compounds **4** and **10**.^{33,34} The topology and parameters of the ligands were obtained using the SwissParam server (Swiss Institute of Bioinformatics, Switzerland) and evaluated with vacuum equilibrations.³⁵ The charge distribution of each ligand was obtained by calculating the Mulliken charges employing Mulliken Population Analyses at the Hartree-Fock level with a 6-31G* basis set, using the PC-GAMESS/Firefly 7.1 software.³⁶⁻³⁸ Subsequently, the complexes obtained by docking the compounds **4** and **10** into the aldolase binding site were solvated in a water box with approximate dimensions of 80 × 80 × 90 Å, with a layer of 10 Å from the protein extremities. Afterwards, both systems were neutralized adding a minimal number of sodium and chloride counter-ions. This procedure provided each system with a total of 47,200 atoms. Protein and ligand ionization states were set based on a pH of 7.0. Long-range electrostatic interactions were calculated through the Particle Mesh Ewald method using a cut-off of 12 Å.³⁹⁻⁴¹

For the two complexes, protein and ligand were then subjected to a restrained run of 100 ps, whereby only water molecules could move in order to avoid possible steric conflicts. The entire systems were then relieved, minimized during 20 ps and equilibrated in simulations of 10 ns using a NTP ensemble, at a temperature of 298 K and pressure of 1.0 atm. Pressure and temperature were retained by the Nosé-Hoover piston pressure control and Langevin Dynamics, respectively.⁴²⁻⁴⁴ The entire simulation process was repeated 5 times for each protein-ligand complex.

Results and Discussion

Biochemical data

The data set consists of a series of alkyl-glycolamido and alkyl-monoglycolate phosphoric esters, which are

structurally related to fructose-bisphosphate (Table 1). The biochemical data (K_i values) were collected under the same experimental conditions, a fundamental requisite for QSAR studies.¹⁸ For modeling purposes, the dependent variables were converted into the corresponding pK_i values ($-\log K_i$), which span approximately three orders of magnitude.

This data set was selected because of its suitability for QSAR modeling, considering the appropriate structural diversity and the quality of the experimental data. In addition, it should be noted that no 3D QSAR studies are available in the literature for this enzyme, emphasizing the importance of the present work.

The selection of training and test sets was executed in such a way that structurally diverse molecules having a wide range of data were included in both sets. From the 38 inhibitors, 30 were selected as members of the training set, for model development and the other 8 (compounds **1**, **3**, **6**, **9**, **14**, **17**, **22** and **31**) as members of the test set for external validation. The predictive ability of the final model was assessed by PLS regression with LOO and LMO cross-validation and also for pK_i prediction for the test set.

Enzyme homology and docking strategy

The aldolase enzyme from *T. brucei* has a 49% of sequence identity compared to the rabbit muscle and human enzymes (Figure 2), and the three proteins have the same tridimensional structural folding pattern.⁸ An analysis of the amino acid residues illustrates the high conservation between the binding site of the rabbit muscle and parasite enzymes (Figure 3). Therefore, it could be expected that inhibitors of the rabbit muscle enzyme would bind in a similar fashion to the *T. brucei* aldolase. In fact, this is the case for mannitol-bisphosphate (MBP), a known inhibitor of the rabbit muscle and *T. brucei* aldolase enzymes, which is structurally related to the data set compounds employed in this work.⁴⁵ The binding mode of MBP is illustrated in Figure 4, in which the crystallographic conformation of the inhibitor into the active site of rabbit muscle aldolase (PDB ID 1ZAJ) is compared with the docking solution upon the parasitic enzyme.⁴⁶ As can be seen, the inhibitor binding poses are very similar for both enzymes. Therefore, the structure of *T. brucei* aldolase was employed to obtain the structural alignment for the CoMFA studies and then, to investigate, through docking and molecular dynamics, the key intermolecular interactions involved in the binding affinity of the inhibitors for the parasitic enzyme.

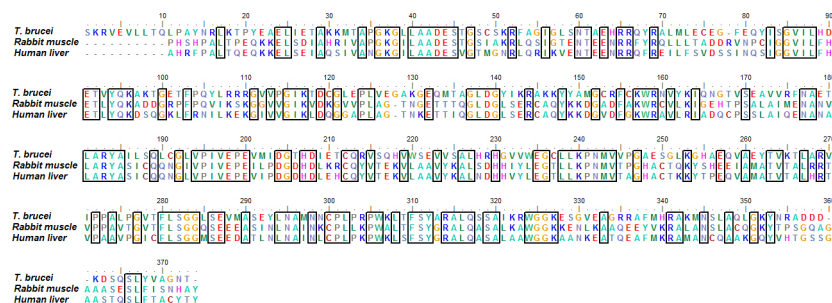


Figure 2. Amino acid sequence alignment of *T. brucei*, rabbit muscle and human liver aldolases. Numbers correspond to amino acid positions of the *T. brucei* enzyme. Conserved amino acid residues are highlighted.

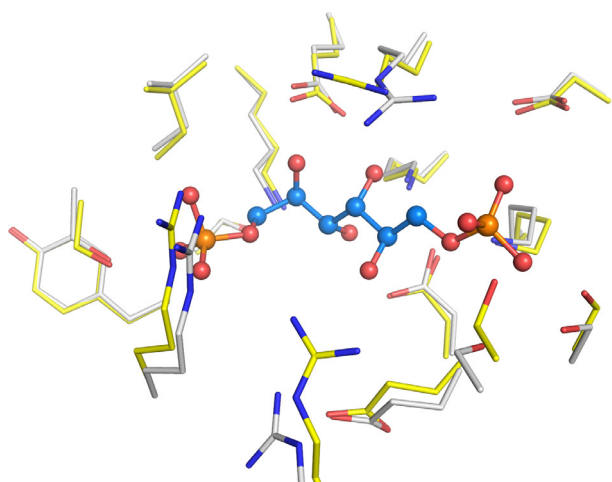


Figure 3. Overlapping of the active site amino acid residues of the aldolase enzymes from rabbit muscle (in yellow) and *T. brucei* (in grey). The inhibitor mannitol-bisphosphate is shown in ball-and-stick representation (in blue).

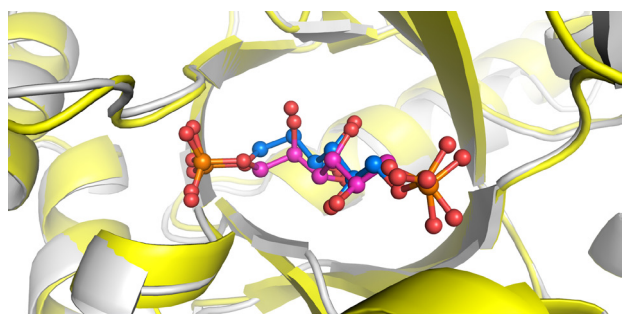


Figure 4. Crystallographic conformation of MBP (in blue) adopted within the rabbit muscle aldolase binding site (in yellow). The MBP inhibitor (in magenta) is docked into the *T. brucei* binding site (in gray).

Table 2. CoMFA results according to the variation of SDC values

Model	q^2	SEP	r^2	SEE	N	F	Fraction	
							S	E
1	0.73	0.37	0.85	0.27	2	81.02	0.45	0.55
2	0.76	0.36	0.95	0.16	4	124.80	0.51	0.49
3	0.80	0.33	0.95	0.16	4	124.82	0.48	0.52

q^2 : leave-one-out (LOO) cross-validated correlation coefficient; N: optimum number of components; r^2 : non-cross-validated correlation coefficient; SEP: standard error of prediction; SEE: standard error of estimate; F : F -test value; S: steric field; E: electrostatic field. Grid spacing of 2 Å in all cases. Model 1, SDC = 0.3; model 2, SDC = 0.6; model 3, SDC = 0.9.

CoMFA statistical results

The aligned structures of the inhibitors (Figure 1) were submitted to several CoMFA analyses, which initially resulted in a model with q^2 of 0.72 and r^2 of 0.77. These results were considered liable to improvement, particularly regarding the fit of the data, represented by the r^2 value. Hence, the region focusing weighted by standard deviation coefficient (SDC) was applied, with values ranging from 0.3 to 0.9. This strategy not only increased both correlation coefficients, leading to improved models, but also resulted in the refinement of the 3D contour maps. The statistical results are presented in Table 2. As can be seen, significant correlation coefficients were obtained for CoMFA models 2 and 3. The best model (3) exhibited q^2 of 0.80 and r^2 of 0.95.

The stability of the models was further confirmed applying the LMO cross validation method, resulting in a q^2 value of 0.73 for 5 cross validation groups. For 10 groups, a value of 0.76 for q^2 was obtained. These values are comparable to those obtained by using LOO cross validation.

Although the cross-validated correlation coefficient may adequately represent the internal consistency and predictive power of the models for untested compounds, the external validation process is the most significant validation method. Therefore, the predictive ability of model 3 was assessed by predicting the affinity of an external test set of 8 compounds. The values of experimental and predicted pK_i

values for the test set compounds are shown in Table 3, and the graphic results are displayed in Figure 5. The good correlation between experimental and predicted pK_i values for the test set compounds indicates the ability of the QSAR model to predict the affinity of novel inhibitors within this structural diversity ($r^2_{\text{pred}} = 0.84$).

Table 3. External validation: experimental and predicted pK_i values for the test set compounds

Inhibitor	pK_i , experimental	pK_i , predicted	Residual ^a
1	2.44	3.00	-0.56
3	2.85	3.13	-0.28
6	3.48	3.02	0.46
9	3.48	3.30	0.18
14	2.10	1.87	0.23
17	1.31	1.60	-0.29
22	1.72	1.77	-0.05
31	1.96	1.86	0.10

^aDifference between experimental and predicted pK_i values.

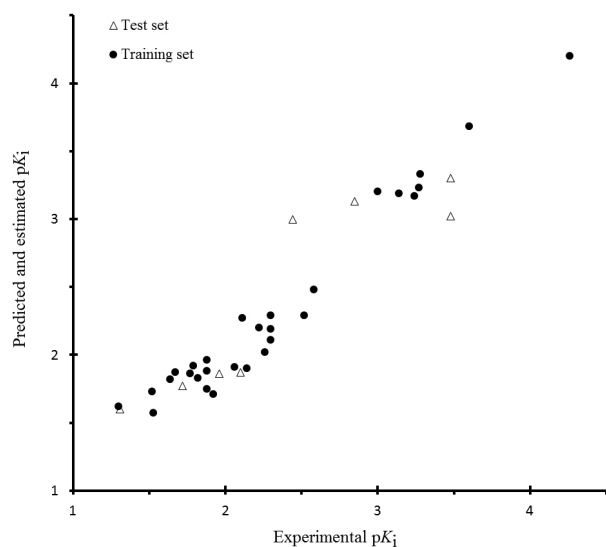


Figure 5. Plot of experimental values of pK_i vs. the corresponding predicted and estimated values for the 30 training (solid circles) and 8 test (open triangles) set compounds.

CoMFA contour maps

The relationships between structure and activity of the data set compounds were further investigated through the interpretation of the CoMFA steric and electrostatic contour maps for model 3 (Figure 6). The contour maps show regions in the 3D space where changes of steric and electrostatic properties correlate with respective changes in the affinity of the inhibitors for the enzyme.

The CoMFA steric contour, represented by green polyhedra, indicates that bulky groups are advantageous

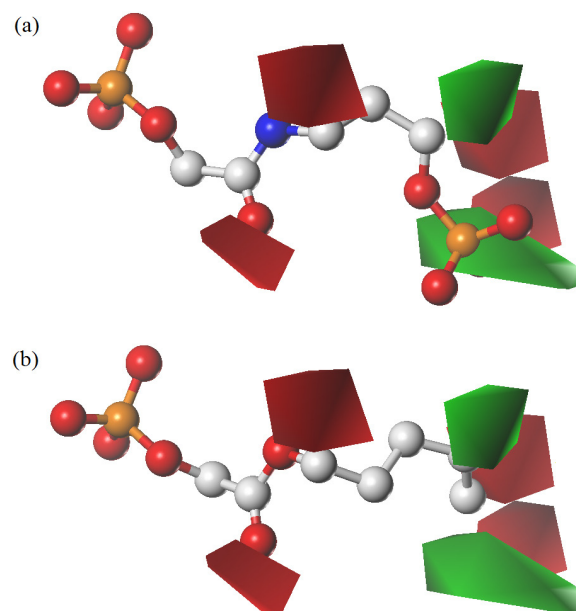


Figure 6. CoMFA contour maps (model 3) for inhibitors **10** (a) and **23** (b).

for activity, while yellow regions advert that bulky groups are detrimental for activity. The map revealed that the compounds possessing only one phosphate group are aligned in such a way that the extremity lacking the other group are oriented always to the same side of the active site. From Figure 6, it can be assumed that the presence of bulky groups in this extremity would be beneficial to the inhibitor affinity. An example is compound **23**, which does not hold a second bulky group (such as phosphate) and has a binding affinity 1000-fold lower than compound **10**.

Regarding the CoMFA electrostatic contour maps, blue regions indicate that electropositive groups are favorable, while red polyhedra give support for electronegative groups. Figure 6 shows red regions surrounding one of the phosphate groups of compound **10** (Figure 6a) and the phosphate lacking extremity of compound **23** (Figure 6b). This finding is consistent with the experimental data since compounds presenting electronegative groups in this region have higher affinity for the enzyme. Moreover, the red regions surrounding the amide and the ester groups of both compounds demonstrate the importance of acceptor groups within the hydrocarbon chain, as indicated by the experimental data and docking studies.

Enzyme-inhibitor interactions

The understanding of protein-ligand interactions is essential for the design and optimization of enzyme inhibitors. In the present study, it was explored an approach that combines 3D QSAR, molecular docking and molecular dynamics to investigate the intermolecular interactions

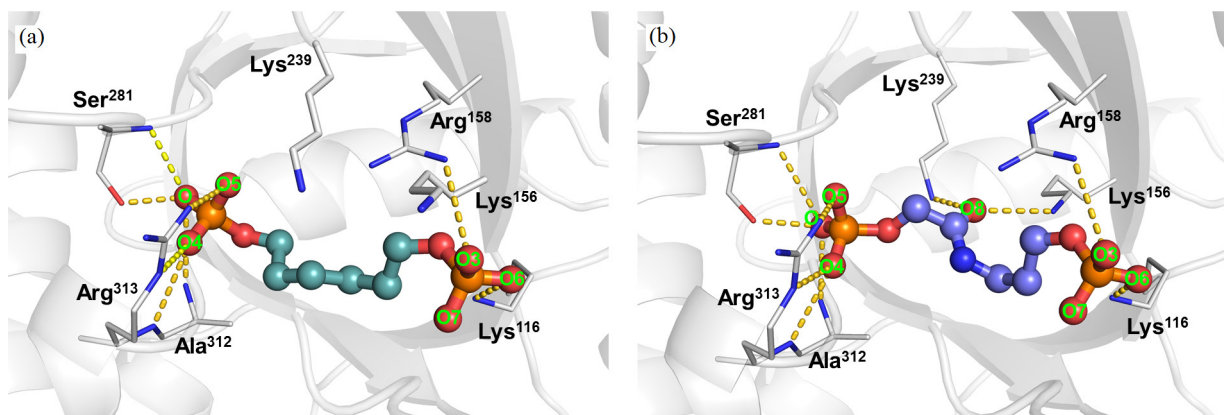


Figure 7. (a) Binding conformation of compound **4** showing electrostatic interactions between the ligand phosphates and the binding site residues of aldolase. (b) Binding conformation of compound **10** showing a very similar interaction pattern to that of compound **4**, with the presence of additional interactions between the amide group of the ligand and the residues Lys156 and Lys239 of the enzyme.

between *T. brucei* aldolase and the data set inhibitors. An analysis of the biochemical data for these compounds shows that the best inhibitors possess two phosphoric esters in their chemical structures (Table 1). Moreover, the presence of an amide substituent within the hydrocarbon chain increases the affinity of the compounds. In order to rationalize these experimental results, identifying key motifs associated to the inhibitory activity and asserting the importance of polar groups within the hydrocarbon chain, it was selected compounds **10** (the highest affinity inhibitor of the data set, $pK_i = 4.26$) and **4** ($pK_i = 3.27$) for molecular dynamics studies. The chemical structure of these compounds differs only for the presence of an amide group within the hydrocarbon chain of compound **10**. The docking conformations of these ligands are depicted in Figure 7.

The binding conformations of the two compounds revealed that the phosphoric esters are necessary for electrostatic interactions with Lys116, Arg158 and Arg313. This moiety also establishes ion-dipole contacts with residues Ser281, Ala312 and Arg313. Furthermore, compound **10** which possess an amide group, establishes additional ion-dipole interactions between its carbonyl oxygen and residues Lys156 and Lys239, as shown in Figure 7b.

The stability of these interactions was further evaluated by molecular dynamics. The average root-mean-square deviations (RMSD) obtained during the simulations from the starting initial conformations of the protein-ligand complexes and for the apoenzyme are depicted in Figure 8a. The plots show very similar RMSD patterns, with equilibrium being achieved mutually after 6 ns. The analysis of RMSD shown in Figure 8b indicates that both ligands are maintained into the binding site of aldolase with the same orientation during the entire simulations. A slightly lower RMSD profile was observed for compound

10 in comparison with compound **4**. The higher stability of compound **10** could be related to its observed higher affinity to the aldolase enzyme (10-fold higher than compound **4**). It was not observed the release of the ligand in any of the 5 simulations for each system. Furthermore, the plot of the root-mean-square fluctuation (RMSF) of all protein residues for both systems (Figure 8c) shows that the binding site residues exhibit the same mobility in the presence of both ligands. As expected, the flexible N-terminal aldolase tail (residues 1-20, Figure 8c) presents high mobility during all simulations.

To verify the conservation of the interaction pattern observed by molecular docking between both ligands and the aldolase binding site, and to evaluate the key elements responsible for the higher affinity of compound **10**, the length of the interactions identified by docking between these two inhibitors and the binding site were evaluated during the MD simulations (Figure 9).

As shown in Figure 9, the intermolecular interactions observed in the initial conformation were well maintained during the simulations. The general profile of all interactions is very similar and linear for both ligand complexes. Furthermore, both aldolase inhibitors seem to be stabilized mainly by common electrostatic interactions. The residue Ser281 establishes two ion-dipole interactions with the phosphate group from both ligands (Figure 7; Figures 9c and d). Furthermore, the observance of three electrostatic interactions involving Arg313 indicates the importance of this residue for the biological activity of this class of inhibitors (Figures 9f, 9g and 9h). Arginine 313 maintains two ion-dipole interactions with the oxygen O4 of the ligands. The first one involves the backbone amino group, and the second one the side chain amino group (Figures 9f and 9h). Also, the protonated nitrogen of Arg313 participates in a strong ionic interaction with the oxygen O5 of the ligands (Figure 9g). Residues Lys116, Arg158 and

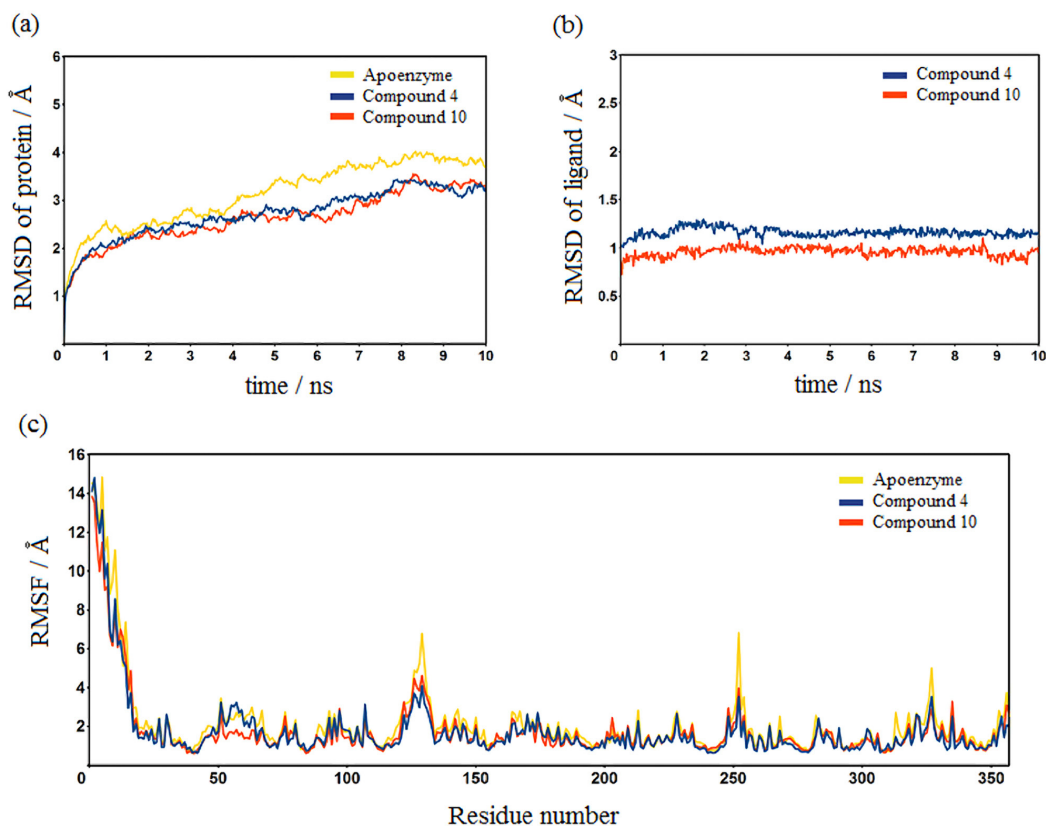


Figure 8. (a) RMSD of aldolase as apoenzyme (yellow line) and in the presence of compounds **4** (blue line) and **10** (orange line) during the MD simulations. (b) RMSD of compounds **4** and **10** during the MD simulations. (c) Average fluctuations of each aldolase residue during the simulations for the apoenzyme and the complexes with compounds **4** and **10**. Each curve is an average of 5 independent simulations.

Ala312 also participate in electrostatic interactions in both ligand complexes, increasing the binding stability.

Essentially, the higher affinity of compound **10** is basically correlated with its additional steady ion-dipole interactions between the oxygen atom O8 of the amide group and the residues Lys156 and Lys239 (Figures 9i and 9j). This finding reveals the importance of the presence of such a polar group in the backbone of this inhibitor.

Structural basis for the design of selective inhibitors

Considering several molecular aspects involved in the binding of the data set inhibitors to the *T. brucei* aldolase enzyme, as well as the information provided by the CoMFA contour maps, it was investigated some key structural aspects which could lead to selectivity towards the parasite enzyme. An analysis of the superimposed structures of the *T. brucei* and human liver (PDB ID 1QO5) aldolases (Figure 10) revealed that the Gly302 and Thr38 residues in the human enzyme are replaced by Ala312 and Ser48 in the parasite enzyme, respectively.⁴⁷ Moreover, it is observed that a neutral phenylalanine residue (Phe79), present in the human aldolase is replaced by a positively charged histidine residue (His88) in the

T. brucei enzyme. The data set inhibitors do not occupy the binding pocket where the Phe79 and His88 residues are located, thus the elongation of the compounds beyond the phosphate moiety through the attachment of a negatively charged or a polar group to provide a better occupation of this pocket would be beneficial to the selectivity towards the *T. brucei* aldolase.

Conclusions

The molecular modeling strategy employed in this work was useful to reveal key molecular characteristics related to the binding affinity and selectivity of this series of alkyl-glycolamido and alkyl-monoglycolate aldolase inhibitors. The statistically significant QSAR model was able to predict the pK_i values for the test set compounds. Moreover, the steric and electrostatic contour maps provided evidences for the importance of electronegative and bulky functional groups in the hydrocarbon scaffold of the data set molecules. The docking studies and MD simulations revealed important molecular insights for the best inhibitor of the series into the binding site of *T. brucei* aldolase. Finally, the examination of the main differences between the human and the parasitic enzyme was profitable to gather

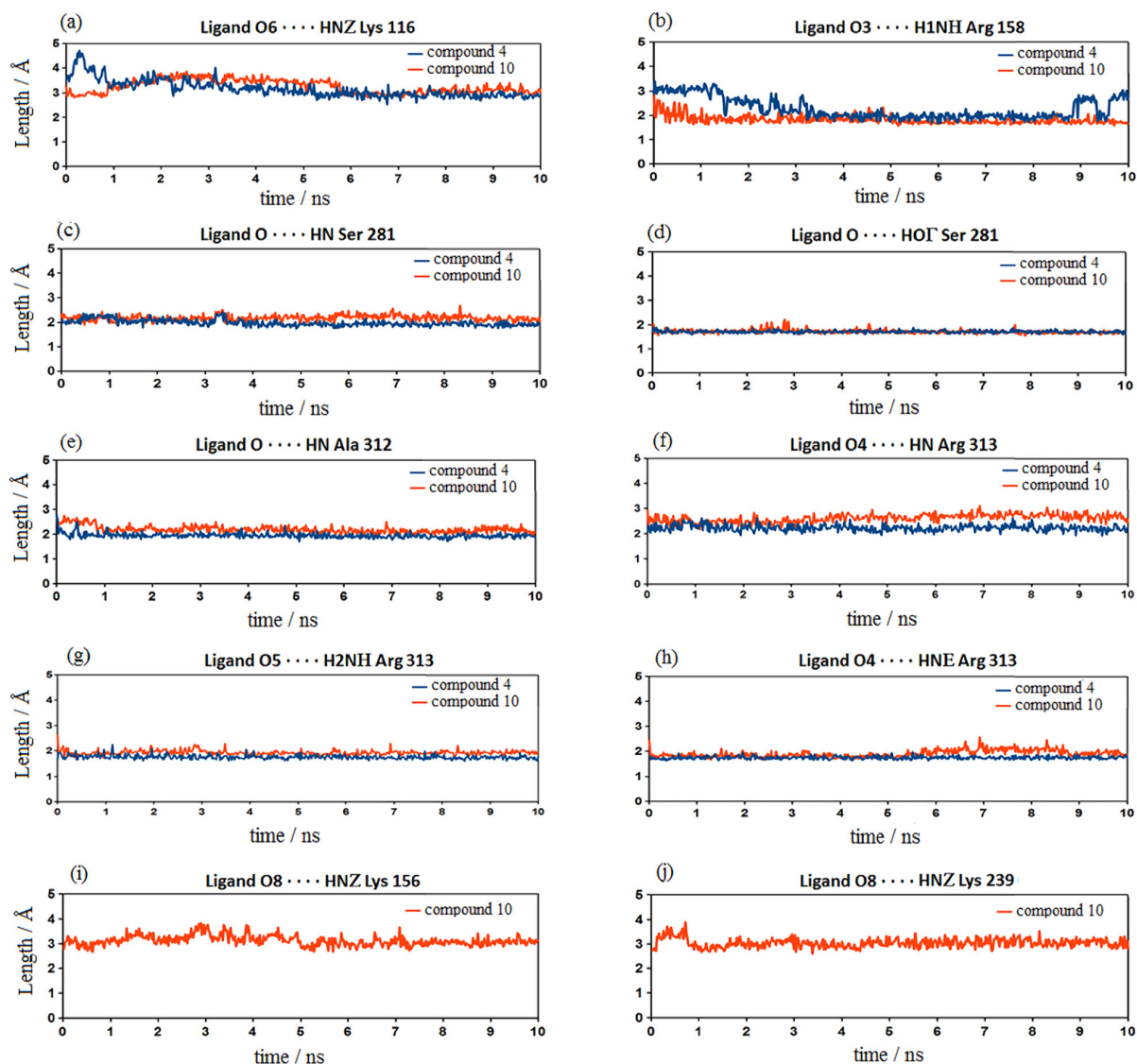


Figure 9. Distance of the intermolecular interactions between the docked ligands and the binding site residues during the MD simulations. Each plot is an average of 5 independent simulations. The interaction lengths between the lysine residues and ligands were calculated considering an average of the distances from all three lysine hydrogen atoms.

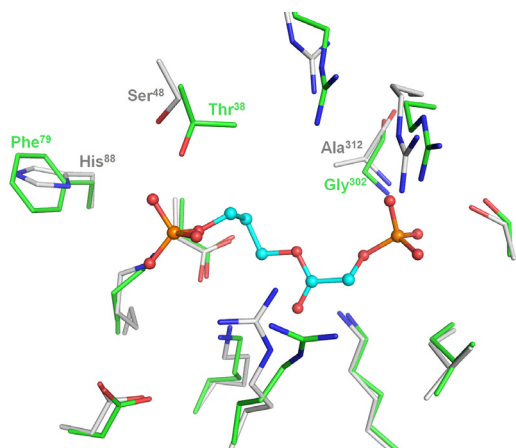


Figure 10. Active site of aldolase from *T. brucei* (in gray) superimposed to the human liver enzyme (in green). In the center, the docked conformation of inhibitor **6** is shown in ball-and-stick.

knowledge on the structural basis underlying the design of new inhibitors with increased affinity and selectivity.

Acknowledgments

We gratefully acknowledge financial support from the State of São Paulo Research Foundation (FAPESP, Fundação de Amparo à Pesquisa do Estado de São Paulo) and the National Council for Scientific and Technological Development (CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico), Brazil.

References

1. Malvy, D.; Chappuis, F.; *Clin. Microbiol. Infect.* **2011**, *17*, 986.

2. Kioy, D.; Jannin, J.; Mattock, N.; *Nat. Rev. Microbiol.* **2004**, *2*, 186.
3. Ridley, R. G.; *EMBO Rep.* **2003**, *4*, 43.
4. Jannin, J.; Cattand, P.; *Curr. Opin. Infect. Dis.* **2004**, *17*, 565.
5. Barrett, M. P.; Vincent, I. M.; Burchmore, R. J.; Kazibwe, A. J.; Matovu, E.; *Future Microbiol.* **2011**, *6*, 1037.
6. Fauman, E. B.; Rai, B. K.; Huang, E. S.; *Curr. Opin. Chem. Biol.* **2011**, *15*, 463.
7. Cáceres, A. J.; Michels, P. A. M.; Hannaert, V.; *Mol. Biochem. Parasitol.* **2010**, *169*, 50.
8. Chudzick, D. M.; Michels, P. A. M.; de Walque, S.; Hol, W. G. J.; *J. Mol. Biol.* **2000**, *300*, 697.
9. Gefflaut, T.; Blonski, C.; Perie, J.; Willson, M.; *Prog. Biophys. Mol. Biol.* **1995**, *63*, 301.
10. Verlinde, C. L.; Hannaert, V.; Blonski, C.; Willson, M.; Périé, J. J.; Fothergill-Gilmore, L. A.; Opperdoes, F. R.; Gelb, M. H.; Hol, W. G.; Michels, P. A.; *Drug Resist. Updat.* **2001**, *4*, 50.
11. Albert, M. A.; Haanstra, J. R.; Hannaert, V.; Van Roy, J.; Opperdoes, F. R.; Bakker, B. M.; Michels, P. A.; *J. Biol. Chem.* **2005**, *31*, 28306.
12. Opperdoes, F. R.; *Ann. Rev. Microbiol.* **1987**, *41*, 127.
13. Bakker, B. M.; Michels, P. A.; Opperdoes, F. R.; Westerhoff, H. V.; *J. Biol. Chem.* **1997**, *6*, 3207.
14. Hoet, S.; Opperdoes, F.; Brun, R.; Quetin-Leclercq, J.; *Nat. Prod. Rep.* **2004**, *21*, 353.
15. Doddareddy, M. R.; Lee, Y. J.; Cho, Y. S.; Choi, K. I.; Koh, H. Y.; Pae, A. N.; *Bioorg. Med. Chem.* **2004**, *12*, 3815.
16. Garcia, I.; Fall, Y.; Gomez, G.; *Curr. Top. Med. Chem.* **2012**, *12*, 895.
17. Andricopulo, A. D.; Yunes, R. A.; Nunes, R. J.; Savi, A. O. S.; Correa, R.; Cruz, A. B.; Cechinel, V.; *Quim. Nova* **1998**, *21*, 573.
18. Ferreira, L. G.; Leitão, A.; Montanari, C. A.; Andricopulo, A. D.; *Med. Chem.* **2011**, *7*, 71.
19. Mascarello, A.; Chiaradia, L. D.; Vernal, J.; Villarino, A.; Guido, R. V. C.; Perizzolo, P.; Poirier, V.; Wong, D.; Martins, P. G. A.; Nunes, R. J.; Yunes, R. A.; Andricopulo, A. D.; Av-Gay, Y.; Terenzi, H.; *Bioorg. Med. Chem.* **2010**, *18*, 3783.
20. Folkers, G.; *Eur. J. Pharm. Sci.* **1994**, *2*, 31.
21. Gasteiger, J.; Marsili, M.; *Tetrahedron* **1980**, *36*, 3219.
22. Postigo, M. P.; Guido, R. V. C.; Oliva, G.; Castilho, M. S.; Pitta, I. D.; Albuquerque, J. F. C.; Andricopulo, A. D.; *J. Chem. Inf. Model.* **2010**, *50*, 1693.
23. Powell, M. J. D.; *Math. Program.* **1977**, *12*, 241.
24. Clark, M.; Cramer III, R. D.; van Opdenbosch, N. J.; *J. Comput. Chem.* **1989**, *10*, 982.
25. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P.; *J. Am. Chem. Soc.* **1985**, *107*, 3902.
26. Luque, F. J.; Orozco, M.; *Chem. Phys. Lett.* **1990**, *168*, 269.
27. Ogata, H.; Takeo, K.; Kuwahara, A.; Suzuno, R.; Fujimoto, M.; Shimizu, J.; *Biochim. Biophys. Acta* **1983**, *742*, 384.
28. Ogata, H.; Fukuda, T.; Yamamoto, K.; Yamasaki, H.; Fujimoto, H.; Fujisaki, S.; Kajigaeshi, S.; *Biochim. Biophys. Acta* **1990**, *1041*, 254.
29. Ogata, H.; Fukuda, T.; Yamamoto, K.; Funakoshi, J.; Takada, K.; Yasue, N.; Fujisaki, S.; Kajigaeshi, S.; *Biochim. Biophys. Acta* **1992**, *1119*, 123.
30. Haiech, J.; Koscielniak, T.; Grassy, G. J.; *J. Mol. Graphics* **1995**, *13*, 46.
31. Bush, B. L.; Nachbar, Jr.; R. B.; *J. Comput. Aided Mol. Des.* **1993**, *7*, 587.
32. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R.; *J. Mol. Biol.* **1997**, *4*, 727.
33. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K.; *J. Comput. Chem.* **2005**, *26*, 1781.
34. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M.; *J. Comp. Chem.* **1983**, *4*, 187.
35. Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O.; *J. Comput. Chem.* **2011**, *32*, 2359.
36. Kosztin, D.; Izrailev, S.; Schulten, K.; *Biophys. J.* **1999**, *76*, 188.
37. Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A.; *J. Comput. Chem.* **1993**, *14*, 1347.
38. Tsareva, D. A.; Osolodkin, D. I.; Shulga, D. A.; Oliferenko, A. A.; Pisarev, S. A.; Palyulin V. A.; Zefirov, N. S.; *Mol. Inf.* **2011**, *30*, 169.
39. Long, D.; Mu, Y.; Yang, D.; *PLoS One* **2009**, *30*, 6081.
40. Sun, T. G.; Hu, J. P.; Li, C. H.; Chen, W. Z.; Wang, C. X.; *J. Mol. Struct. THEOCHEM* **2005**, *725*, 9.
41. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T. H. L.; Lee, H.; Pedersen, L. G.; *J. Chem. Phys.* **1995**, *103*, 8577.
42. Martínez, L.; Polikarpov, I.; Skaf, M. S.; *J. Phys. Chem. B* **2008**, *112*, 10741.
43. Yang, Y.; Qin, J.; Liu, H.; Yao, X.; *J. Chem. Inf. Model.* **2011**, *51*, 680.
44. Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R.; *J. Chem. Phys.* **1995**, *103*, 4613.
45. Mabilia-Bassiloua, C.; Arthus-Cartier, G.; Hannaert, V.; Therisod, H.; Sygusch, J.; Therisod, M.; *ACS Med. Chem. Lett.* **2011**, *2*, 804.
46. St-Jean M.; Lafrance-Vanasse, J.; Liotard, B.; Sygusch, J.; *J. Biol. Chem.* **2005**, *280*, 27262.
47. Dalby, A. R.; Tolan, D. R.; Littlechild, J. A.; *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2001**, *57*, 1526.

Submitted: October 3, 2012

Published online: January 11, 2013