

Structural Basis for Selective Inhibition of Trypanosomatid Glyceraldehyde-3-Phosphate Dehydrogenase: Molecular Docking and 3D QSAR Studies

Rafael V. C. Guido, Glaucius Oliva, Carlos A. Montanari, and Adriano D. Andricopulo*

Laboratório de Química Medicinal e Computacional, Centro de Biotecnologia Molecular Estrutural, Instituto de Física de São Carlos, Universidade de São Paulo, Av. Trabalhador São-carlense 400, 13560-970, São Carlos-SP, Brazil, and Grupo de Química Medicinal de Produtos Naturais, Instituto de Química de São Carlos, Universidade de São Paulo, Av. Trabalhador São-carlense 400, 13566-970, São Carlos-SP, Brazil

Received December 6, 2007

The glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is as an attractive target for the development of novel antitrypanosomatid agents. In the present work, comparative molecular field analysis and comparative molecular similarity index analysis were conducted on a large series of selective inhibitors of trypanosomatid GAPDH. Four statistically significant models were obtained ($r^2 > 0.90$ and $q^2 > 0.70$), indicating their predictive ability for untested compounds. The models were then used to predict the potency of an external test set, and the predicted values were in good agreement with the experimental results. Molecular modeling studies provided further insight into the structural basis for selective inhibition of trypanosomatid GAPDH.

1. INTRODUCTION

Parasitic diseases are a major global cause of illness, morbidity, long-term disability, and death, with severe medical and psychological consequences for millions of men, women, and children.^{1–3} According to the World Health Organization, in both tropical and subtropical regions worldwide, leishmaniasis and trypanosomiasis (sleeping sickness and Chagas' disease) are serious and life-threatening diseases, affecting about 30 million people, with nearly another 400 million living in high-risk areas.^{4–6} In spite of the alarming health, economic, and social consequences of these parasitic infections, the limited existing drug therapy (e.g., antimonials, arsenicals, and nitroheterocyclic drugs) suffers from a combination of drawbacks including poor efficacy, resistance, and serious side effects. Therefore, there is an urgent need for new drugs that can overcome resistance and are safe and effective for use in humans.

The identification of novel molecular targets as well as safe and affordable drug candidates is of utmost importance in drug discovery.^{7–11} The glycolytic pathway of trypanosomatids has been investigated as a promising biochemical target for drug intervention.¹² The bloodstream form of trypanosomatids has no functional tricarboxylic acid cycle and is highly dependent on glycolysis for ATP production.^{13,14} This vital dependence on glycolysis as a source of energy makes the glycolytic enzymes attractive targets for drug design. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH; EC 1.2.1.12), one of the key enzymes in the glycolytic cascade, is a homotetramer that catalyzes the reversible NAD^+ -dependent oxidative phosphorylation of glyceraldehydes-3-phosphate to 1,3-diphosphoglycerate in the presence of inorganic phosphate. The glycolytic GAPDH enzyme from the pathogenic parasites *Trypanosoma cruzi*,¹⁵ *Trypanosoma brucei*,¹⁶ and

*Leishmania mexicana*¹⁷ are closely related (about 90% sequence identity) and possess important structural differences for drug design when compared to the homologue protein of the mammalian host (about 45% sequence identity).^{18,19}

Structure- and ligand-based approaches have become vital components of modern drug design.²⁰ Quantitative structure–activity relationships (QSARs) have been successfully used as valuable tools to assist the design of small-molecule inhibitors or receptor ligands having promise of utility in clinical medicine.^{21–27} In the present work, comparative molecular field analysis (CoMFA) and comparative molecular similarity index analysis (CoMSIA) have been performed to investigate the QSAR of a large series of adenosine analogs, which explore significant structural differences between trypanosomatids and human GAPDHs, leading to a selective inhibition of the parasite enzymes. Molecular modeling studies were also employed to study the molecular basis underlying the selective inhibition of trypanosomatid GAPDH.

2. RESULTS AND DISCUSSION

2.1. Chemical and Biological Data. The data set employed for the 3D QSAR CoMFA and CoMSIA analyses contains 70 adenosine analogs which were collected from five publications reported by one laboratory.^{28–32} The *in vitro* IC_{50} values employed in this work were measured under the same experimental conditions, a fundamental requirement for QSAR studies.^{24–26} Importantly, the reported statistical error limits have been calculated and add up to 10%. The family of GAPDH enzyme inhibitors under investigation has moderate structural diversity. The adenosine analogs bear bulk substituents linked to the N⁶ and 9 positions of the adenine system, exploring significant differences observed in the NAD^+ -binding-site vicinity of the parasite enzymes.

* Corresponding author. Tel.: +55 16 3373-8095. Fax: +55 16 3373-9881. E-mail: aandrico@ifsc.usp.br.

The structures and corresponding pIC_{50} values ($-\log \text{IC}_{50}$, where the IC_{50} is the concentration required for 50% inhibition of GAPDH) for the whole set of inhibitors are included in Table 1. The pIC_{50} values span about 3.5 orders of magnitude and are acceptably distributed across the range of values (Figure 1). The generation of consistent statistical models depends on the composition of the training and test sets. From the original data set, 56 inhibitors (compounds 1–56, Table 1) were selected as members of the training set for model generation, whereas the other 14 inhibitors (compounds 57–70, Table 1) were selected as members of the test set for external model validation. The compounds of both training and test sets were carefully selected in order to ensure appropriate property coverage on the entire range of pIC_{50} values. A statistical cluster analysis confirmed that the composition of both training and test sets is representative of the whole data set. The partial least squares (PLS) and leave-one-out (LOO) cross-validation methods were used for all 3D QSAR analyses.

2.2. Molecular Alignment Strategies. Due to the high dependency of trypanosomatids on glycolysis for survival, the energy production could be blocked through the action of a potent competitive inhibitor of an enzyme of the glycolytic pathway.³³ Glycolytic enzymes are potential targets for antitrypanosomatid drug design provided that reasonable parasite versus host selectivity can be obtained.^{28,33} It is important to note that the trypanosomatid GAPDHs possess low sequence identity in comparison with their mammalian host, as shown in Figure 2. The sequence alignments reveal that there are several highly conserved residues among the parasite enzymes that substantially differ from those of the human homologue. It is likely that SAR, QSAR, and structure-based drug design techniques would be useful for the design of GAPDH inhibitors that selectively target the trypanosomatid enzymes.²⁸

The determination of spatial molecular alignments is a crucial step in 3D QSAR studies, since the analyses are highly dependent on the quality of the alignments.^{22,23} Thus, molecular docking studies were carried out using the programs FlexX³⁴ and GOLD (Genetic Optimisation for Ligand Docking)^{35,36} in order to generate specialized molecular alignments for selective QSAR model building. The crystallographic coordinates of the inhibitor N^6 -(1-naphthalene methyl)-2'-deoxy-2'-(3,5-dimethoxy benzamido) (NMDBA, compound 36, Table 1) bound to the *L. mexicana* GAPDH binding site were employed for validating the docking studies.³⁷ Both FlexX and GOLD successfully docked the NMDBA inhibitor in good agreement with the experimentally determined X-ray crystal structure, placing the adenine moiety, as well as the N^6 and 9 substituents close to the bioactive conformation (Figure 3A and B). The heavy atom root-mean-square deviation for the best docking solution in comparison with the crystallographic conformation was 1.43 and 1.39 Å, according to FlexX and GOLD, respectively. On the basis of that, the molecular alignments were performed by applying the implemented scoring functions followed by the selection of the top-ranked docking solutions generated for each ligand of the data set. The final FlexX and GOLD alignments, respectively, are depicted in Figure 3C and D. As can be seen, the top-ranked conformations are reasonably well predicted within the *L. mexicana* GAPDH active site, confirming the robustness of the applied docking procedures.^{38–40}

2.3. 3D QSAR Analysis. **2.3.1. CoMFA and CoMSIA Models.** In order to investigate the structural and chemical features related to the biological activity of this series of trypanosomatid GAPDH inhibitors, the molecular alignments derived from both FlexX and GOLD docking simulations were quantitatively analyzed using the CoMFA^{41,42} and CoMSIA^{43,44} methods. These 3D QSAR techniques are based on the assumption that changes in binding affinities of ligands are related to changes in molecular properties represented by molecular fields. In the present study, the two molecular alignments were submitted to CoMFA and CoMSIA analyses, generating four QSAR models referred to as CoMFA_{FlexX}, CoMFA_{GOLD}, CoMSIA_{FlexX}, and CoMSIA_{GOLD}. The statistical results of the QSAR analyses are presented in Table 2.

Both 3D structural alignments yielded significant statistical CoMFA models, as indicated by the conventional non-cross-validated correlation (r^2) and cross-validated correlation (q^2_{LOO}) coefficients (CoMFA_{FlexX}, $r^2 = 0.93$ and $q^2_{\text{LOO}} = 0.83$; CoMFA_{GOLD}, $r^2 = 0.93$ and $q^2_{\text{LOO}} = 0.82$). The corresponding CoMSIA models derived for the same training set are also of good statistical quality (CoMSIA_{FlexX}, $r^2 = 0.92$ and $q^2_{\text{LOO}} = 0.72$; CoMSIA_{GOLD}, $r^2 = 0.93$ and $q^2_{\text{LOO}} = 0.79$). Although the CoMFA models had slightly better statistical parameters in comparison with the corresponding CoMSIA models, in accordance with similar results reported in the literature for other data sets,^{43–47} the results shown in Table 2 indicate that the models were not substantially affected by the different alignment protocols (FlexX and GOLD). This may be due to the fact that the training set molecules incorporate sufficient structural information to allow privileged molecular recognition within the binding site of *L. mexicana* GAPDH, resulting in similar field descriptors that explain the variance of the 3D QSAR models.

2.3.2. QSAR Model Validation. Leave-many-out cross-validation (q^2_{LMO}) employing 10 or 5 groups (LMO₁₀ or LMO₅, respectively) was performed as a more rigorous test for the stability and statistical significance of the models. Accordingly, the data set was divided into 10 and 5 randomly selected groups of equal size, and subsequently, each group was left out once during the cross-validation process. Each model was evaluated 25 times by measuring its accuracy in predicting the activity of the remaining 10% and 20% data set compounds (Table 3). The results confirm the stability and reliability of the 3D QSAR models, considering that the statistical values obtained for the LMO analyses are comparable to those of the LOO analysis.

Progressive scrambling of the data set was applied to determine the sensitivity of the QSAR models to chance correlations, and to test the stability of the models.²³ The results of the progressive scrambling further confirmed consistency of the models as defined by the critical slope, and optimum statistics for cross-validated standard deviation of error prediction and Q^2 obtained at the end of several different runs. Finally, to validate the models' accuracy, the bootstrapping function was used to determine the error on the r^2 (r^2_{boot}) and the standard error of estimate (SEE_{boot}) of each model. An average r^2_{boot} value of 0.95 (SEE_{boot} of about 0.25) was obtained for both CoMFA and CoMSIA models. These results indicate that the QSAR models generated in this study are very stable and possess substantial statistical significance and high internal predictive ability.
























Table 1. Chemical Structures and Corresponding pIC₅₀ Values for a Series of Inhibitors of *L. mexicana* GAPDH

Training Set Compounds					
Cpd	Structure	pIC ₅₀	Cpd	Structure	pIC ₅₀
1		3.30	2		2.52
3		3.15	4		2.40
5		2.80	6		2.22
7		3.40	8		3.60
9		3.12			

Training Set Compounds				
Cpd	R ¹	R ²	R ³	pIC ₅₀
10	CH ₃	H	H	2.22
11	H	H		2.40
12	H		H	2.43
13	H		H	3.15
14	H		H	3.15
15	H		H	2.74
16	H		H	3.44
17	H		H	3.47
18	H		H	3.62

Training Set Compounds							
Cpd	R ¹	R ²	R ³	R ⁴	R ⁵	R ⁶	pIC ₅₀
19	H	H	H	H	H	H	2.48
20	H	H	H	H	C ₆ H ₅	H	3.70
21	H	H	H	OCH ₂ CH ₃	H	H	3.40
22	H	H	H	OCH ₃	OH	H	3.30
23	H	Br	H	H	H	H	2.52
24	H	H	H	OH	OH	OH	3.60
25	H	H	H	OH	H	OH	3.19
26		H	H	OCH ₃	H	OCH ₃	4.70
27		H	H	OCH ₃	H	H	4.80
28		H	H	OCH ₃	H	OCH ₃	4.60
29		H	H	OCH ₃	H	H	4.60
30		H	H	OCH ₃	H	OCH ₃	4.60

Table 1. Continued

Training Set Compounds								Training Set Compounds							
Cpd	R ¹	R ²	R ³	R ⁴	R ⁵	R ⁶	pIC ₅₀	Cpd	R ¹	R ²	R ³	R ⁴	R ⁵	R ⁶	pIC ₅₀
31		H	H	OCH ₃	H	H	4.60	51		H	H	OCH ₃	H	OCH ₃	5.70
32		H	H	OCH ₃	H	H	5.26	52		H	H	Cl	H	Cl	4.60
33		H	H	OCH ₃	H	H	3.82	53		H	H	OCH ₃	H	H	5.70
34		H	H	OCH ₃	H	H	4.10	54		H	H	Cl	H	Cl	4.60
35		H	H	Cl	H	H	5.00	55		H	H	OCH ₃	H	H	5.30
36		H	H	OCH ₃	H	OCH ₃	5.70	56		H	H	OCH ₃	H	H	4.08
37		H	H	Cl	H	Cl	4.92								
38		H	H	OAc	H	OAc	5.00								
39		H	H	OH	H	OH	5.30								
40		H	H	H	H	OH	4.10								
41		H	H	OAc	H	OAc	4.60								
42		H	H	OH	H	OCH ₃	4.43								
43		H	H	OCH ₃	H	OH	5.00								
44		H	H	N(CH ₃) ₂	H	H	4.60								
45		H	H	OCH ₃	H	OCH ₃	5.70								
46		H	H	Cl	H	Cl	4.60								
47		H	H	OCH ₃	H	OCH ₃	4.22								
48		H	H	OCH ₃	H	H	5.22								
49		H	H	Cl	H	H	5.40								

Test Set Compounds								
Cpd	Structure	pIC ₅₀	Cpd	Structure	IC ₅₀	Cpd	Structure	pIC ₅₀
57								

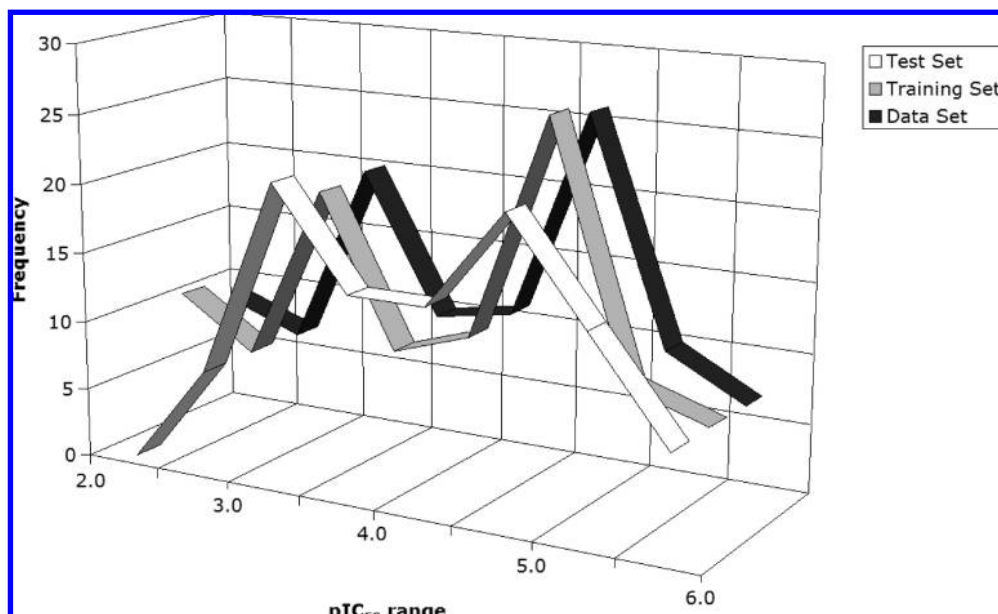


Figure 1. Distribution of pIC_{50} values for the training set, test set, and complete data set.

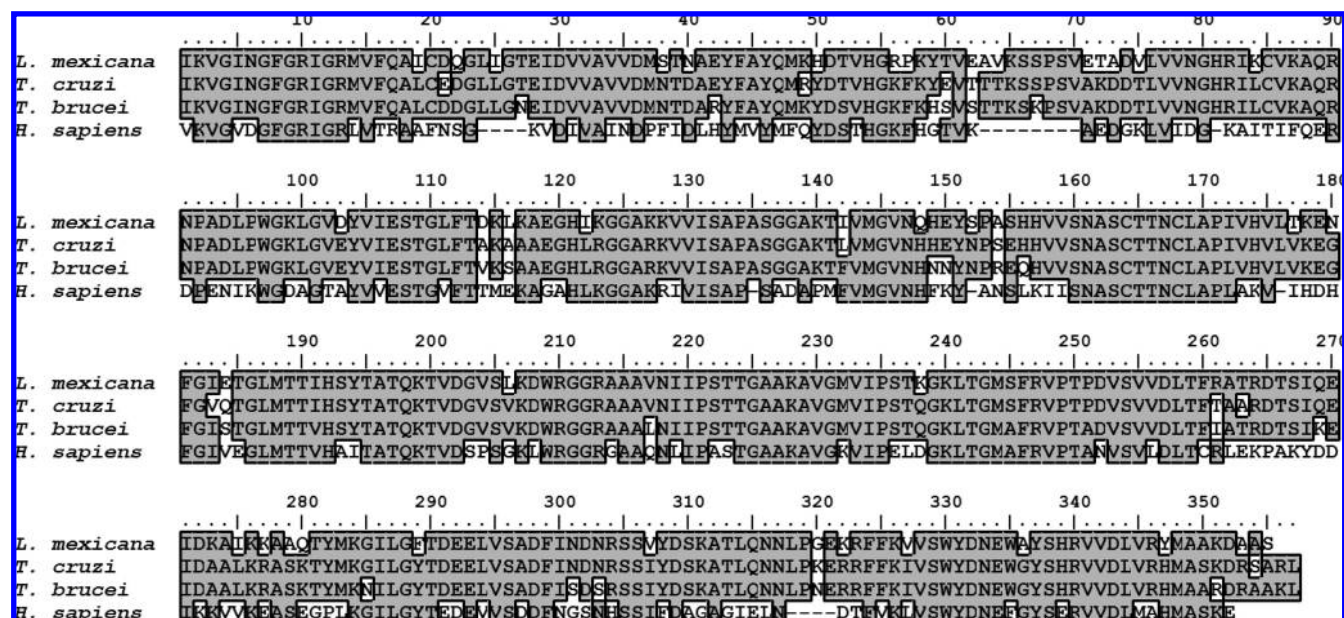


Figure 2. Sequence alignments of *L. mexicana*, *T. cruzi*, *T. brucei*, and *H. sapiens* GAPDHs. Numbers above the sequences refer to the *L. mexicana* GAPDH sequence. Residues conserved among the different species are highlighted in gray.

Although the q^2_{LOO} and q^2_{LMO} procedures may give a suitable representation of the predictive power of the models for untested GAPDH inhibitors, the external validation process can be considered the most valuable validation method.^{21–23} Therefore, the predictive ability of the models was assessed by predicting the potency of an external test set of 14 compounds (57–70, Table 1), which were completely excluded from model generation. On the basis of the appropriate representation of chemical diversity and distribution of the biological property (Figure 1), the test set meets the requirements for the purpose of external model validation.^{21–27} The results of the external validation are listed in Table 4, and the graphic results are simultaneously displayed in Figure 4.

Predictive r^2 values of 0.86, 0.83, 0.80, and 0.78 were obtained for CoMFA_{FlexX}, CoMSIA_{FlexX}, CoMFA_{GOLD}, and CoMSIA_{GOLD}, respectively (Table 3). The good agreement

between experimental and predicted pIC_{50} values for the test set compounds indicates the reliability of the QSAR models, as expressed by the high predictive r^2 values. This is probably the best statistical parameter to assess the external predictive power of QSAR models.^{48–51} Hence, in addition to the good statistical quality and internal predictive ability, the models show excellent predictive power for novel GAPDH inhibitors within this structural diversity.

2.3.3. 3D Contour Maps. The understanding of protein–ligand interactions is essential for the design of new inhibitors with improved potency.²⁰ Besides predicting the property value of interest for untested molecules, 3D QSAR models should also provide insights into molecular properties directly related to biological activity. One prominent feature of the 3D QSAR methods employed in this study is that they allow easy visualization of regions in space responsible for increases or decreases in the values of a particular type of

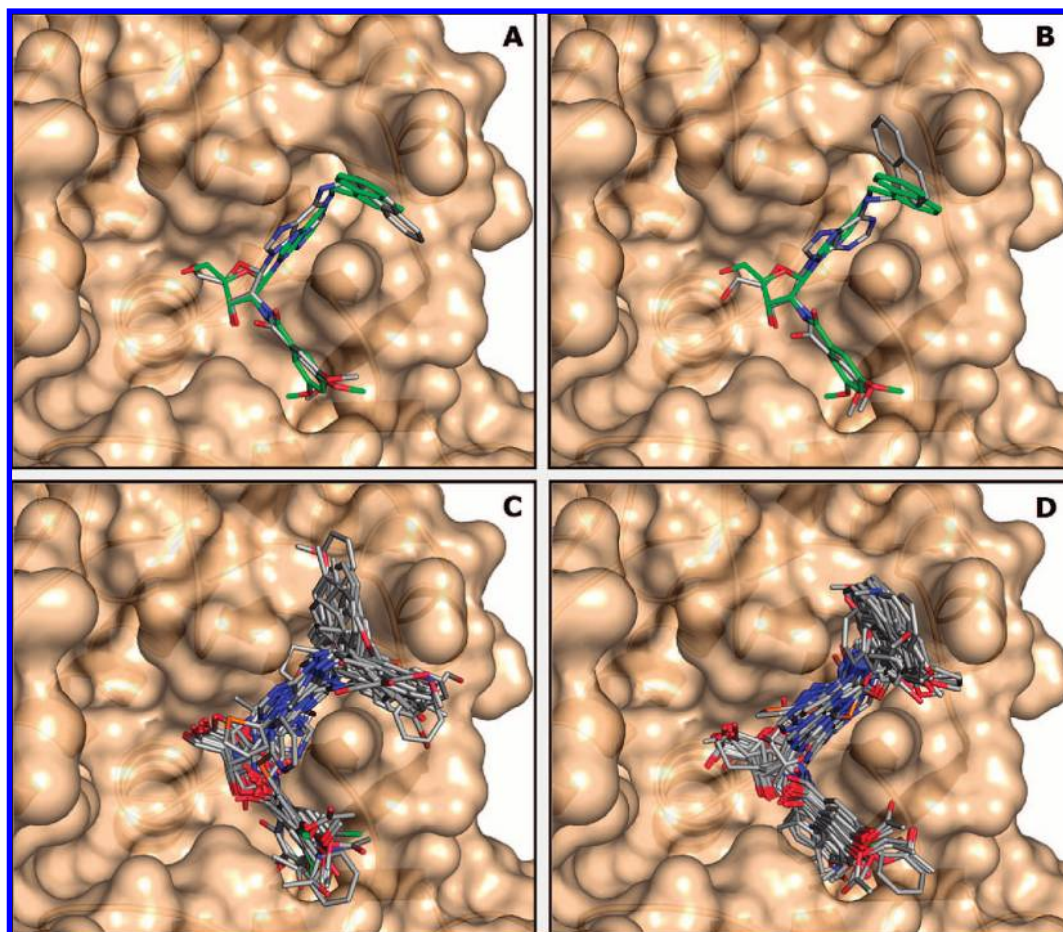


Figure 3. X-ray crystal structure 1I32 of the NMDBA inhibitor (green) bound to *L. mexicana* GAPDH and the best docking solution of NMBDA (gray) generated by FlexX (A) and GOLD (B). FlexX (C) and GOLD (D) data set molecular alignments are shown inside the Connolly surface of the adenosine portion of the NAD⁺ binding pocket.

Table 2. CoMFA and CoMSIA Statistical Results for the FlexX and GOLD 3D Alignments^a

statistical parameters	FlexX		GOLD	
	CoMFA	CoMSIA	CoMFA	CoMSIA
q^2_{LOO}	0.83	0.72	0.82	0.79
r^2	0.93	0.92	0.93	0.93
SEE	0.27	0.30	0.29	0.29
F	259	153	218	167
N	3	4	3	4
fraction (%)				
S	44	18	48	15
E	56	26	52	27
H		26		26
H donor		18		21
H acceptor		12		11

^a q^2 , leave-one-out (LOO) cross-validated correlation coefficient; r^2 , non-cross-validated correlation coefficient; N , optimum number of components; SEE, standard error of estimate; F , F test value; **S**, steric field; **E**, electrostatic field; **H**, hydrophobic field.

Table 3. Statistical Data for Validation Procedure of CoMFA and CoMSIA Models^a

statistical parameters	FlexX		GOLD	
	CoMFA	CoMSIA	CoMFA	CoMSIA
q^2 LMO ₁₀	0.82 ± 0.01	0.72 ± 0.02	0.82 ± 0.01	0.79 ± 0.01
q^2 LMO ₅	0.82 ± 0.01	0.70 ± 0.03	0.82 ± 0.01	0.78 ± 0.02
r^2_{boot}	0.95 ± 0.01	0.95 ± 0.01	0.95 ± 0.01	0.95 ± 0.01
SEE _{boot}	0.24 ± 0.11	0.25 ± 0.11	0.23 ± 0.11	0.23 ± 0.11
predictive r^2	0.86	0.83	0.80	0.78

^a LMO, mean and standard deviation values of q^2 obtained after 25 independent runs of cross-validation protocols. Bootstrapping, mean, and standard deviation values of r^2 and SEE obtained after 100 simulation samplings.

dependent variable (e.g., pIC₅₀). Considering that both alignment protocols generated models of similar statistical quality, the discussion about the CoMFA and CoMSIA contour maps will be restricted to the models derived with the FlexX alignment, because they show slightly better statistical parameters and predictive power. The CoMFA (steric and electrostatic) and CoMSIA (steric, electrostatic, hydrophobic, and hydrogen-bond donor and acceptor) con-

tour maps are displayed as PLS standard deviation × coefficient maps in Figures 5 and 6, respectively.

2.3.3.1. CoMFA Contour Maps. The CoMFA steric contour map is shown in Figure 5A. The green contour surrounding the naphthyl substituent of the inhibitor **36** (pIC₅₀ = 5.70) indicates that bulky substituents would be favorable in this region. Conversely, unfavorable steric regions located near the 8 position of the adenine moiety suggest that bulky substituents at this position are detrimental for potency. In fact, in the proposed binding mode of compound **23** (pIC₅₀ = 2.52), the bromine substituent is oriented toward this unfavorable region. Thus, the absence of bulky groups in the sterically favorable region combined with the presence of a steric unfavorable substituent in the

Table 4. Experimental and Predicted Activities (pIC_{50}) with Residual Values for the Test Set Compounds

cpd	exp ^a	FlexX				GOLD			
		CoMFA		CoMSIA		CoMFA		CoMSIA	
		pred ^b	res ^c	pred ^b	res ^c	pred ^b	res ^c	pred ^b	res ^c
57	2.62	3.18	-0.56	3.31	-0.69	3.11	-0.49	2.89	-0.27
58	3.52	3.45	0.08	3.65	-0.13	3.42	0.10	3.26	0.26
59	3.30	3.30	0.00	3.17	0.13	2.68	0.62	2.80	0.50
60	3.22	2.95	0.27	3.13	0.09	3.38	-0.16	3.17	0.05
61	3.07	3.17	-0.10	3.29	-0.22	3.17	-0.10	3.11	-0.04
62	3.82	3.60	0.22	3.27	0.55	3.08	0.74	3.28	0.54
63	5.30	4.94	0.36	5.25	0.05	4.86	0.44	5.20	0.10
64	5.40	4.93	0.47	4.93	0.47	4.88	0.52	4.63	0.77
65	5.70	5.66	0.04	5.42	0.28	5.53	0.17	5.30	0.40
66	4.30	4.62	-0.32	4.75	-0.45	4.96	-0.66	5.05	-0.75
67	4.60	5.07	-0.47	4.91	-0.31	4.66	-0.06	5.04	-0.44
68	4.00	4.56	-0.56	4.45	-0.45	4.36	-0.36	4.40	-0.40
69	4.74	4.46	0.28	4.27	0.47	4.84	-0.10	5.12	-0.38
70	5.00	4.57	0.43	4.44	0.56	4.55	0.46	4.46	0.54

^a Experimental value. ^b Predicted value. ^c Residual: the difference between experimental and predicted values.

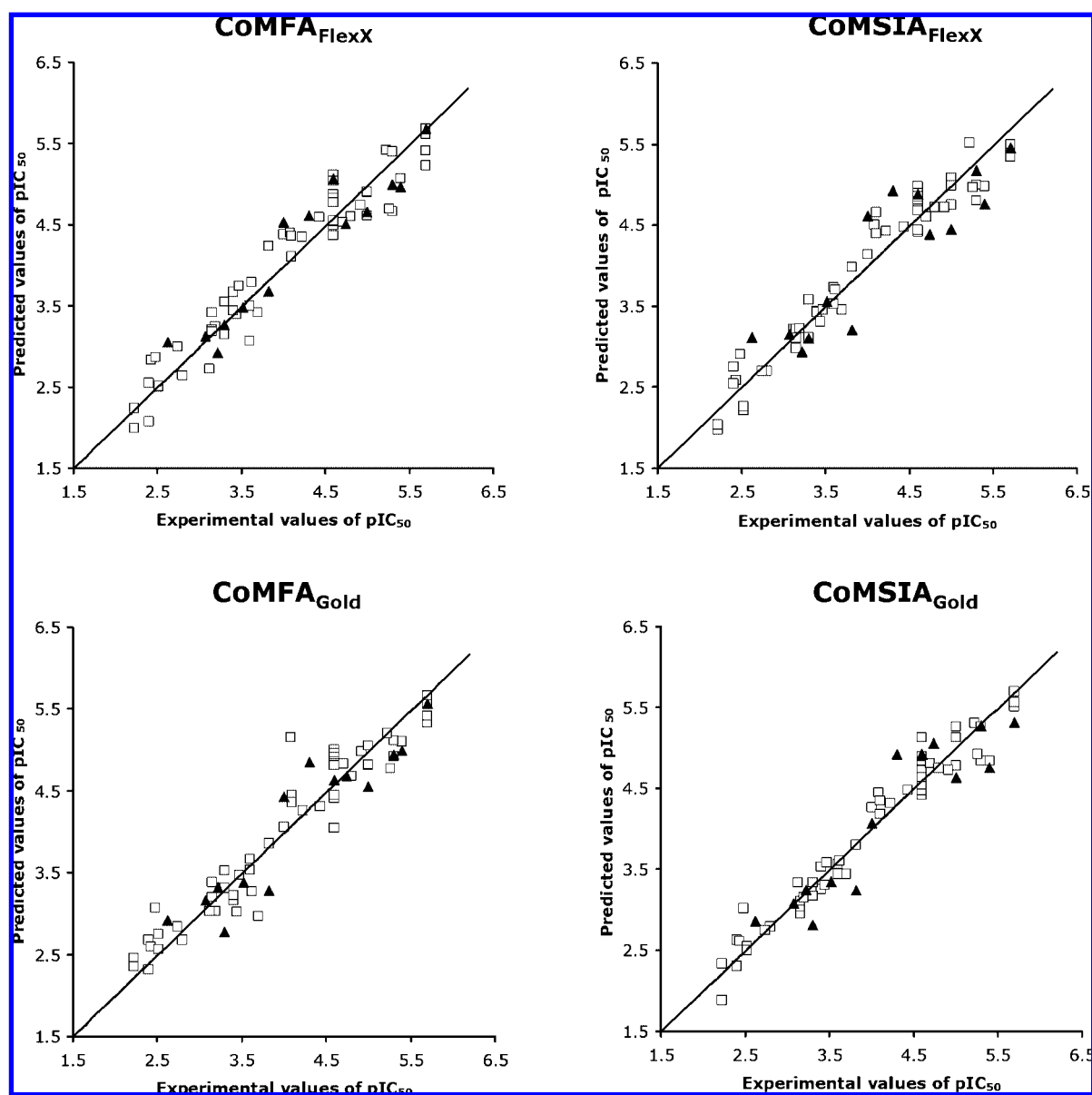


Figure 4. Plot of predicted values of pIC_{50} versus the corresponding experimental values for the 56 training set inhibitors (open square) and 14 test set inhibitors (solid triangles).

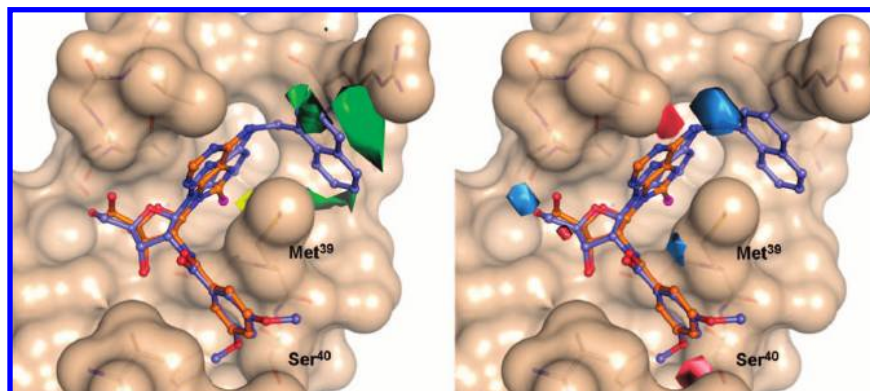


Figure 5. 3D contour maps (standard deviation \times coefficient) for CoMFA_{FlexX}. The inhibitors **36** (in light blue) and **23** (in orange) are represented by ball-and-stick models in the active site of *L. mexicana* GAPDH. The protein Connolly surface is indicated. The hydrogen atoms were omitted for clarity. (A) CoMFA steric contour maps are shown in yellow and green (-0.005 and 0.035 kcal/mol, respectively); (B) CoMFA electrostatic contour maps are shown in red and blue (-0.05 and 0.05 kcal/mol, respectively).

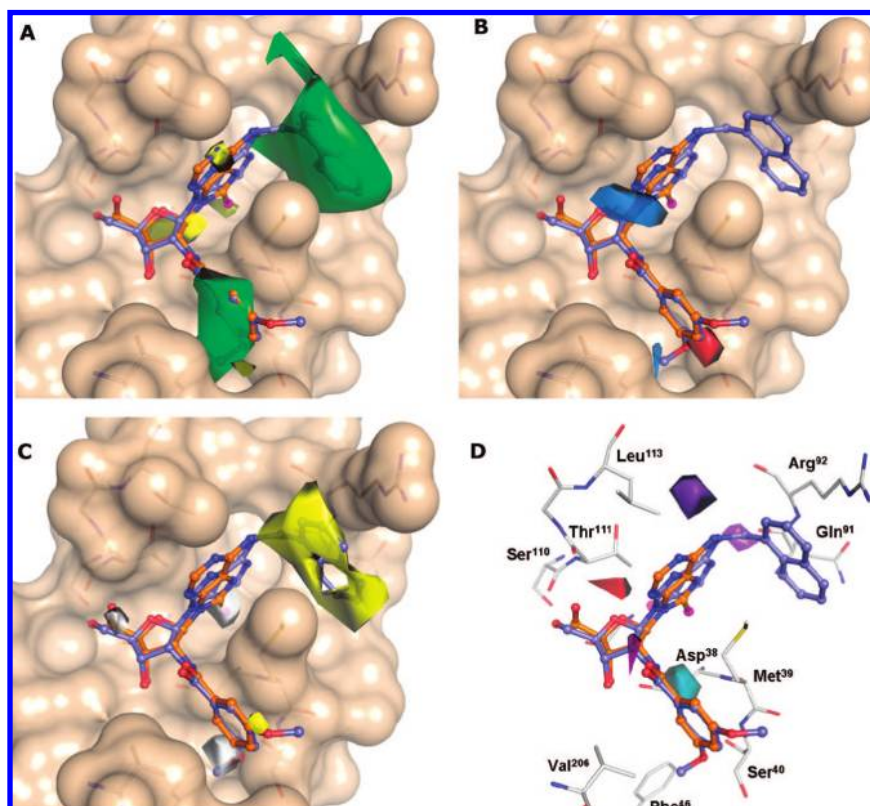


Figure 6. 3D contour maps (standard deviation \times coefficient) for CoMSIA_{FlexX}. The GAPDH inhibitors model within the *Lm*-GAPDH binding site. The inhibitors **36** (in light blue) and **23** (in orange) are represented by ball-and-stick models in the active site of *L. mexicana* GAPDH. The protein-solvent-accessible surface is indicated as a solid surface. The hydrogen atoms were omitted for clarity. (A) CoMSIA steric contour maps are shown in yellow and green (-0.0014 and 0.008 kcal/mol, respectively); (B) CoMSIA electrostatic contour maps are shown in red and blue (-0.03 and 0.02 kcal/mol, respectively). (C) CoMSIA hydrophobic contour maps are shown in white and yellow (-0.01 and 0.015 kcal/mol, respectively); (D) CoMSIA hydrogen-bond donor and acceptor contour maps. H-donor contour fields are shown in purple and cyan (-0.015 and 0.020 kcal/mol, respectively). H-acceptor contour fields are shown in red and magenta (-0.07 and 0.015 kcal/mol, respectively).

adenine ring could explain the poor inhibitory activity exhibited by compound **23** (Figure 5A). The CoMFA electrostatic contour plots are shown in Figure 5B. The map highlights three favorable regions for electropositive groups (blue contour). Two of them are solvent-exposed, one being near the 2 position of the naphthyl ring and the other near the 6'-hydroxyl of the ribose ring. The third favorable region can be found surrounding the sulfur in the side chain of Met39, indicating that electropositive groups at this position may enhance the inhibitory potency. Furthermore, there are two favorable regions for negatively charged groups: one,

solvent exposed, near the N⁶-position of the adenine moiety, and the other, located near to the side-chain hydroxyl group of Ser40. The 3-methoxy benzamide substituent of **36** is close to this electronegatively favorable region. The contour map also agrees with previous SAR studies, which indicated that the absence of electronegative groups in the benzamide ring (compounds **19** and **23**) is remarkably unfavorable for inhibitory potency within this series of trypanosomatid GAPDH inhibitors.^{28,29}

2.3.3.2. CoMSIA Contour Maps. The chemical environment of the enzyme binding site is represented by five

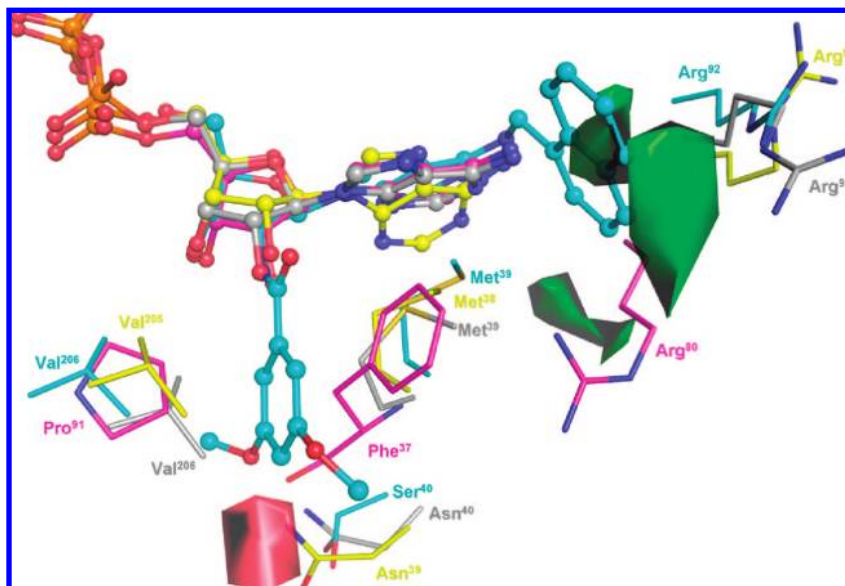


Figure 7. 3D contour maps (standard deviation \times coefficient) for the CoMFA_{FlexX} model showed within the active site of the superposed crystallographic structures of trypanosomatid GAPDHs. The GAPDH structures are shown as *L. mexicana*, in cyan (PDB: 1I32); *T. brucei* in yellow (PDB: 1GGA); *T. cruzi* in gray (PDB: 1ML3); and human, in magenta (PDB: 1U8F). The inhibitor **36** (cyan) and the cofactor NAD⁺ are illustrated as ball-and-stick models. The hydrogen atoms were omitted for clarity. The CoMFA steric favorable map is shown in green, whereas the favorable CoMFA electronegative contour map is shown in red (+0.035 and -0.05 kcal/mol, respectively).

molecular fields in the CoMSIA method (steric, electrostatic, hydrophobic, and hydrogen-bond donor and acceptor). Figure 6A shows the CoMSIA steric contour plots, which agree with the contour plots of the corresponding CoMFA map, highlighting two important steric regions (green isopleths). The first one is positioned between the side chains of Met39 and Arg92, while the second one lies in the lower portion of the binding pocket, at the so-called “selective cleft”, between the chains of Met39 and Val206. Ligands possessing bulky groups in these regions have enhanced inhibitory potency (compounds **26–55**, **63–70**).^{30–32} The CoMSIA unfavorable steric map also highlights the 8-bromine substituent from compound **23** as detrimental for activity.

The CoMSIA electrostatic contour map of Figure 6B shows that inhibitors having either negatively charged groups in the red regions or positively charged groups in the blue regions would substantially enhance inhibitory potency toward the parasite enzymes. As previously noted for the CoMFA electrostatic contour map of Figure 5A, compound **36** orients the electronegative oxygen of the 3-methoxyl benzamide group in a favorable region for negative charges. In addition, the methyl of the 3-methoxy group is located in the favorable positively charged lower portion of the binding site.

The yellow and white isopleths in Figure 6C indicate that hydrophobic and hydrophilic regions correlate with inhibitory potency. There are two hydrophobically favorable regions (yellow contours); one encompasses the N⁶ substituent of the adenine system, while the other surrounds the 5-methoxyl of the benzamide moiety. The former region partially overlaps with the steric volume favoring bulk groups, thus, suggesting that ligands with bulk and hydrophobic substituents at this site would be beneficial for enhanced potency. This corroborates with the chemical environment of the binding site and agrees with experimental results.^{28–32} The other region favoring hydrophobic contacts is located near the side chain of Met39. In addition, the contour map displays two unfavorable hydrophobic regions. One region can be

observed around the bromine substituent of compound **23**, while the other partly overlaps with the favored partially negatively charged region, suggesting that electron-withdrawing groups with hydrophilic properties would be beneficial at this region.

The CoMSIA hydrogen-bond contour map is illustrated in Figure 6D, being discriminated for donor (cyan and purple) or acceptor (magenta and red) regions. The presence of cyan regions near the carboxyl group in the side chain of Asp38 denotes the regions where hydrogen-bond acceptors might exist in the receptor, and thus a hydrogen-bond-donating substituent could be favored. In fact, the carboxyl group of the Asp38 is an important interaction point for molecular recognition toward the trypanosomatid GAPDHs.³⁷ A magenta isopleth, favorable for hydrogen-acceptor groups, is in good agreement with the binding site environment, suggesting the main-chain carbonyl group of Gln91 as a reasonable site for hydrogen-bond formation.

2.3.4. 3D QSAR Selectivity Map. The structure-based GAPDH inhibitor design can take advantage of the substantial structural differences between trypanosomatid and human enzymes. Selective inhibitors of *L. mexicana*, *T. brucei*, and *T. cruzi* GAPDH enzymes were obtained on the basis of adenosine analogs bearing functional groups that bind in the intersubunit (selectivity cleft) of the NAD⁺-binding site.^{28–32,37} Due to the fact that the human enzyme possesses a more narrow “selectivity cleft”, these inhibitors selectively bind to the trypanosomatid GAPDHs. In order to obtain further insights into the molecular basis underlying the selective inhibition of trypanosomatid GAPDH enzymes, a detailed analysis of the most relevant 3D QSAR fields was carried out through the superposition of the binding sites of *T. cruzi*,¹⁵ *T. brucei*,¹⁶ *L. mexicana*,¹⁷ and human¹⁸ GAPDHs (Figure 7). For this study, we have employed the corresponding steric and electrostatic fields of the CoMFA_{FlexX} model (the best QSAR model generated), which have significant complementarity to the GAPDH binding site.

As can be seen, the CoMFA electrostatic field descriptor explained 56% of the variance, while the proportion of the steric descriptor accounted for the remaining 44% (Table 2). According to the aforementioned CoMFA and CoMSIA results, the biological potency of the inhibitors is directly affected by the electrostatic fields. The electronegatively favorable region (red contour) located within the superposed binding sites is in accordance with the chemical environment of the GAPDH parasite enzymes. The map is near the NH_2 group in the side chain of the Asn residues (Asn40 in *T. cruzi* and Asn39 in *T. brucei*), and also near the OH group in the side chain of the Ser40 residue in *L. mexicana* GAPDH. Conversely, this electronegatively favorable region near the carbonyl group in the main chain of Phe37 of the human homologue suggests that para-substituted benzamides with partially negative charged groups may be exploited to enhance both potency and selectivity toward the trypanosomatid enzymes. The CoMFA steric field is also in good agreement with the binding site of the parasite enzymes, indicating that bulky substituents are able to fit between the side chains of the Arg92 and Met39 (Met38 in *T. brucei*) residues. In contrast, this region is sterically hindered in the human enzyme by a bulkier side chain of the Phe37 residue.

3. CONCLUSION

In this study, CoMFA and CoMSIA QSAR models were developed on the basis of the docked conformations of 70 adenosine analogs as trypanosomatid GAPDH inhibitors. The 3D QSAR models described possess high internal and external consistency. The good correlation between experimental and predicted pIC_{50} values for the test set compounds further indicated the robustness of the QSAR models. The CoMFA and CoMSIA contour maps, representing the effects of the steric, electrostatic, hydrophobic, and hydrogen-bond donor and acceptor molecular fields, are in good agreement with the chemical environment of the GAPDH binding site and, therefore, could be useful to guide further structural modification as well as the structure-based design of new selective trypanosomatid-selective inhibitors of GAPDH. The integration of molecular docking and 3D QSAR is a useful tool for the generation of predictive models and for the identification of key structural features responsible for selectivity and biological potency.

4. MATERIALS AND METHODS

4.1. Data Set. The data set used for the QSAR studies contains 70 inhibitors of GAPDH, which were selected from the literature.^{28–32} The structures and corresponding pIC_{50} values for the complete set of inhibitors are included in Table 1. The complete data set was divided into training (compounds 1–56, Table 1) and test (compounds 57–70, Table 1) sets in the ratio of 5:1 (20%), and the pIC_{50} values were used as dependent variables in the QSAR analyses.

4.2. Computational Approach. The QSAR modeling analyses, calculations, and visualizations for CoMFA and CoMSIA were performed using the SYBYL 7.3 package (Tripos Inc., St. Louis, MO) running on Red Hat Enterprise Linux workstations. The 3D structures of the inhibitors were constructed using standard geometric parameters of the molecular modeling software package SYBYL 7.3. Each single-optimized conformation of each molecule in the data set was

energetically minimized employing the Tripos force field⁵² and the Powell conjugate gradient algorithm⁵³ with a convergence criterion of 0.05 kcal/mol·Å. Partial atomic charges were calculated by the Gasteiger–Hückel method in the SYBYL 7.3 program.⁵⁴ Molecular modeling studies were carried out with FlexX (Tripos Inc., St. Louis, MO) and GOLD 2.1 (Cambridge Crystallographic Data Centre, Cambridge, U.K.). With the default parameters of FlexX and GOLD, the docking procedures were repeated 30 times for each data set compound. The implemented FlexX and GOLDScore scoring functions were employed to select the representative conformation for each compound. The FlexX and GOLD top-ranked poses were then used to produce the statistically significant alignments. A statistical cluster analysis was carried out with Tsar 3D, version 3.3 (Accelrys, San Diego, CA), using the complete linkage clustering method (Euclidean distances) with no data standardization.

4.3. Molecular Modeling and Structural Alignment.

The data set molecular alignment protocols applied were based on the top-ranked conformation selected by the FlexX and GOLD 2.1 docking programs. The crystallographic data of *L. mexicana* GAPDH in complex with NMDBA (compound 51, Table 1) were used as a molecular target for all docking simulations (PDB code: 1I32).³⁷ Hydrogen atoms were added in standard geometry using the Biopolymer module as implemented in SYBYL 7.3. The histidine, glutamine, and asparagine residues in the active binding site were manually checked for possible flipped orientation, protonation, and tautomeric states with a Pymol 0.99 (DeLano Scientific, San Carlos, CA) side-chain wizard script. The active site was defined incorporating all amino acid residues within a radius sphere of 8.0 Å centered on the bound ligand. In both docking approaches, default parameters were employed, and only the top-ranked conformation of each inhibitor was considered for 3D QSAR studies.

4.4. 3D QSAR: CoMFA and CoMSIA Analyses. The CoMFA and CoMSIA analyses were carried out on the basis of the interaction energies between a suitable probe atom and the aligned ligand atoms calculated within a grid box of 2 Å spacing. The grid box embedded all ligands with a margin of at least 4 Å in each direction. A positively charged sp^3 carbon probe atom was employed to calculate steric and electrostatic CoMFA fields using SYBYL standard parameters (TRIPOS standard field, dielectric constant 1/ r , cutoff 30 kcal/mol). CoMSIA fields were computed for steric, electrostatic, hydrophobic, and hydrogen bonding properties, using a probe of charge +1, a radius of 1, hydrophobicity and hydrogen bonding properties of +1, and an attenuation factor of 0.3 for the Gaussian distance-dependent function. According to each alignment, the pIC_{50} values were correlated with the field descriptors using SAMPLS⁵⁵ in a LOO cross-validation analysis. After determining the optimal number of components, a PLS analysis was performed without cross-validation, applying no column filtering. A progressive scrambling method⁵⁶ was employed to determine the sensitivity of the QSAR models to chance correlations. This validation technique addresses the overly optimistic cross-validation or response randomization results for redundant data sets. In this approach, small random perturbations are introduced into a data set. This causes the nominal predictivity of unstable models to fall off rapidly, whereas robust models are relatively stable. LMO with 10 (LMO₁₀)

and 5 (LMO₅) randomly selected groups were used as a more rigorous test to assess model stability and statistical significance. Each random cross-validation run was repeated 25 times to obtain mean values for q^2 and the corresponding SDEP. The bootstrapping procedure was used to validate each model. In this procedure, n random selections of the original set of n objects were performed several times (100 times was used to produce consistent statistical information) to simulate different sampling from a larger set of objects. In each individual run, some objects may not be included in the PLS analysis (same method to determine the q^2), whereas other objects can be included more than once. Confidence intervals for each term were estimated through this procedure, providing an independent measure for the stability of the PLS model.^{57,58} The CoMFA and CoMSIA contour maps were generated by interpolation of the pairwise products between the PLS coefficients and standard deviations of the corresponding descriptors values. External model validation was performed with a test set of 14 compounds, which were not considered for QSAR model generation. After generation of the PLS training set models, the dependent variables (pIC₅₀) were predicted for the test set compounds, allowing predictive- r^2 values to be determined for the individual 3D QSAR models.

ACKNOWLEDGMENT

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 Pesquisa e Desenvolvimento), Brazil. We also thank Prof. Gerhard Klebe and his research group (AGKlebe) from the University of Marburg—Germany, for important scientific support and very stimulating discussions.

REFERENCES AND NOTES

- (1) Watkins, B. M. Drugs for the Control of Parasitic Diseases: Current Status and Development. *Trends Parasitol.* **2003**, *19*, 477–478.
- (2) Nicholson, B. D.; Walley, J. D.; Baguley, D. S. Leishmaniasis, Chagas disease and Human African Trypanosomiasis revisited: Disease Control Priorities in Developing Countries. *Trop. Med. Int. Health* **2006**, *11*, 1339–1340.
- (3) McKerrow, J. H.; Caffrey, C.; Kelly, B.; Loke, P.; Sajid, M. Proteases in Parasitic Diseases. *Annu. Rev. Pathol. Mech. Dis.* **2006**, *1*, 497–536.
- (4) Croft, S. L.; Coombs, G. H. Leishmaniasis - Current Chemotherapy and Recent Advances in the Search for Novel Drugs. *Trends Parasitol.* **2003**, *19*, 502–508.
- (5) *Control of Chagas Disease*; World Health Organization, Tech. Rep. Ser.: Geneva, Switzerland, 2002; Vol. 905, pp 1–109.
- (6) Sanz-Rodriguez, C. E.; Concepcion, J. L.; Pekerar, S.; Oldfield, E.; Urbina, J. A. Bisphosphonates as Inhibitors of *Trypanosoma cruzi* Hexokinase: Kinetic and Metabolic Studies. *J. Biol. Chem.* **2007**, *282*, 12377–12387.
- (7) Andricopulo, A. D.; Montanari, C. A. Structure-Activity Relationships for the Design of Small-Molecule Inhibitors. *Mini-Rev. Med. Chem.* **2005**, *5*, 585–593.
- (8) Marton, M. J.; DeRisi, J. L.; Bennett, H. A.; Iyer, V. R.; Meyer, M. R.; Roberts, C. J.; Stoughton, R.; Burchard, J.; Slade, D.; Dai, H.; Bassett, D. E., Jr.; Hartwell, L. H.; Brown, P. O.; Friend, S. H. Drug Target Validation and Identification of Secondary Drug Target Effects Using DNA Microarrays. *Nat. Med.* **1998**, *4*, 1293–1301.
- (9) Brinkman, R. R.; Dube, M. P.; Rouleau, G. A.; Orr, A. C.; Samuels, M. E. Human Monogenic Disorders - a Source of Novel Drug Targets. *Nat. Rev. Genet.* **2006**, *7*, 249–260.
- (10) Wadman, M. Drive for Drugs Leads to Baby Clinical Trials. *Nature* **2006**, *440*, 406–407.
- (11) Reichert, J. M. Trends in Development and Approval Times for New Therapeutics in the United States. *Nat. Rev. Drug Discovery* **2003**, *2*, 695–702.
- (12) Verlinde, C. L. M. J.; Hannaert, V.; Blonski, C.; Willson, M.; Perie, J. J.; Fothergill-Gilmore, L. A.; Opperdoes, F. R.; Gelb, M. H.; Hol, W. G.; Michels, P. A. M. Glycolysis as a Target for the Design of New Anti-Trypanosome Drugs. *Drug Resist. Updates* **2001**, *4*, 50–65.
- (13) Bakker, B. M.; Mensonides, F. I.; Teusink, B.; van Hoek, P.; Michels, P. A. M.; Westerhoff, H. V. Compartmentation Protects Trypanosomes from the Dangerous Design of Glycolysis. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2087–2092.
- (14) Michels, P. A. M.; Hannaert, V.; Bringaud, F. Metabolic Aspects of Glycosomes in Trypanosomatidae - New Data and Views. *Parasitol. Today* **2000**, *16*, 482–489.
- (15) Souza, D. H. F.; Garratt, R. C.; Araujo, A. P.; Guimaraes, B. G.; Jesus, W. D.; Michels, P. A. M.; Hannaert, V.; Oliva, G. *Trypanosoma cruzi* Glycosomal Glyceraldehyde-3-Phosphate Dehydrogenase: Structure, Catalytic Mechanism and Targeted Inhibitor Design. *FEBS Lett.* **1998**, *424*, 131–135.
- (16) Vellieux, F. M.; Hajdu, J.; Verlinde, C. L. M. J.; Groendijk, H.; Read, R. J.; Greenhough, T. J.; Campbell, J. W.; Kalk, K. H.; Littlechild, J. A.; Watson, H. C.; Hol, W. G. J. Structure of Glycosomal Glyceraldehyde-3-phosphate Dehydrogenase from *Trypanosoma brucei* Determined from Laue Data. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 2355–2359.
- (17) Kim, H.; Feil, I. K.; Verlinde, C. L. M. J.; Petra, P. H.; Hol, W. G. Crystal Structure of Glycosomal Glyceraldehyde-3-Phosphate Dehydrogenase from *Leishmania mexicana*: Implications for Structure-Based Drug Design and a New Position for the Inorganic Phosphate Binding Site. *Biochemistry* **1995**, *34*, 14975–14986.
- (18) Jenkins, J. L.; Tanner, J. J. High-Resolution Structure of Human D-Glyceraldehyde-3-Phosphate Dehydrogenase. *Acta Crystallogr., Sect. D* **2006**, *62*, 290–301.
- (19) Ismail, S. A.; Park, H. W. Structural Analysis of Human Liver Glyceraldehyde-3-Phosphate Dehydrogenase. *Acta Crystallogr., Sect. D* **2005**, *61*, 1508–1513.
- (20) Guido, R. V. C.; Oliva, G.; Andricopulo, A. D. Virtual Screening and Its Integration with Modern Drug Design Technologies. *Curr. Med. Chem.* **2008**, *15*, 37–46.
- (21) Castilho, M. S.; Guido, R. V. C.; Andricopulo, A. D. 2D Quantitative Structure-Activity Relationship Studies on a Series of Cholesteryl Ester Transfer Protein Inhibitors. *Bioorg. Med. Chem.* **2007**, *15*, 6242–6252.
- (22) Salum, L. B.; Polikarpov, I.; Andricopulo, A. D. Structural and Chemical Basis for Enhanced Affinity and Potency for a Large Series of Estrogen Receptor Ligands: 2D and 3D QSAR Studies. *J. Mol. Graphics Modell.* **2007**, *26*, 434–442.
- (23) Honorio, K. M.; Garratt, R. C.; Polikarpov, I.; Andricopulo, A. D. 3D QSAR Comparative Molecular Field Analysis on Nonsteroidal Farnesoid X Receptor Activators. *J. Mol. Graphics Modell.* **2007**, *25*, 921–927.
- (24) Farutin, V.; Masterson, L.; Andricopulo, A. D.; Cheng, J.; Riley, R.; Hakimi, R.; Frazer, J.; Cordes, E. H. Structure-Activity Relationships for a Class of Inhibitors of Purine Nucleoside Phosphorylase. *J. Med. Chem.* **1999**, *42*, 2422–2431.
- (25) Honorio, K. M.; Garratt, R. C.; Andricopulo, A. D. Hologram Quantitative Structure-Activity Relationships for a Series of Farnesoid X Receptor Activators. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3119–3125.
- (26) Guido, R. V. C.; Trossini, G. H. G.; Castilho, M. S.; Oliva, G.; Ferreira, E. I.; Andricopulo, A. D. Structure-Activity Relationships for a Class of Selective Inhibitors of the Major Cysteine Protease from *Trypanosoma cruzi*. *J. Enzyme Inhib. Med. Chem.* **2008**. DOI: 10.1080/14756360701810322.
- (27) Castilho, M. S.; Postigo, M. P.; de Paula, C. B.; Montanari, C. A.; Oliva, G.; Andricopulo, A. D. Two- and Three-Dimensional Quantitative Structure-Activity Relationships for a Series of Purine Nucleoside Phosphorylase Inhibitors. *Bioorg. Med. Chem.* **2006**, *14*, 516–527.
- (28) Aronov, A. M.; Suresh, S.; Buckner, F. S.; Van Voorhis, W. C.; Verlinde, C. L. M. J.; Opperdoes, F. R.; Hol, W. G.; Gelb, M. H. Structure-Based Design of Submicromolar, Biologically Active Inhibitors of Trypanosomatid Glyceraldehyde-3-Phosphate Dehydrogenase. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 4273–4278.
- (29) Van Calenbergh, S.; Verlinde, C. L. M. J.; Soenens, J.; De Bruyn, A.; Callens, M.; Blaton, N. M.; Peeters, O. M.; Rozenskik, J.; Hol, W. G.; Herdewijn, P. Synthesis and Structure-Activity Relationships of Analogs of 2'-deoxy-2'-(3-methoxybenzamido)adenosine, a Selective Inhibitor of Trypanosomal Glycosomal Glyceraldehyde-3-Phosphate Dehydrogenase. *J. Med. Chem.* **1995**, *38*, 3838–3849.
- (30) Aronov, A. M.; Verlinde, C. L. M. J.; Hol, W. G.; Gelb, M. H. Selective Tight Binding Inhibitors of Trypanosomal Glyceraldehyde-3-Phosphate Dehydrogenase Via Structure-Based Drug Design. *J. Med. Chem.* **1998**, *41*, 4790–4799.

- (31) Verlinde, C. L. M. J.; Callens, M.; Van Calenbergh, S.; Van Aerschot, A.; Herdewijn, P.; Hannaert, V.; Michels, P. A. M.; Opperdoes, F. R.; Hol, W. G. Selective Inhibition of Trypanosomal Glyceraldehyde-3-Phosphate Dehydrogenase by Protein Structure-Based Design: Toward New Drugs for the Treatment of Sleeping Sickness. *J. Med. Chem.* **1994**, *37*, 3605–3613.
- (32) Bressi, J. C.; Verlinde, C. L. M. J.; Aronov, A. M.; Shaw, M. L.; Shin, S. S.; Nguyen, L. N.; Suresh, S.; Buckner, F. S.; Van Voorhis, W. C.; Kuntz, I. D.; Hol, W. G.; Gelb, M. H. Adenosine Analogues as Selective Inhibitors of Glyceraldehyde-3-Phosphate Dehydrogenase of Trypanosomatidae Via Structure-Based Drug Design. *J. Med. Chem.* **2001**, *44*, 2080–2093.
- (33) Bakker, B. M.; Michels, P. A. M.; Opperdoes, F. R.; Westerhoff, H. V. What Controls Glycolysis in Bloodstream form *Trypanosoma brucei*. *J. Biol. Chem.* **1999**, *274*, 14551–14559.
- (34) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.
- (35) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. D. Development and Validation of a Genetic Algorithm for Flexible Docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (36) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved Protein-Ligand Docking Using GOLD. *Proteins* **2003**, *52*, 609–623.
- (37) Suresh, S.; Bressi, J. C.; Kennedy, K. J.; Verlinde, C. L. M. J.; Gelb, M. H.; Hol, W. G. Conformational Changes in *Leishmania mexicana* Glyceraldehyde-3-Phosphate Dehydrogenase Induced by Designed Inhibitors. *J. Mol. Biol.* **2001**, *309*, 423–435.
- (38) Bissantz, C.; Folkers, G.; Rognan, D. Protein-Based Virtual Screening of Chemical Databases. 1. Evaluation of Different Docking/Scoring Combinations. *J. Med. Chem.* **2000**, *43*, 4759–4767.
- (39) Wang, R.; Lu, Y.; Wang, S. Comparative Evaluation of 11 Scoring Functions for Molecular Docking. *J. Med. Chem.* **2003**, *46*, 2287–2303.
- (40) Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A Critical Assessment of Docking Programs and Scoring Functions. *J. Med. Chem.* **2006**, *49*, 5912–5931.
- (41) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5867.
- (42) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Recent Advances in Comparative Molecular Field Analysis (CoMFA). *Prog. Clin. Biol. Res.* **1989**, *291*, 161–165.
- (43) Klebe, G.; Abraham, U.; Mietzner, T. Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules to Correlate and Predict Their Biological Activity. *J. Med. Chem.* **1994**, *37*, 4130–4146.
- (44) Klebe, G.; Abraham, U. Comparative Molecular Similarity Index Analysis (CoMSIA) to Study Hydrogen-Bonding Properties and to Score Combinatorial Libraries. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 1–10.
- (45) Doddareddy, M. R.; Jung, H. K.; Cha, J. H.; Cho, Y. S.; Koh, H. Y.; Chang, M. H.; Pae, A. N. 3D QSAR Studies on T-type Calcium Channel Blockers Using CoMFA and CoMSIA. *Bioorg. Med. Chem.* **2004**, *12*, 1613–1621.
- (46) Amin, E. A.; Welsh, W. J. Highly Predictive CoMFA and CoMSIA Models for Two Series of Stromelysin-1 (MMP-3) Inhibitors Elucidate S1' and S1-S2' Binding Modes. *J. Chem. Inf. Model.* **2006**, *46*, 1775–1783.
- (47) Zhu, L. P.; Ye, D. Y.; Tang, Y. Structure-Based 3D-QSAR Studies on Thiazoles as 5-HT₃ Receptor Antagonists. *J. Mol. Model.* **2007**, *13*, 121–131.
- (48) Oprea, T. I.; Waller, C. L. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley: New York, 1997; Vol. 11, Chapter 3, pp. 127–182.
- (49) Golbraikh, A.; Tropsha, A. Beware of q²! *J. Mol. Graphics Model.* **2002**, *20*, 269–276.
- (50) Golbraikh, A.; Shen, M.; Xiao, Z.; Xiao, Y. D.; Lee, K. H.; Tropsha, A. Rational Selection of Training and Test Sets for the Development of Validated QSAR Models. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 241–253.
- (51) Doweyko, A. M. 3D-QSAR Illusions. *J. Comput.-Aided Mol. Des.* **2004**, *18*, 587–596.
- (52) Clark, M.; Cramer, R. D., III; Van Opdenbosch, N. Validation of the General Purpose Tripos 5.2 Force Field. *J. Comput. Chem.* **1989**, *10*, 982–1012.
- (53) Powell, M. J. D. Restart Procedures for the Conjugate Gradient Method. *Math. Program.* **1977**, *12*, 241–254.
- (54) Gasteiger, J.; Marsili, M. Iterative Partial Equalization of Orbital Electronegativity - A Rapid Access to Atomic Charges. *Tetrahedron* **1980**, *36*, 3219–3228.
- (55) Bush, B. L.; Nachbar, R. B., Jr. Sample-Distance Partial Least Squares: PLS Optimized For Many Variables, with Application to CoMFA. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 587.
- (56) Clark, R. D.; Fox, P. C. Statistical variation in progressive scrambling. *J. Comput.-Aided Mol. Des.* **2004**, *17*, 1–14.
- (57) Cramer, R. D.; Bunce, J. D.; Patterson, D. E.; Frank, I. E. Cross-Validation, Bootstrapping, and Partial Least-Squares Compared with Multiple-Regression in Conventional QSAR Studies. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18–25.
- (58) Wehrens, R.; van der Linden, W. E. Bootstrapping Principal Component Regression Models. *J. Chemom.* **1997**, *11*, 157–171.

CI700453J