3D-QSAR CoMFA of a Series of DABO Derivatives as HIV-1 Reverse Transcriptase Non-Nucleoside Inhibitors

Monique Araújo de Brito,*[,†] Carlos Rangel Rodrigues,[‡] José Jair Vianna Cirino,[†,#]
Ricardo Bicca de Alencastro,[†] Helena Carla Castro,[§] and Magaly Girão Albuquerque*[,†]

Programa de Pós-Graduação em Química Orgânica, Departamento de Química Orgânica, Laboratório de Modelagem Molecular (LabMMol), Instituto de Química, CCMN, UFRJ, Rio de Janeiro, RJ, Brazil, CEP 21941-909, Departamento de Medicamentos, Laboratório de Modelagem Molecular e QSAR (ModMolQSAR), Programa de Pós-Graduação em Ciências Farmacêuticas, Faculdade de Farmácia, CCS, UFRJ, Rio de Janeiro, RJ, Brazil, CEP 21941-590, and Departamento de Biologia Celular e Molecular, Laboratório de Antibióticos, Bioquímica e Modelagem Molecular (LABioMol), Instituto de Biologia, CEG, UFF, Niterói, RJ, Brazil, CEP 24210-130

A series of 74 dihydroalkoxybenzyloxopyrimidines (DABOs), a class of highly potent non-nucleoside reverse transcriptase inhibitors (NNRTIs), was retrieved from the literature and studied by comparative molecular field analysis (CoMFA) in order to derive three-dimensional quantitative structure−activity relationship (3D-QSAR) models. The CoMFA study has been performed with a training set of 59 compounds, testing three alignments and four charge schemes (DFT, HF, AM1, and PM3) and using defaults probe atom (Csp^3, +1 charge), cutoffs (30 kcal.mol^{-1} for both steric and electrostatic fields), and grid distance (2.0 Å). The best model ($N = 59$), derived from Alignment 1 and PM3 charges, shows $q^2 = 0.691$, $SE_{cv} = 0.475$, optimum number of components $= 6$, $r^2 = 0.930$, SEE $= 0.226$, and F-value $= 115.544$. The steric and electrostatic contributions for the best model were 43.2% and 56.8%, respectively. The external predictive ability ($r^2_{pred} = 0.918$) of the resultant best model was evaluated using a test set of 15 compounds. In order to design more potent DABO analogues as anti-HIV/AIDS agents, attention should be taken in order to select a substituent for the 4-oxopyrimidine ring, since, as revealed by the best CoMFA model, there are a steric restriction at the C2-position, a electron-rich group restriction at the C6-position (*para*-substituent of the 6-benzyl group), and a steric allowed region at the C5-position.

INTRODUCTION

The human immunodeficiency virus type-1 (HIV-1) is a retrovirus that infects cells of the human immune system causing the globally disseminated disease named acquired immunodeficiency syndrome (AIDS). HIV-1 reverse transcriptase (RT) is a key enzyme responsible for the conversion of the viral single-stranded RNA genome into double-stranded DNA. HIV-1 RT exhibits two enzymatically distinct activities: DNA polymerase activity, synthesizing DNA from either RNA or DNA templates, and ribonuclease-H (RNaseH) activity that specifically degrades the RNA strand in RNA/DNA hybrids.[1,2] HIV-1 RT is only active as a p66/p51 heterodimer. The p66 subunit (560 aa, 66 kDa) contains both DNA polymerase and RNaseH domains at distinct regions,[1,2] while the p51 subunit (440 aa, 51 kDa) is derived from the p66 by proteolytic cleavage but keeping a no functional DNA polymerase domain.[3,4]

RT is one of the most studied targets into the HIV-1 replicative cycle in order to develop new drug candidates against the HIV/AIDS infection. Drug discovery research targeting HIV-1 RT have been extensively enhanced in the past two decades.[5] Currently approved drugs targeting HIV-1 RT to treat HIV/AIDS infection fall into two classes: nucleoside/nucleotide analog RT inhibitors (NRTIs/NtRTIs) and non-nucleoside RT inhibitors (NNRTIs).

NRTIs/NtRTIs are competitive inhibitors (of the natural deoxynucleotide substrates) that function as chain terminators in the DNA elongation process, thus preventing HIV from multiplying. Zidovudine (ZDV, azidothymidine or AZT), a NRTI, was the first antiretroviral drug approved by the U.S. Food and Drug Administration (FDA)[6,7] for the treatment of HIV/AIDS patients in 1986. Several others NRTIs are currently approved by the FDA:[7] didanosine (dideoxyinosine or ddI, 1991), zalcitabine (dideoxycytidine or ddC, 1992), stavudine (d4T, 1994), lamivudine (3TC, 1995), abacavir (ABC, 1998), emtricitabine (FTC, 2003), and tenofovir (a NtRTI, 2001). However, their application is limited due to the high cytotoxicity (inhibition to other normal polymerization process in vivo) and the emergence of drug-resistant mutants.[8]

On the other hand, NNRTIs, a chemically diverse set of compounds, are noncompetitive substrate inhibitors that bind at an allosteric site (non-nucleoside binding site, NNBS) and also block the process of DNA synthesis, with high specificity to HIV-1 RT. NNRTIs have gained a definitive and important place as anti-HIV/AIDS drugs due to their unique

* Corresponding author e-mail: moniquebrito@yahoo.com.br (M.A.B.); girao@acd.ufrj.br (M.G.A.).
† Instituto de Química.
‡ Faculdade de Farmácia.
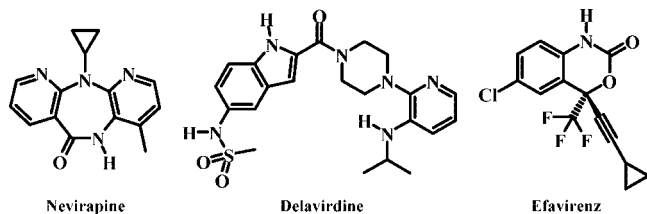# Current address: Petrobras/CENPES.
§ Instituto de Biologia.

3D-QSAR CoMFA of a Series of DABO Derivatives

J. Chem. Inf. Model., Vol. 48, No. 8, 2008 **1707**

Figure 1. Structures of nevirapine, delavirdine, and efavirenz: NNRTI drugs approved by the FDA to treat HIV/AIDS infection.

Figure 2. Structures of some NNRTIs of the HEPT (MKC-442 or emivirine) and DABO (NH- and S-DABO) series.

antiviral profile, being highly potent, selective, and with low toxicity.[5] Nevirapine (NVP, 1996), delavirdine (DLV, 1997), and efavirenz (EFV, 1998) are the three NNRTIs currently approved by the FDA[7] (Figure 1). However, similar to NRTIs, NNRTI's Achilles heel is the emergence of drug-resistant mutants, and there is no final treatment solution to AIDS yet.[4,8,9] The current treatment for HIV/AIDS infection is a regimen known as highly active antiretroviral therapy (HAART), which consists of the combinations of three or four drugs belonging to at least two classes of antiretroviral agents, such as NRTIs/NtRTIs, NNRTIs, and protease inhibitors (PIs).[4,6,10]

Comparative molecular field analysis (CoMFA),[11] a well known three-dimensional quantitative structure−activity relationship (3D-QSAR) method, has been established in many studies as a versatile and powerful tool in rational drug design and related applications.[12−19] The CoMFA approach involves the alignment of molecules in a structurally and pharmacologically reasonable manner on the basis of the assumption that each compound acts via a common binding mode.[11,13,16,20] Therefore, in order to develop 3D-QSAR CoMFA models, we should have some knowledge or hypothesis regarding active conformations of the molecules under study as a prerequisite for structural alignment. CoMFA, based on the assumption that enzyme−inhibitor interactions are primarily noncovalent (i.e., considering reversible enzyme inhibitors) and shape-dependent, derives a 3D-QSAR model by systematically sampling the steric and electrostatic interactions by means of the Lennard-Jones and Coulomb potentials, respectively, between a series of aligned molecules and an array of probe atoms arranged in a grid.[21]

A 3D-QSAR model is a mathematical expression that relates the variation of the biological response in a series of related compounds to the variation in their three-dimensional chemical structure.[16,20,22,23] The relation between the spatial interactions (independent variables) and the biological response (dependent variable) is established by use of the partial least-squares (PLS) regression method.[20,23,24] By careful selection of the biological data set (the training and test set compounds) and careful model construction (e.g., the trial alignment and the putative active conformation), a 3D-QSAR study can lead to a predictive model that could be used to predict the biological activity values of new compounds prior to their synthesis, which is the primary goal of any drug design process in the medicinal chemistry field.[23,25−29] Since several cocrystal structures of NNRTI bound to HIV-1 RT have been determined,[30,1,4,5] such as for MKC-442 or emivirine (6-benzyl-1-(ethoxymethyl)-5-isopropyluracil) (Figure 2), a NNRTI of the hydroxyethoxyphenylthiothymine (HEPT) series, which is available in the Protein Data Bank (PDB)[31] under the PDB code 1RT1,[32]
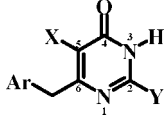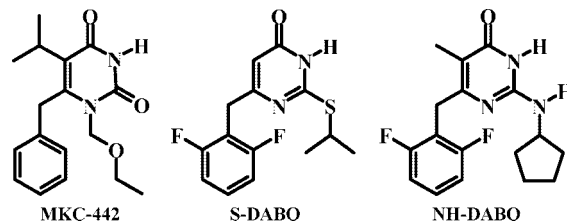
this information can be applied in various molecular modeling studies to develop new and more potent compounds.

The dihydroalkoxybenzyloxopyrimidine (DABO, Figure 2) series is a class of highly potent NNRTIs, structurally related to HEPT derivatives.[1,4] DABOs were first disclosed in 1992 and further developed during the following years, being characterized by the presence of a pyrimidine nucleobase (uracil or thymine) containing the 6-benzyl group and a 2-O-alkyl, 2-S-alkyl, or 2-NH-alkyl chain.[33] S-DABOs (S-alkyl, Figure 2) showed increased anti-HIV-1 activity comparing with original DABOs, and recently, isosteric replacement from -S- to -NH- moiety furnished NH-DABOs (NH-alkyl, Figure 2), with increased potency.[34] Since, DABOs synthesis, pharmacology, and medicinal chemistry is still of interest,[35−37] we are proposing in this work the first 3D-QSAR CoMFA model for this class of compounds.

Therefore, on the basis of the 3D-QSAR CoMFA model for a series of DABO analogues, we attempted to elucidate a structure−activity relationship to provide useful guidelines for the design of new DABO derivatives as potent inhibitors of HIV-1 RT, based on putative bioactive conformation and pharmacophore hypothesis. Because of the lack of structural data supporting a specific active conformation for the DABO class, we might assume that a compound is active in an energetically minimized conformation.[20] In addition, since the cocrystal structure of MKC-442 bound to HIV-1 RT is available in the PDB, this served as a template upon which other analogues were overlaid. Therefore, we are assuming that the MKC-442 and the DABO analogues share the same binding mode. Moreover, the combination of X-ray investigations on the complexes with CoMFA studies allows a much more accurate interpretation of the CoMFA contour maps and leads also to a better knowledge of the enzyme-ligand interaction.

MATERIALS AND METHODS

Biological Data. We selected 74 DABO derivatives from the literature[38−41] for the CoMFA study. The compounds biological profile was evaluated in HIV-1 RT affinity binding assays, using ^{3}[H]-dGTP as the radioligand. Compounds IC_{50} values were transformed in pIC_{50} (-$LogIC_{50}$), which ranged from 4.23 to 7.52 M. A test set of 15 compounds (**60**-**74**, Table 1), used to determine the external predictivity of the resulting CoMFA model, was removed from the original data set that evenly spanned the antiviral activity range as well as the chemical structural diversity of the database. The remaining 59 compounds (**1**-**59**, Table 1) constitute the training set used to derive the CoMFA model. Compounds containing a stereogenic center (Y=S-*sec*-butyl, Table 1), and thus corresponding to a racemate, were defined with the *R* absolute configuration, and the original IC_{50} values were

Table 1. Structures of the S- and NH-DABO Derivatives and the Corresponding HIV-1 RT Inhibitory Potencies (pIC_{50}, M)[38–41]

no.[a]	X[b]	Ar	Y[c]	pIC_{50}[d]	no.[a]	X[b]	Ar	Y[c]	pIC_{50}[d]
1	Me	2-naphtyl	S-*sec*-Bu	4.23	**38**	H	2,6-di-F-Ph	S-Me	6.10
2	H	1-naphtyl	S-cyclopentyl	4.31	**39**	Me	2-Cl-Ph	S-*sec*-Bu	6.10
3	Me	1-naphtyl	S-cyclopentyl	4.35	**40**	Me	2-F-Ph	S-*sec*-Bu	6.10
4	Me	4-F-Ph	S-*sec*-Bu	4.59	**41**	Me	3-NO_2−Ph	S-*sec*-Bu	6.10
5	Me	4-Cl-Ph	S-*sec*-Bu	4.77	**42**	H	2-F-Ph	S-*sec*-Bu	6.22
6	H	1-naphtyl	S-*sec*-Bu	4.79	**43**	H	3-NO_2−Ph	S-*sec*-Bu	6.22
7	H	2-naphtyl	S-*sec*-Bu	4.83	**44**	H	2,6-di-Cl-Ph	S-*tert*-Bu	6.22
8	H	4-F-Ph	S-*sec*-Bu	4.83	**45**	H	2,6-di-Cl-Ph	S-*n*-Bu	6.30
9	H	4-Cl-Ph	S-*sec*-Bu	5.02	**46**	H	2,6-di-Cl-Ph	S-cyclopentyl	6.40
10	H	Ph	S-*tert*-Bu	5.07	**47**	H	2,6-di-F-Ph	S-*n*-Bu	6.70
11	H	3-Me-Ph	S-*tert*-Bu	5.09	**48**	H	2,6-di-F-Ph	S-*tert*-Bu	6.70
12	Me	3-Me-Ph	S-*sec*-Bu	5.27	**49**	H	2,6-di-Cl-Ph	S-*sec*-Bu	6.70
13	Me	2,6-di-Cl-Ph	S-cyclohexyl	5.31	**50**	Me	2,6-di-Cl-Ph	S-*sec*-Bu	6.92
14	Me	Ph	S-Me	5.31	**51**	H	2,6-di-F-Ph	S-*sec*-Bu	7.00
15	Me	Ph	S-*sec*-Bu	5.32	**52**	Me	2,6-di-F-Ph	S-*sec*-Bu	7.00
16	Me	3-Me-Ph	S-*tert*-Bu	5.34	**53**	H	2,6-di-F-Ph	S-cyclohexyl	7.05
17	Me	Ph	S-cyclohexyl	5.37	**54**	Me	2,6-di-F-Ph	S-*tert*-Bu	7.05
18	H	3-Cl-Ph	S-*sec*-Bu	5.42	**55**	H	2,6-di-F-Ph	S-cyclopentyl	7.10
19	Me	4-NO_2−Ph	S-*sec*-Bu	5.44	**56**	Me	2,6-di-F-Ph	S-cyclopentyl	7.10
20	Me	3-Me-Ph	S-cyclopentyl	5.47	**57**	H	2,6-di-F-Ph	NH-cyclopentyl	7.15
21	H	2-Cl-Ph	S-*sec*-Bu	5.49	**58**	H	2,6-di-F-Ph	S-*iso*-Pr	7.30
22	Me	3-F-Ph	S-*sec*-Bu	5.52	**59**	Me	2,6-di-F-Ph	NH-cyclopentyl	7.52
23	H	2,6-di-Cl-Ph	S-Me	5.52	**60**	Me	1-naphtyl	S-*sec*-Bu	4.35
24	H	Ph	S-cyclohexyl	5.52	**61**	H	2-naphtyl	S-cyclohexyl	4.48
25	H	3-Me-Ph	S-*iso*-Pr	5.54	**62**	H	Ph	S-*sec*-Bu	5.27
26	H	Ph	S-cyclopentyl	5.55	**63**	Me	Ph	S-cyclopentyl	5.47
27	H	3-Me-Ph	S-cyclohexyl	5.59	**64**	H	3-Me-Ph	S-cyclopentyl	5.59
28	Me	3-Me-Ph	S-Me	5.60	**65**	Me	Ph	S-*iso*-Pr	5.60
29	Me	3-Me-Ph	S-*iso*-Pr	5.60	**66**	H	3-Me-Ph	S-*sec*-Bu	5.62
30	H	4-NO_2−Ph	S-*sec*-Bu	5.62	**67**	Me	3-Cl-Ph	S-*sec*-Bu	5.74
31	Me	3-Me-Ph	S-cyclohexyl	5.66	**68**	H	3-F-Ph	S-*sec*-Bu	5.92
32	Me	Ph	S-*tert*-Bu	5.72	**69**	H	2-NO_2−Ph	S-*sec*-Bu	6.22
33	Me	2,6-di-Cl-Ph	S-cyclopentyl	5.80	**70**	H	2,6-di-Cl-Ph	S-cyclohexyl	6.40
34	H	2,6-di-Cl-Ph	S-*iso*-Pr	5.89	**71**	Me	2,6-di-F-Ph	S-Me	6.70
35	Me	2,6-di-Cl-Ph	S-*iso*-Pr	5.94	**72**	Me	2,6-di-F-Ph	S-*n*-Bu	7.05
36	Me	2,6-di-Cl-Ph	S-*n*-Bu	5.94	**73**	Me	2,6-di-F-Ph	S-cyclohexyl	7.15
37	Me	2,6-di-Cl-Ph	S-*tert*-Bu	5.96	**74**	Me	2,6-di-F-Ph	S-*iso*-Pr	7.30

[a] Underlined numbers correspond to test set compounds (**60**-**74**). [b] DABOs pyrimidine nucleobase uracil (X=H) or thymine (X=Me). [c] S-DABO (Y=S-alkyl) and NH-DABO (Y=NH-alkyl) series. [d] The original IC_{50} values of compounds containing a stereogenic center (Y=S-*sec*-butyl) were multiplied by two.

multiplied by two. Since this stereocenter is located at the lateral alkyl chain, and considering it of minor importance, we have arbitrarily considered the *R*-enantiomer as the eutomer. Table 1 shows the chemical structures and biological activity values (pIC_{50}) of the 74 compounds. Figure S1 (Supporting Information) shows the potency range of the training set (pIC_{50} from 4.23 to 7.52 M) and test set (pIC_{50} from 4.35 to 7.30 M).

Computational Methods. *3D-Structures Construction and Optimization.* The structures of the entire set of DABO derivatives (Table 1) were built according to the structure of the template compound, MKC-442 (Figure 2), available in the PDB under code 1RT1[32] using the Build option in Spartan′06 (v. 1.0.1).[42] Figure 3 shows the superposition of MKC-442 and the most potent NH-DABO derivative, compound **59**. All structures were fully geometry optimized at the AM1 semiempirical level of theory available in the Spartan program.

Partial Atomic Charge Assignment. In a CoMFA approach, the electrostatic contribution to the final 3D-QSAR model depends on the scheme used for the partial atomic charge assignment, due the charge variation according to the theoretical approach used to derived them. Therefore, in order to assess the effect of partial atomic charge assignment scheme on the overall quality of CoMFA models, we have assigned the inhibitors charges using the restrained molecular electrostatic potential (RESP) fitting procedure. The charges were fitted to a RESP calculated at four different levels of theory: (i) DFT (B3LYP/6−31G*), (ii) *ab initio* (HF/6−31G*), and (iii) AM1 and (iv) PM3 semiempirical methods available in the Spartan program. The optimized data set compounds were imported into the Sybyl 7.0,[43] in which the subsequent CoMFA calculations were performed.

CoMFA Methodology, Pharmacophore Hypothesis and Alignment. The positioning of the molecules inside the fixed lattice is by far the most important aspect in CoMFA, since the interaction energies depend strongly on the relative position of the functional groups present in the ligands.[11,44] Based on the binding mode of MKC-442 and the superposition of this reference compound with the derivative **59**
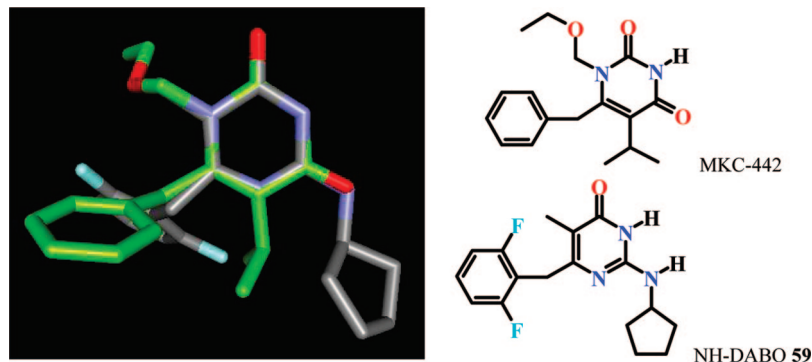
Figure 3. Superposition of MKC-442 (carbon atoms colored in green) and NH-DABO **59** (carbon atoms colored in gray) 3D-structures.
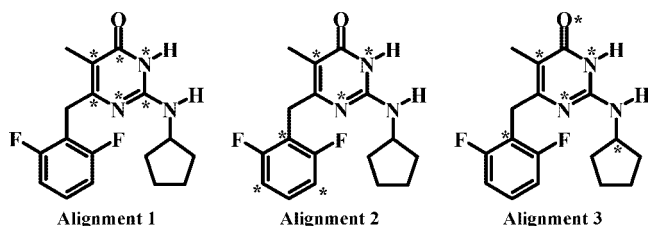

Figure 4. Template compound (NH-DABO **59**) marked with atoms used for superposition in the three trial alignments.

(Figure 3), we derived a pharmacophore hypothesis to orient the superposition of all DABO derivatives and to afford a consistent alignment.

Therefore, in order to obtain the best superposition of the DABO compounds, we have tested three alignments for the subsequent QSAR analysis, using as template compound **59**. As shown in Figure 4, Alignment 1 embraces the six atoms (N1, C2, N3, C4, C5, and C6) of the 4-oxopyrimidine ring, Alignment 2 comprises only three atoms (N1, N3, and C5) of the 4-oxopyrimidine ring plus three atoms (C1′, C3′, and C5′) of the 6-benzyl group, and Alignment 3 consists of the same three atoms of the 4-oxopyrimidine ring used in Alignment 2 plus the oxygen atom of the 4-carbonyl group, the carbon atom (C1′) of the *ipso* position of the 6-benzyl group, and the carbon atom (C1) of the cyclopentyl group. Substituents on the 6-benzyl group were oriented in the same relative direction for optimal overlap with each another. Compounds were aligned via the Alignment Database command in Sybyl 7.0.[43]

Grid Distance and 3D-Field Calculations. Following standard CoMFA procedures, each compound of the training set was inserted into a virtual three-dimensional box, which was extended at least 4.0 Å in all directions from the set of aligned molecules, according to the three trial alignments, and the default grid distance (D=2.0 Å). To explore the effect of different variables on the predictive ability of the CoMFA model, separate CoMFA models were constructed that varied in terms of their alignment scheme, partial-charge assignment schemes, cutoffs levels, and probe atom types.

Regression Analysis. Linear regression equations were obtained by PLS analysis, correlating changes in the independent variables (i.e., steric and electrostatic fields) with changes in the dependent-variable (pIC_{50}) of the inhibitor sets. PLS reduces the dimensionality of the problem by constructing latent variables that are composed of linear combinations of the original independent variables. Leave-one-out cross-validation was used to assess the internal predictive ability of the CoMFA model. In this technique,

compounds are systematically excluded from the data set, and the pIC_{50} value of each removed compound is predicted by a new model (same equation with new coefficients) derived from the remaining compounds in the training set. Cross-validation yielded the optimum number of principal components (PC) together with the highest cross-validated q^2 value.

The PLS analyses were then repeated without cross-validation using the optimum number of PC, giving the final CoMFA model from which the conventional r^2 values, noncross-validated standard errors of estimate (SSE), F-value (Fisher's test value), and related statistical parameters were computed. As recommended by ref 45, the optimum number of PC was chosen as that which corresponds to the first minimum in the standard error of cross-validation (SE_{cv}). To avoid chance correlations, no more than N/5 components were extracted.[21]

Column filtering, for any column of computed energies with a variation less than 2.0 kcal.mol^{-1}, was applied as needed to reduce computation time without negatively affecting the quality of the models. The CoMFA model was represented as a color contour map depicting regions of descriptor fields that contribute significantly to that model.

The results from cross-validation analysis were expressed as the cross-validated r^2 value (r^2_{cv} or q^2), which is defined according to eqs 1 and 2 (where Y is the observed activity value).

$$r^2_{cv} = 1 - \frac{PRESS}{\sum (Y - Y_{mean})^2} \qquad (1)$$

$$PRESS = \sum (Y - Y_{pred})^2 \qquad (2)$$

Predictive Ability. To validate the derived CoMFA model, the overall predictive ability of that analysis was evaluated by the term r^2_{pred}, which was calculated according to eq 3.

$$r^2_{pred} = \frac{SD - PRESS}{SD} \qquad (3)$$

In eq 3, SD is the sum of squared deviation between the biological activities of the test set molecule and the mean activity value of the training set molecules, and PRESS is the prediction error sum of squares derived from the leave-one-out method. The uncertainty of prediction is defined according to eq 4 (where k is the number of variables in the model and n is the number of compounds used in the study).

1710 *J. Chem. Inf. Model., Vol. 48, No. 8, 2008*

BRITO ET AL.

$$\text{SPRESS} = \left[\frac{\text{PRESS}}{n - k - 1}\right]^{1/2} \qquad (4)$$

RESULTS AND DISCUSSIONS

3D-QSAR CoMFA models were generated using a series of 74 dihydroalkoxybenzyloxopyrimidine (DABO) non-nucleoside HIV-1 RT inhibitors from the literature.[38–41] The corresponding structures and biological activity values of the data set are described in Table 1. The *in vitro* RT inhibitory activity (pIC_{50}, M) was used as the dependent variable and the CoMFA fields, as the independent variables.

3D-QSAR models were generated from a training set of 59 compounds and externally validated by predicting the potencies of a test set of 15 compounds. The selection of test set and training set compounds was done to include moderate and high potent compounds occurring in roughly equal proportions in both series (see Figure S1, Supporting Information).

Selection of bioactive conformations and their alignments are the two most crucial steps in CoMFA. Not only do they often significantly influence the results but also they are critical in the design of new molecules. The alignments tested showed good overlap between different regions of the molecules. As an example, Figure 5 depicts the structural Alignment 1 of S-DABO and NH-DABO derivatives.

All the cross-validated results were analyzed by considering the fact that the value of q^2 above 0.3 indicates that the probability of chance correlation is less than 5%.[44] Statistical results of CoMFA models obtained, testing the three alignments and the four partial atomic charge assignment schemes (DFT, HF, AM1, and PM3) but keeping defaults cutoff (30 kcal.mol^{-1} for both steric and electrostatic fields), probe atom (Csp^3, +1 charge), and grid spacing (2.0 Å), are presented in Table 2.

It can be seen in Table 2 that there is not a significant difference among statistical values of CoMFA models generated by the four schemes of partial atomic charge assignment (DFT, HF, AM1 and PM3), using the three trial alignments. In fact, they show qualitatively similar CoMFA maps (data not shown), with positive and negative contributions distributed in similar regions around the molecules.

However, the best model was obtained with Alignment 1 and PM3 derived charges, using defaults cutoff, probe atom,
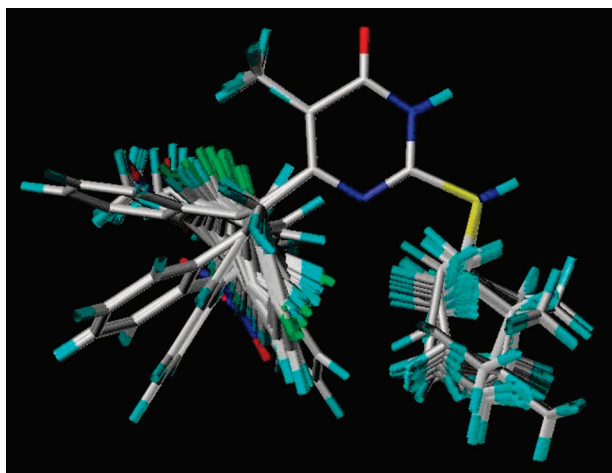

Figure 5. Superposition of the 59 compounds from the training set of S- and NH-DABOs according to Alignment 1.

Table 2. CoMFA Statistical Results of the Three Trial Alignments for S- and NH-DABOs, Using Four Partial Atomic Charge Assignment Schemes, and Defaults Cutoff (30 kcal.mol^{-1} for Both Steric and Electrostatic Fields), Probe Atom (Csp^3, +1 Charge), and Grid Spacing (2.0 Å)

q[a]	q^{2b}	SE$_{cv}$[c]	PC[d]	r^{2e}	SEE[f]	F-value[g]	% S[h]	% E[i]
Alignment 1								
DFT	0.611	0.533	6	0.927	0.230	110.437	50.2	49.8
HF	0.617	0.528	6	0.942	0.205	141.498	41.3	58.7
AM1	0.674	0.488	6	0.931	0.225	116.223	42.7	57.3
PM3	0.691	0.475	6	0.930	0.226	115.544	43.2	56.8
Alignment 2								
DFT	0.575	0.546	4	0.875	0.296	94.897	50.5	49.5
HF	0.614	0.530	6	0.940	0.210	135.118	42.5	57.5
AM1	0.649	0.501	5	0.911	0.252	109.094	40.2	59.8
PM3	0.681	0.478	5	0.920	0.239	121.982	41.2	58.8
Alignment 3								
DFT	0.607	0.535	6	0.923	0.238	103.164	50.0	50.0
HF	0.613	0.531	6	0.940	0.209	136.485	43.8	56.2
AM1	0.659	0.499	6	0.927	0.230	110.485	43.5	56.5
PM3	0.683	0.481	6	0.929	0.228	112.782	43.6	56.4

[a] The partial atomic charges (q) were fitted to a restrained molecular electrostatic potential (RESP) calculated at four different levels of theory: DFT (B3LYP/6−31G*), ab initio (HF/6−31G*), and (AM1 and PM3) semiempirical methods. [b] Cross-validated r^2 (q^2). [c] Standard error of cross-validation (SE$_{cv}$). [d] Optimum number of principal components (PC). [e] Coefficient of determination (r^2). [f] Standard errors of estimate (SSE). [g] Fisher's test value (F-value). [h] Steric contribution (% S). [i] Electrostatic contribution (% E).

and grid spacing. The stepwise F-test performed on the PRESS values justified six components, with SE$_{cv}$ = 0.475 and q^2 = 0.691. As expected, the corresponding non cross-validated analysis yields a better data fitting (SEE = 0.226 and r^2 = 0.930). The % ratio of electrostatic and steric fields contribution in the final model was 56.8% and 43.2%, respectively, indicating a nearly equivalent influence of these two fields on ligand−receptor interactions. However, there is a slight predominance of the electrostatic contribution, which is interesting, since the RT non-nucleoside binding site (NNBS) is characterized as a hydrophobic pocket,[30] where steric features should be predominating. Probably, this result is reflecting a subtle influence of the electrostatic character due to the almost equivalent overall molecular volume (and molecular shape) of this DABO series, which is ranging from 4.23 to 7.52 M in the training set.

We also investigated the effect of different probe-atoms on the resulting q^2 of the model, with compounds assigned with PM3 charges (the best ones) and considering the default cutoff and grid spacing values. Table 3 shows statistical results of CoMFA models derived for the three probe atoms (Csp^3, +1 charge; Osp^3, −1 charge; and H, +1 charge). Although similar with the others, the best q^2, 0.695, was obtained for the sp^3 oxygen with charge −1; however, there was a great decrease in r^2 and Fisher's values when comparing with the Csp^3 with +1 charge, showing a lower level of internal consistency and significance. Therefore, Csp^3 with +1 charge remains the best choice for this analysis.

Using the smoothing function to apply the cutoffs levels, we assessed the effect of different cutoffs on the resulting q^2 of the best model, with PM3 charges, that were truncated to both steric and electrostatic contributions in 30 (default), 20, and 10 kcal.mol^{-1}. The best statistical results were obtained with the default cutoff (Table 4).

Table 3. CoMFA Statistical Results According to Alignment 1 for S- and NH-DABOs, Using Three Probe Atoms (Csp³, +1 Charge; Osp³, −1 Charge; and H, +1 Charge), and Defaults Cutoff (30 kcal.mol⁻¹ for Both Steric and Electrostatic Fields) and Grid Spacing (2.0 Å)

PA^a	q^{2b}	SE$_{cv}$^{c}	PC^d	r^e	SEE^f	F-value^g	% S^h	% E^i
Csp³ +1	0.691	0.475	6	0.930	0.226	115.544	43.2	56.8
Osp³ −1	0.695	0.462	4	0.842	0.333	71.939	43.7	56.3
H +1	0.678	0.476	4	0.862	0.312	84.058	0	100

^{a} Probe atom (PA) and net charge. ^{b} Cross-validated r² (q²). ^{c} Standard error of cross-validation (SE$_{cv}$). ^{d} Optimum number of principal components (PC). ^{e} Coefficient of determination (r²). ^{f} Standard errors of estimate (SSE). ^{g} Fisher's test value (F-value). ^{h} Steric contribution (% S). ^{i} Electrostatic contribution (% E).

Table 4. CoMFA Statistical Results According to Alignment 1 for S- and NH-DABOs, Using Three Cutoff Values (30, 20, and 10 kcal/mol⁻¹) for Both Steric and Electrostatic Fields, and Defaults Probe Atom (Csp³, +1 Charge) and Grid Spacing (2.0 Å)

cutoff^a	q^{2b}	SE$_{cv}$^{c}	PC^d	r^{2e}	SEE^f	F-value^g	% S^h	% E^i
30	0.691	0.475	6	0.930	0.226	115.544	43.2	56.8
20	0.685	0.478	6	0.921	0.240	100.727	46.0	54.0
10	0.726	0.447	6	0.924	0.235	105.724	47.2	52.8

^{a} Cutoff values for both steric and electrostatic fields. ^{b} Cross-validated r² (q²). ^{c} Standard error of cross-validation (SE$_{cv}$). ^{d} Optimum number of principal components (PC). ^{e} Coefficient of determination (r²). ^{f} Standard errors of estimate (SSE). ^{g} Fisher's test value (F-value). ^{h} Steric contribution (% S). ^{i} Electrostatic contribution (% E).

The significance and utility of QSAR models is generally checked by predicting the activity values of a set of compounds, named test set, which are not included in model development. The 15 molecules from the test set (**60**-**74**) were constructed and minimized as described for the training set ones. They were aligned with the most potent derivative (**59**) using atom-based rms fitting. The potency of the test set was predicted with the best CoMFA model and exhibited a high external predictive ability (r²$_{pred}$=0.918). Considering the external predictivity of compounds in the test set, the performance of the model can be termed as great. Approximately 80% of compounds potencies were very well predicted with residual values lower than 0.50. Statistically significant predictions support the validity of derived models in predicting the potencies of newer derivatives.

Table 5 shows predicted potencies (pIC$_{50Pred}$) for the training and test set compounds, according to the best CoMFA model, as well as the observed pIC$_{50}$ (pIC$_{50Obs}$) and the respective calculated residues (pIC$_{50Obs}$ - pIC$_{50Pred}$). Considering outliers those compounds that had a residue larger than twice the standard deviation of the model (residue > |0.452|), there are two outliers in the training set, compounds **46** (residue = 0.48) and **50** (residue = 0.72) and three outliers in the test set, compounds **68** (residue = 0.51), **69** (residue = 0.54), and **70** (residue = 0.62) (Table 5). Figures 6 and 7 show a plot with predicted and observed pIC$_{50}$ values for the training and test sets, respectively.

According to statistical results and test set predictive ability of the best model, we chose to evaluate graphically CoMFA maps (i.e., steric and electrostatic contour plots) generated with Alignment 1, PM3 derived charges, and defaults cutoff (30 kcal.mol⁻¹), probe atom (Csp³ with +1 charge), and grid spacing (2.0 Å). The contour plots are to be considered as a

Table 5. Observed and Predicted pIC$_{50}$ Values and Residues (pIC$_{50Obs}$ - pIC$_{50Pred}$) for the Training (**1**-**59**) and Test (**60**-**74**) Sets According to the Best 3D-QSAR CoMFA Model According to Alignment 1 for S- and NH-DABOs, Using PM3 Derived Charges, and Defaults Cutoff (30 kcal.mol⁻¹ for Both Steric and Electrostatic Fields), Probe Atom (Csp3, +1 Charge), and Grid Spacing (2.0 Å)

no.	pIC$_{50Obs}$	pIC$_{50Pred}$	res	no.	pIC$_{50Obs}$	pIC$_{50Pred}$	res
1	4.23	4.23	0.00	38	6.10	6.26	−0.16
2	4.31	4.54	−0.23	39	6.10	5.93	0.17
3	4.35	4.23	0.12	40	6.10	6.06	0.04
4	4.59	4.64	−0.05	41	6.10	5.89	0.21
5	4.77	5.02	−0.25	42	6.22	6.07	0.15
6	4.79	4.64	0.15	43	6.22	6.32	−0.10
7	4.83	4.90	−0.07	44	6.22	6.08	0.14
8	4.83	4.58	0.25	45	6.30	6.15	0.15
9	5.02	4.93	0.09	46	6.40	5.92	0.48
10	5.07	5.19	−0.12	47	6.70	6.73	−0.03
11	5.09	5.45	−0.36	48	6.70	6.83	−0.13
12	5.27	5.53	−0.26	49	6.70	6.27	0.43
13	5.31	5.67	−0.36	50	6.92	6.20	0.72
14	5.31	5.15	0.16	51	7.00	7.01	−0.01
15	5.32	5.53	−0.21	52	7.00	7.17	−0.17
16	5.34	5.47	−0.13	53	7.05	7.06	−0.01
17	5.37	5.42	−0.05	54	7.05	6.89	0.16
18	5.42	5.41	0.01	55	7.10	7.05	0.05
19	5.44	5.47	−0.03	56	7.10	7.13	−0.03
20	5.47	5.46	0.01	57	7.15	7.26	−0.11
21	5.49	5.85	−0.36	58	7.30	7.11	0.19
22	5.52	5.55	−0.03	59	7.52	7.44	0.08
23	5.52	5.87	−0.35	**60**	4.35	4.40	−0.05
24	5.52	5.55	−0.03	**61**	4.48	4.73	−0.25
25	5.54	5.47	0.07	**62**	5.27	5.43	−0.16
26	5.55	5.56	−0.01	**63**	5.47	5.44	0.03
27	5.59	5.49	0.10	**64**	5.59	5.54	0.05
28	5.60	5.35	0.25	**65**	5.60	5.42	0.18
29	5.60	5.47	0.13	**66**	5.62	5.39	0.23
30	5.62	5.69	−0.07	**67**	5.74	5.65	0.09
31	5.66	5.53	0.13	**68**	5.92	5.41	0.51
32	5.72	5.39	0.33	**69**	6.22	5.68	0.54
33	5.80	5.83	−0.03	**70**	6.40	5.78	0.62
34	5.89	6.25	−0.36	**71**	6.70	6.78	−0.08
35	5.94	6.19	−0.25	**72**	7.05	6.95	0.10
36	5.94	6.19	−0.25	**73**	7.15	7.15	0.00
37	5.96	6.10	−0.14	**74**	7.30	7.24	0.06
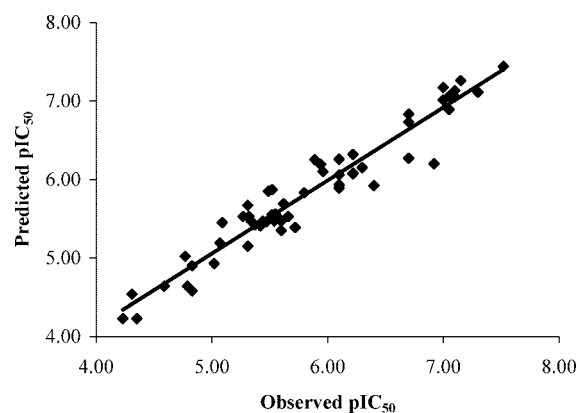
Figure 6. Observed vs predicted pIC$_{50}$ for the training set (**1**-**59**) according to the best 3D-QSAR CoMFA model for S- and NH-DABOs, using Alignment 1, PM3 derived charges, and defaults cutoff (30 kcal.mol⁻¹ for both steric and electrostatic fields), probe atom (Csp³, +1 charge), and grid spacing (2.0 Å).

representation of the lattice points, where the difference in field values is strongly associated with difference in receptor binding affinity. Though CoMFA contour maps cannot be used as receptor maps, they still generate many useful interpretations. In order to illustrate the best CoMFA model,
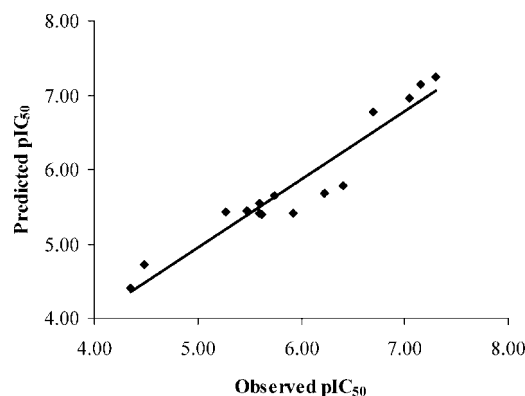
Figure 7. Observed vs predicted pIC_{50} for the test set (**60**-**74**) according to the best 3D-QSAR CoMFA model, using Alignment 1, PM3 derived charges, and defaults cutoff (30 kcal.mol^{-1} for both steric and electrostatic fields), probe atom (Csp3, +1 charge), and grid spacing (2.0 Å).

Figure 8 shows the steric (Figure 8-A) and electrostatic (Figure 8-B) contour maps around the most potent compound, NH-DABO **59** (residue = 0.08).

In the CoMFA steric field contour map (Figure 8-A), areas where steric bulk substituents increase the potency are represented by green polyhedrons, while areas where steric bulk substituents decrease the potency are represented by yellow polyhedrons. Using derivative **59** as a reference compound, we can note steric fields around the substituents at the C2, C5, and C6 positions of the 4-oxopyrimidine ring (Figure 8-A). With respect to the C2-position, a steric map shows a green region around this position, but close to it there are also yellow fragmental fields, indicating careful substituent group selection for this region.

Concerning the C5-position, there is a sterically favorable region (green) around it, which would increase the biological potency, making this region suitable for future modifications. However, the size of this group should not be too large because of the limited volume of the corresponding binding pocket.

Finally, with respect to the C6-position of the 4-oxopyrimidine ring, CoMFA steric map shows a large yellow region around the aromatic ring, indicating a steric restriction at this position. The size of the aromatic ring at this region should not be too large. Indeed, compounds that have a naphthalene ring at this position, instead of a phenyl ring, have the worst activity profile of the database (pIC_{50} values ranging from 4.23 to 4.83 M).
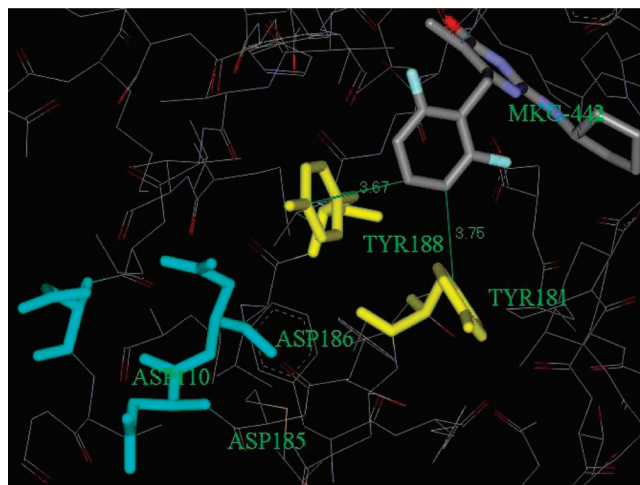

Figure 9. Stereoview of the binding mode of NH-DABO **59** (stick model and colored by element) in the HIV-1 RT non-nucleoside binding site. All residues are represented in lines and colored by element, except Tyr181 and Tyr188 (in close contact with the benzyl group of **59**), which are represented in stick and colored in yellow, and Asp110, Asp185, and Asp186 (catalytic residues), which are represented in stick and colored in blue. Hydrogens are omitted for clarity.

In addition, analyzing S- and NH-DABO compounds docked in the RT non-nucleoside binding site (NNBS), we see that better substituents in this region should be actually aromatic groups due the observed pi-pi interactions among this aromatic group and the aromatic rings of Tyr181 and Tyr188. In fact, the most potent compound, **59**, docked in the NNBS showed, after minimization, a distance of 3.75 Å and 3.67 Å between its aromatic ring and the aromatic rings of Tyr181 and Tyr188, respectively (Figure 9). These amino acids are pivotal in the inhibitory process, since the RT catalytic site is contiguous to the NNBS (allosteric site), where Tyr181 and Tyr188 are in close contact with the catalytic residues (Asp110, Asp185, and Asp186), there is a reorganization of the catalytic site after inhibitor binding to the allosteric site, impairing the catalytic function.[1,2,6]

In the CoMFA electrostatic field contour map (Figure 8-B), areas where electronegative groups increase the potency are represented in red, while areas where electronegative groups decrease the potency are represented in blue. Using derivative **59** as reference, we can note blue (close to *ortho* and *para* positions) and red (close to the *meta* position) electrostatic fields around the substituents at the C6-position of the
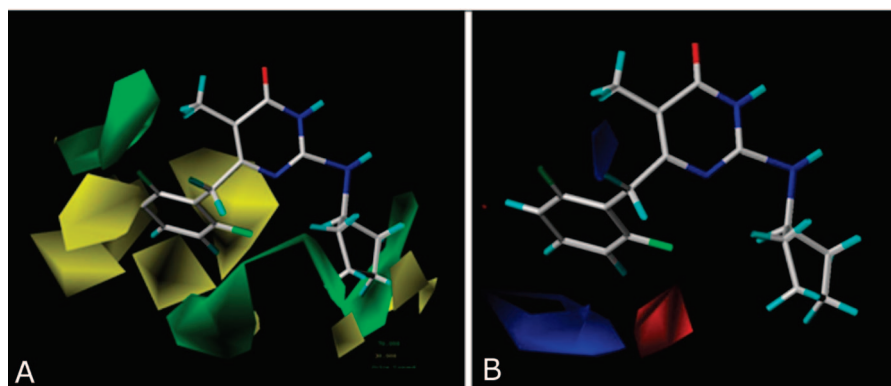

Figure 8. The CoMFA contour maps for compound NH-DABO **59** (stick model and colored by element). (A) Steric map indicating areas where bulk is predicted to increase (green) or decrease (yellow) potency. (B) Electrostatic map indicating where high electron density (negative charge) (red) and low electron density (positive charge) (blue) regions are expected to increase potency.
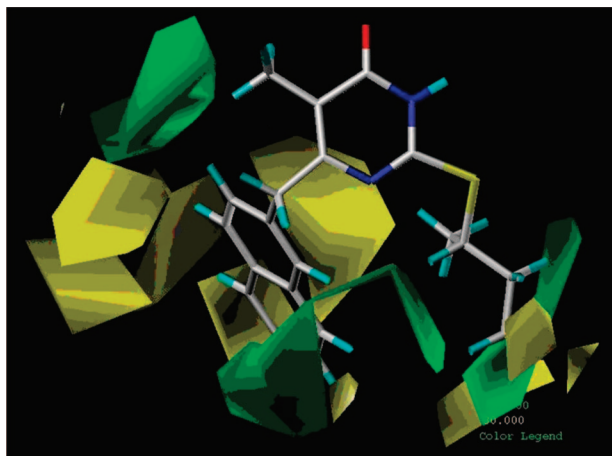
3D-QSAR CoMFA OF A SERIES OF DABO DERIVATIVES

J. Chem. Inf. Model., Vol. 48, No. 8, 2008 **1713**

Figure 10. The CoMFA steric contour map for compound S-DABO **1** (stick model and colored by element), indicating areas where bulk is predicted to increase (green) or decrease (yellow) potency.

4-oxopyrimidine ring (Figure 8-B), indicating careful substituent group selection.

The most potent compounds from the training and test set, derivatives **59** (pIC_{50Obs} = 7.52, residue = 0.08) and **74** (pIC_{50Obs} = 7.30, residue = 0.06), respectively, were predicted with great accuracy using the best CoMFA model. In addition, the least potent ones from the training and test set, derivatives **1** (pIC_{50Obs} = 4.23, residue = 0) and **60** (pIC_{50Obs} = 4.35, residue = −0.05), respectively, were also predicted with great accuracy. Both compounds have low inhibitory profile, probably due to the presence of a bulky naphthalene ring (C6-position of the 4-oxopyrimidine ring), positioned in a volume restricted subpocket of the NNBS, as is shown in the CoMFA steric map (yellow areas around the naphthyl ring) of compound **1** (Figure 10).

Concerning the outliers from the training set, compound **50** (pIC_{50Obs} = 6.92 M) is the worst predicted by the CoMFA model (pIC_{50Pred} = 6.20 M), showing the largest residual value (residue = 0.72) from both training and test sets. By observing this outlier compound inside the CoMFA electrostatic model, as shown in Figure 11, we can see a blue region near one chlorine atom of the 2,6-dichlorophenyl group, indicating that negatively charged groups around it decrease the potency. This proximity, not observed for fluorine analog **52** (residue = −0.17), implicated in a poor potency prediction, is probably due to the larger size of the chlorine group compared with fluorine. Similar behavior occurred with outlier **46** (residue = 0.48), which could be compared with fluorine analog **55** (residue = 0.05) and also with compound **49** (residue = 0.43), compared with fluorine analog **51** (residue = −0.01).

Concerning the outliers from the test set, derivatives **70** (residue = 0.62), **69** (residue = 0.54) and **68** (residue = 0.51) were the worst predicted. Comparing compound **70** (Ar=2,6-dichlorophenyl) with the fluorine analog **73** (Ar=2,6-difluorphenyl, residue = 0), we can explain its behavior by the same reason given before for compound **50**. However, for compounds **69** (Ar=2-nitro-phenyl) and **68** (Ar=3-fluorphenyl), both of them containing an asymmetric substitution pattern at the C6-position of the 4-oxopyrimidine ring, it is not clear the reason for their outlier behavior, since there are other compounds *ortho*- or *meta*-monosubstituted, but that are still well predicted.
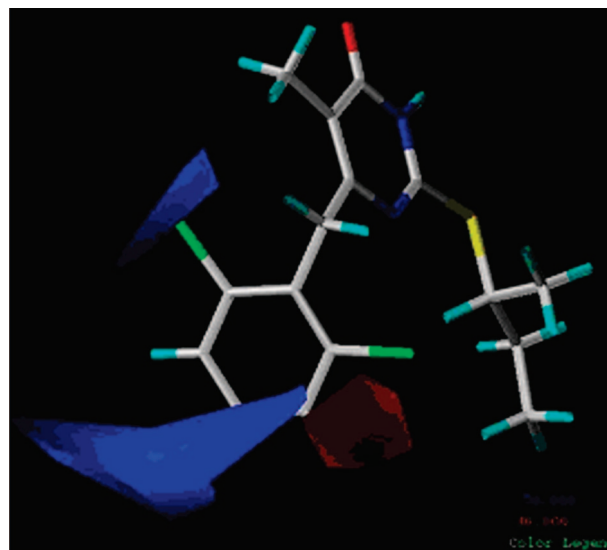

Figure 11. The CoMFA electrostatic contour map for compound S-DABO **50** (stick model and colored by element), indicating areas where high electron density (negative charge) (red) and low electron density (positive charge) (blue) regions are expected to increase potency.

CONCLUSIONS

In the present study, we have performed a 3D-QSAR analysis of 74 structurally related S- and NH-DABO derivatives from the literature by the CoMFA method. It provided good models with predictive ability and self-consistency for the design of new potent DABO analogues as anti-HIV/AIDS agents. In general, all models, developed with a training set of 54 compounds, correlated well with experimental results.

Based on average residuals, the best CoMFA model obtained with PM3 derived charges, according to Alignment 1 and using defaults cutoff (30 kcal.mol^{-1} for both steric and electrostatic fields), probe atom (Csp^3, +1 charge), and grid spacing (2.0 Å), has slightly higher predictive ability (q^2=0.691) than other charges schemes (DFT, HF, and AM1), cutoff values (20 and 10 kcal.mol^{-1}), and probe atom types (Osp^3, −1 charge; and H, +1 charge). The best CoMFA model exhibited good internal consistency in terms of r^2 (0.930) and SEE (0.226) and good predictive ability in terms of q^2 (0.691) and SE_{cv} (0.475). The steric and electrostatic contributions for the best model were 43.2% and 56.8%, respectively. In addition, 15 compounds (test set) were used to evaluate the external predictive ability (r^2_{pred} = 0.918) of the best model.

The most relevant structural conclusions of this study are the following: (i) the steric restriction at the C2-position of the 4-oxopyrimidine ring, which implies in a careful substituent selection; (ii) the restriction of electron-rich groups as a *para*-substituent of the phenyl ring at the C6-position of the 4-oxopyrimidine ring, and (iii) the allowed presence of more bulky substituents at the C5-position of the 4-oxopyrimidine ring, that would increase biological potency, making this region suitable for future modifications.

Overall, the proposed CoMFA 3D-QSAR model provides the possibility to achieve reliable predictions on the inhibitory potency of DABO derivatives and should be useful for assisting the design of new structurally related compounds, shedding light on important ligand−receptor binding char-

1714 *J. Chem. Inf. Model., Vol. 48, No. 8, 2008*

BRITO ET AL.

acteristic, in order to develop potential HIV-1 reverse transcriptase non-nucleoside inhibitors for HIV/AIDS drug therapy.

ACKNOWLEDGMENT

Supporting Information Available: Correlation between activities (pIC_{50}) of the training (blue) and test (red) sets of compounds (Figure S1). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) De Clercq, E. New Approaches Toward Anti-HIV Chemotherapy. *J. Med. Chem.* **2005**, *48*, 1297–1313.

(2) Castro, H. C.; Loureiro, N. I. V.; Pujol-Luz, M.; Souza, A. M. T.; Albuquerque, M. G.; Santos, D. O.; Cabral, L. M.; Frugulhetti, I. C.; Rodrigues, C. R. HIV-1 Reverse Transcriptase: A Therapeutical Target in the Spotlight. *Curr. Med. Chem.* **2006**, *13*, 313–324.

(3) Balzarini, J. Current Status of the Non-Nucleoside Reverse Transcriptase Inhibitors of Human Immunodeficiency Virus Type 1. *Curr. Top. Med. Chem.* **2004**, *4*, 921–944.

(4) De Clercq, E. HIV-Chemotherapy and -Prophylaxis: New Drugs, Leads and Approaches. *Int. J. Biochem. Cell Biol.* **2004**, *36*, 1800–1822.

(5) De Clercq, E. New Developments in Anti-HIV Chemotherapy. *B.B.A. Mol. Basis. Dis.* **2002**, *1587*, 258–275.

(6) Flexner, C. Antiretroviral Agents and Treatment of HIV Infection In *Goodman & Gilman: The Pharmacological Basis of Therapeutics,* 12th ed.;Brunton, L. L., Ed.; McGraw Hill: New York, NY, 2006; Section VIII, Chapter 9, pp 1273−1314.

(7) United States Food and Drug Administration. FDA Approved HIV AIDS Drugs. http://www.fda.gov/oashi/aids/virals.html (accessed May 20, 2008).

(8) De Clercq, E. New Developments in Anti-HIV Chemotherapy. *Il Farmaco* **2001**, *56*, 3–12.

(9) De Clercq, E. Perspectives of Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs) in the Therapy of HIV Infection. *Il Farmaco* **1999**, *54*, 26–4.

(10) WHO/UNAIDS Global Summary of the HIV/AIDS Epidemic. WHO Report. http://www.who.int/mediacentre/news/releases/2007/pr61/en/index.html (accessed May 20, 2008).

(11) Cramer, R. D. III; 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*, 5959–5967.

(12) Jayatilleke, P. R. N.; Nair, A. C.; Zauhar, R.; Welsh, W. J. Computational Studies on HIV-1 Protease Inhibitors: Influence of Calculated Inhibitor-Enzyme Binding Affinities on the Statistical Quality of 3D-QSAR CoMFA Models. *J. Med. Chem.* **2000**, *43*, 4446–4451.

(13) Sippl, W.; Holtje, H. D. Structure-Based 3D-QSAR - Merging the Accuracy of Structure-Based Alignments with the Computational Efficiency of Ligand-Based Methods. *J. Mol. Struct. (Theochem)* **2000**, *503*, 31–50.

(14) Hannongbua, S.; Nivasanond, K.; Lawtrakul, L.; Pungpo, P.; Wolschann, P. 3D-Quantitative Structure-Activity Relationships of HEPT Derivatives as HIV- 1 Reverse Transcriptase Inhibitors, Based on Ab Initio Calculations. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 848–855.

(15) Avery, M. A.; Alvim-Gaston, M.; Rodrigues, C. R.; Barreiro, E. J.; Cohen, F. E.; Sabnis, Y. A.; Woolfrey, J. R. Structure-Activity Relationships of the Antimalarial Agent Artemisinin 6. The Development of Predictive in vitro Potency Models Using CoMFA and HQSAR Methodologies. *J. Med. Chem.* **2002**, *45*, 292–303.

(16) Akamatsu, M. Current State and Perspectives of 3D-QSAR. *Curr. Top. Med. Chem.* **2002**, *12*, 1381–1394.

(17) Aboye, T. L.; Sobhia, M. E.; Bharatam, P. V. 3D-QSAR Studies of Pyruvate Dehydrogenase Kinase Inhibitors Based on a Divide and Conquer Strategy. *Bioorg. Med. Chem.* **2004**, *12*, 2709–2715.

(18) Bhongade, B. A.; Gadad, A. K. 3D-QSAR CoMFA/CoMSIA Studies on Urokinase Plasminogen Activator (UPA) Inhibitors: A Strategic Design in Novel Anticancer Agents. *Bioorg. Med. Chem.* **2004**, *12*, 2797–2805.

(19) 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.

(20) (a) Kubinyi, H. QSAR and 3D-QSAR in Drug Design. Part 1: Methodology. *Drug Discovery Today* **1997**, *11*, 457–467. (b) Kubinyi, H. QSAR and 3D-QSAR in Drug Design. Part 2: Applications and Problems. *Drug Discovery Today* **1997**, *12*, 538–546.

(21) Melville, J. L.; Hirst, J. D. On the Stability of CoMFA Models. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1294–1300.

(22) Gaudio, A. C.; Zandonade, E. Proposition, Validation and Analysis of QSAR Models. *Quím. Nova* **2001**, *24*, 658–671.

(23) Leach, A. R. *Molecular Modelling. Principles and Applications*, 2nd ed.; Pearson Education: Harlow, England, 2001; 744 pp.

(24) (a) Geladi, P.; Kowalski, B. R. Partial Least-Squares Regression: A Tutorial. *Anal. Chim. Acta* **1986**, *186*, l–17. (b) Geladi, P.; Kowalski, B. R. An Example of 2-Block Predictive Partial Least Squares Regression with Simulated Data. *Anal. Chim. Acta* **1986**, *186*, 19–32.

(25) Amaral, A. T.; Montanari, C. A. 25 Years of Medicinal Chemistry in Brazil. *Quim. Nova* **2002**, *25*, 39–44, Suppl. 1.

(26) Oprea, T. I.; Marshall, G. R. Receptor-Based Prediction of Binding Affinities In *3D QSAR in Drug Design*; Kubinyi, H., Folkers, G., Martin, Y. C., Eds.; Kluwer Academic Publishers: London, 2002; Vol 2, pp 35−61.

(27) Thomas, G. The SAR and QSAR Approaches to Drug Design In *Fundamentals of Medicinal Chemistry*; John Wiley & Sons: England, 2003; pp 71−92.

(28) Selassie, C. D. History of Quantitative Structure-Activity Relationships In *Burger's Medicinal Chemistry and Drug Discovery,* 10th ed.; Abraham, D. J., Ed.; Wiley-Interscience: NJ, 2003; Vol *1*, pp 1−48.

(29) Tropsha, A. Recent Trends in Quantitative Structure-Activity Relationships. In *Burger's Medicinal Chemistry and Drug Discovery,* 10th ed.; Abraham, D. J., Ed.; Wiley-Interscience: NJ, 2003; Vol *1*, pp 49−76.

(30) Ren, J. S.; Esnouf, R.; Garman, E.; Somers, D.; Ross, C.; Kirby, I.; Keeling, J.; Darby, G.; Jones, Y.; Stuart, D.; Stammers, D. High Resolution Structures of HIV-1 RT from Four RT-Inhibitor Complexes. *Nat. Struct. Biol.* **1995**, *2*, 293–302.

(31) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(32) Hopkins, A. L.; Ren, J.; Esnouf, R. M.; Willcox, B. E.; Jones, E. Y.; Ross, C.; Miyasaka, T.; Walker, R. T.; Tanaka, H.; Stammers, D. K.; Stuart, D. I. Complexes of HIV-1 Reverse Transcriptase with Inhibitors of the HEPT Series Reveal Conformational Changes Relevant to the Design of Potent Non-Nucleoside Inhibitors. *J. Med. Chem.* **1996**, *39*, 1589–1600.

(33) Artico, M.; Massa, S.; Mai, A.; Marongiu, M. E.; Piras, G.; Tramontano, E.; La Colla, P. 3,4-Dihydro-2-Alkoxy-6-Benzyl-4-Oxopyrimidines (DABOs): A New Class of Specific Inhibitors of Human Immunodeficiency Virus Type 1. *Antiviral Chem. Chemother.* **1993**, *4*, 361–368.

(34) D'Cruz, O. J.; Uckun, F. M. Novel Tight Binding PETT, HEPT and DABO-Based Non-Nucleoside Inhibitors of HIV-1 Reverse Transcriptase. *J. Enzyme Inhib. Med. Chem.* **2006**, *21*, 329–350.

(35) Wang, Y.; Chen, F.-E.; Balzarini, J.; De Clercq, E.; Pannecouque, C. Non-Nucleoside HIV-1 Reverse-Transcriptase Inhibitors. Part 10. Synthesis and Anti-HIV Activity of 5-Alkyl-6-(1-Naphthylmethyl)Pyrimidin-4(3H)-Ones with a Mono- or Disubstituted 2-Amino Function as Novel Dihydro-Alkoxy-Benzyl-Oxopyrimidine (DABO) Analogues. *Chem. Biodivers.* **2008**, *5*, 168–176.

(36) Mugnaini, C.; Alongi, M.; Togninelli, A.; Gevariya, H.; Brizzi, A.; Manetti, F.; Bernardini, C.; Angeli, L.; Tafi, A.; Bellucci, L.; Corelli, F.; Massa, S.; Maga, G.; Samuele, A.; Facchini, M.; Clotet-Codina, I.; Armand-Ugon, M.; Este, J. A.; Botta, M. Dihydro-Alkylthio-Benzyl-Oxopyrimidines as Inhibitors of Reverse Transcriptase: Synthesis and Rationalization of the Biological Data on Both Wild-Type Enzyme and Relevant Clinical Mutants. *J. Med. Chem.* **2007**, *50*, 6580–6595.

(37) Aly, Y. L.; Pedersen, E. B.; La Colla, P.; Loddo, R. Novel Synthesis and Anti-HIV-1 Activity of 2-Arylthio-6-Benzyl-2,3-Dihydro-1*H*-Pyrimidin-4-Ones (Aryl S-DABOs). *Synthesis* **2007**, *13*, 1955–1960.

(38) Mai, A.; Artico, M.; Sbardella, G.; Massa, S.; Loi, A. G.; Tramontano, E.; Scano, P.; La Colla, P. Synthesis and Anti-HIV-1 Activity of Thio Analogues of Dihydro-Alkoxy-Benzyl-Oxopyrimidines. *J. Med. Chem.* **1995**, *38*, 3258–3263.

(39) Mai, A.; Artico, M.; Sbardella, G.; Quartarone, S.; Massa, S.; Loi, A. G.; De Montis, A.; Scintu, F.; Putzolu, M.; La Colla, P. Dihydro-(Alkylthio)-(Naphthylmetil)-Oxopyrimidines: Novel Non-Nucleoside Reverse Transcriptase Inhibitors of the S-DABO Series. *J. Med. Chem.* **1997**, *40*, 1447–1454.

(40) Mai, A.; Artico, M.; Sbardella, G.; Massa, S.; Novellino, E.; Greco, G.; Loi, A. G.; Tramontano, E.; Marongiu, M. E.; La Colla, P. 5-Alkyl-

2-(Alkylthio)-6-(2,6-Dihalophenylmethyl)-3,4-Dihydropyrimidin-4(3*H*)-Ones: Novel Potent and Selective Dihydro-Alkoxy-Benzyl-Oxopyrimidine Derivatives. *J. Med. Chem.* **1999**, *42*, 619–627.

(41) Ragno, R.; Mai, A.; Sbardella, G.; Artico, M.; Massa, S.; Musiu, C. Computer-Aided Design, Synthesis, and Anti-HIV-1 Activity in vitro of 2-Alkylamino-6-[1-(2,6-difluorophenyl)alkyl]-3,4-dihydro-5-alkylpyrimidin-4(3H)-ones as Novel Potent Non-Nucleoside Reverse Transcriptase Inhibitors, Also Active Against the Y181C Variant. *J. Med. Chem.* **2004**, *47*, 928–934.

(42) Spartan 06; Wavefunction Inc.: Irvine, CA 92612, U.S.A., 2004.

(43) Sybyl 7.2; Tripos Associates Inc.: 1699 South Hanley Road, St. Louis, MO 63144.

(44) Peterson, S. D.; Schaal, W.; Karlén, A. Improved CoMFA Modeling by Optimization of Settings. *J. Chem. Inf. Model.* **2006**, *46*, 355–364.

(45) Kubinyi, H.; Abraham, U. Practical Problems in PLS Analyses. In *3D-QSAR in Drug Design. Theory, Methods and Applications*; Kubinyi, H., Ed.; ESCOM: Leiden, The Netherlands, 1993; pp 717−728.