European Journal of Medicinal Chemistry 44 (2009) 4344–4352

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Original article

3D-QSAR CoMFA/CoMSIA models based on theoretical active conformers of HOE/BAY-793 analogs derived from HIV-1 protease inhibitor complexes

Elaine Fontes Ferreira da Cunha [a,b,c,*], Wolfgang Sippl [b], Teodorico de Castro Ramalho [a], Octavio Augusto Ceva Antunes [c,†], Ricardo Bicca de Alencastro [c], Magaly Girão Albuquerque [c,**]

[a] Universidade Federal de Lavras (UFLA), Departamento de Química, 37200-000, Caixa Postal 3037, Lavras, MG, Brazil
[b] Martin-Luther-Universität Halle-Wittenberg (MLU), Department of Pharmaceutical Chemistry, Wolfgang-Langenbeck-Str. 4, 06120 Halle (Saale), Germany
[c] Universidade Federal do Rio de Janeiro (UFRJ), Instituto de Química (IQ), Programa de Pós-Graduação em Química (PGQu), Ilha do Fundão, 21941-909 Rio de Janeiro, RJ, Brazil

ARTICLE INFO

Article history:
Received 18 April 2008
Received in revised form
27 April 2009
Accepted 15 May 2009
Available online 23 May 2009

Keywords:
3D-QSAR CoMFA/CoMSIA
HIV-1 protease inhibitor HOE/BAY-793
Aspartyl-protease
Peptidomimetic
C2-Symmetric diol
Molecular modeling

ABSTRACT

The three-dimensional quantitative structure–activity relationships (3D-QSAR) of a series of HOE/BAY-793 analogs (C_2-symmetric diol peptidomimetics), developed by Budt and co-workers [Bioorg. Med. Chem. 3 (1995) 559] as inhibitors of HIV-1 protease (HIV-PR), were studied using Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA). Theoretical active conformers for these peptidomimetics were generated, derived from modeled protease inhibitor complexes, in order to orient the compounds superposition and to afford a consistent alignment. The best CoMFA model ($N = 27$, $q^2 = 0.637$, $R^2 = 0.991$) showed contributions of the steric (45.7%) and electrostatic (54.3%) fields to the activity, while the best CoMSIA model ($N = 27$, $q^2 = 0.511$, $R^2 = 0.987$) showed contributions of the electrostatic (68.5%) and hydrogen bond donor (37.5%) fields. The models were also external validated using four compounds (test set) not included in the model generation process. The statistical parameters from both models indicate that the data are well fitted and have high predictive ability. Moreover, the resulting 3D CoMFA/CoMSIA contour maps provide useful guidance for designing highly active ligands. The CoMFA/CoMSIA models were also compared with previous 4D-QSAR models [E.F.F. da Cunha, M.G. Albuquerque, O.A.C. Antunes, R.B. de Alencastro, QSAR Comb. Sci. 24 (2005), 240–253.].

1. Introduction

The HIV-1 protease (HIV-PR) performs an essential step in the life cycle of the human immunodeficiency virus, the causative agent of the acquired immunodeficiency syndrome (AIDS), catalyzing the cleavage of the gag and gag-pol precursors into mature proteins. The functional form of the HIV-PR is a homodimer, containing 99 amino acid residues in each monomer unit; one related to the other by a two-fold crystallographic axis [1]. There is a unique active site for substrate cleavage, which is located at the interface between the adjacent subunits. Since the HIV-PR belongs to the aspartyl-protease family, it involves a pair of catalytic triads (Asp-Thr-Gly), where the two catalytic aspartyl side-chains (one from each monomer) are closely opposed and form a symmetric dyad. The binding site has an extended ravine shape with an entry controlled by two flaps that form a flexible gate for the approaching ligand (substrate or inhibitor) [2,3].

Because of its essential function, HIV-PR is an excellent target for anti-HIV drug therapy. There are nine HIV-1 protease inhibitors (PIs) currently approved by the U.S. Food and Drug Administration (FDA): saquinavir (the first PI approved in 1995), ritonavir (1996), indinavir (1996), nelfinavir (1997), amprenavir (1999), lopinavir (2002), atazanavir (2003), and more recently, tipranavir and darunavir (2005). These compounds share the same structural scaffold, i.e., a hydroxyethylene moiety (or a hydroxyethene group in the case of tipranavir) instead of a natural peptide bond (amide). These peptidomimetics represent non-scissile substrate analogs for the HIV-PR. The effectiveness of these PIs has been hampered by the emergence of drug-resistant and cross-resistant mutants. Therefore, a definitive cure for AIDS has yet to be found [4], and the

* Corresponding author. Tel./fax: +55 35 3829 1271.
** Corresponding author. Tel./fax: +55 21 2562 7132.
E-mail addresses: elaine_cunha@ufla.br (E.F.F. da Cunha), wolfgang.sippl@pharmazie.uni-halle.de (W. Sippl), teo@ufla.br (T.C. Ramalho), bicca@iq.ufrj.br (R.B. de Alencastro), magaly@iq.ufrj.br (M.G. Albuquerque).
† Deceased (June 1st, 2009)

0223-5234/$ – see front matter © 2009 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2009.05.016

development of a new generation of PIs, that effectively addresses the issue of resistance, requires a better understanding of the interactions between HIV-PR and its ligands (substrate or inhibitor).

In the development of new HIV-1 PIs, Budt and co-workers described a new series of peptidomimetics with C_2-symmetry and a vicinal diol group as the central unit, with the ability to bind to the HIV-1 protease active site [5]. This study leads to HOE/BAY-793 (Hoechst/Bayer) (Fig. 1), presenting outstanding potency, especially in HIV-1 infected cell cultures [5]. In addition, structure–activity relationship (SAR) studies, developed by Budt and co-workers, show that all three side-chains in this series (P1-P3′, Fig. 1) need to be lipophilic, in order to achieve in vitro potent inhibitory activity [5]. In cases where the enzyme tolerates hydrophilic substituents, drastic reductions in anti-HIV activity (in cell culture assay) are observed, most likely due to insufficient cell penetration [5].

Quantitative structure–activity relationship (QSAR) studies may be useful in the search for sites on molecules capable of being modified into better specific ligands [6]. This methodology is mostly used to correlate structural descriptors with biological properties, but it can also be applied to predict the activity value of non-synthesized compounds structurally related to the training sets. It is a mathematical model of correlation, statistically validated, between the variation on chemical structure and the activity profile of a series of compounds. With the advent of molecular modeling, three-dimensional (3D) descriptors replaced the didactical physicochemical and bidimensional (2D) descriptors [7].

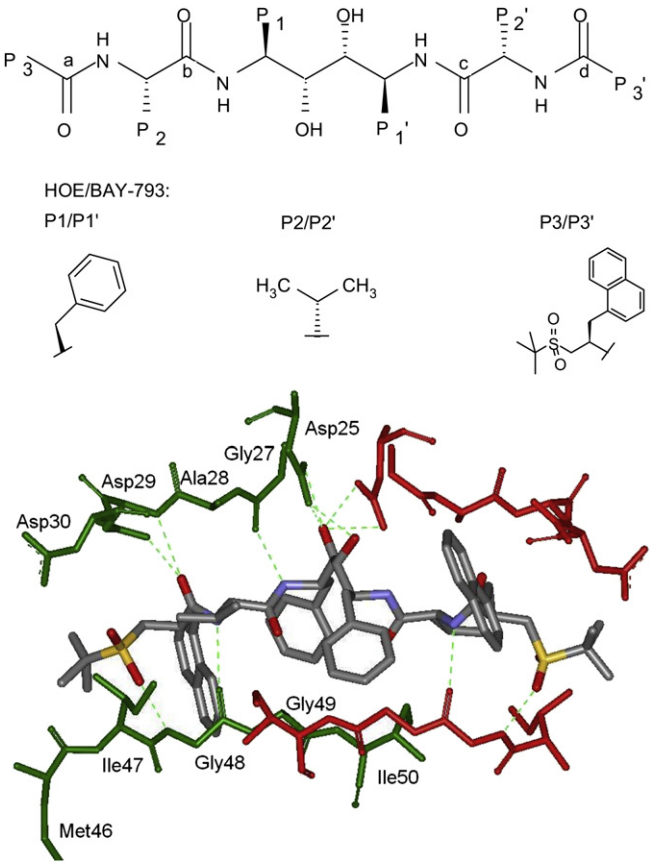

Fig. 1. General structure of HOE/BAY-793 analogs, showing P1-P3′ substituents for compound **31** (HOE/BAY-793) and the carbon atoms (a, b, c, and d) used to define Alignment 3 (multifit alignment, MFA). Close view of HOE/BAY-793 bound into the binding site of HIV-1 protease, showing inhibitor–protease hydrogen bonding interactions.

Among the 3D-QSAR methods, the Comparative Molecular Field Analysis (CoMFA), proposed by Cramer and co-workers in 1988, is extensively used in the current practice of drug discovery [8]. The CoMFA approach describes the molecular properties by 3D steric (Lennard–Jones) and electrostatic (Coulomb) fields, evaluated over a lattice of points. Partial least-squares (PLS) method is used in order to correlate the variation of these properties with the variation of the biological response [8]. Therefore, one of the advantages of CoMFA is the ability to predict the biological potency of the molecules by deriving a relation between steric-electrostatic properties and biological responses in the form of contour maps.

A similar approach to the computation of molecular potential fields has been described as the Comparative Molecular Similarity Indices Analysis (CoMSIA) by Klebe and co-workers in 1994, in which a probe atom is used to calculate similarity indices, at regularly spaced grid points, for the pre-aligned molecules [9]. Unlike CoMFA, CoMSIA uses a Gaussian-type distance-dependent function to assess five fields of different physicochemical properties (i.e., steric, electrostatic, hydrophobic, and hydrogen bonding donor and acceptor) [9]. Similarly to the CoMFA, the fields are evaluated by PLS analysis [9]. The advantage of CoMSIA over CoMFA method, it is that no singularities occur at the atomic positions, since the fields are evaluated using a Gaussian function, and therefore, no arbitrary definitions of cutoff limits should be used [9]. Moreover, the CoMSIA approach provides additional fields, measuring the compounds hydrophobicity and the ability to perform H-bonds.

The graphic representation of the CoMFA/CoMSIA fields describes the 'favorable' or 'unfavorable' contribution of a region of interest surrounding the ligands to the target property [10]. In order to obtain a reliable model, a spatial orientation of the ligands towards one another, representing the relative differences in the binding geometry at the protein-binding site, has to be found. Therefore, the choice of a good level of superposition for the studied molecules will definitely determine the success of the molecular field analysis [11].

In the present study, 3D-QSAR CoMFA/CoMSIA models were developed for a series of peptidomimetics that was synthesized and evaluated by Budt and co-workers as HIV-1 protease inhibitor [5]. In order to develop 3D-QSAR CoMFA/CoMSIA models, a hypothesis concerning the active conformation of the molecules under study should be stated as a prerequisite for structural alignment. Therefore, since the crystal structure of HOE/BAY-793 bound to HIV-PR is available (Fig. 1), it was used as a template to build the analogs, and to generate the theoretical active conformers for these peptidomimetics, derived from modeled protease–inhibitor complexes, in order to orient the compounds superposition and to afford a consistent alignment.

Moreover, the combination of the X-ray investigations on the complexes with the CoMFA/CoMSIA studies allows a much more accurate interpretation of the generated contour maps and leads to a better knowledge of the enzyme–inhibitor interaction. The obtained CoMFA/CoMSIA contour maps revealed the significance of steric, electrostatic, and hydrogen bond donor fields to the biological response. The structural variations in the molecular fields at particular regions in the space were studied and the 3D-QSAR models generated gives an insight in the design of more potent and, thereby, better anti-HIV-1 protease agents. Additionally, these models were compared with a previous 4D-QSAR model developed using this same series of compounds [12].

2. Methodology

Molecular modeling techniques described herein were performed on a Silicon Graphics O2 workstation (Silicon Graphics, Inc.) using SYBYL 6.22 (TRIPOS, Inc.) software [13].

2.1. Biological data

The biological data set was retrieved from a series of 43 HIV-1 PIs developed by Budt and co-workers [5], but only 31 compounds were selected from this series in order to afford a consistent database in terms of both chemical structure and biological data. The CoMFA/CoMSIA models were developed using 27 compounds (**1**–**15**, **17**–**18**, **20**–**24**, **26**–**29**, and **31**, Table S1, Supplemental Material) as training set, and externally validated using four compounds (**16**, **19**, **25**, and **30**, Table S1) as test set.

The biological activities of these compounds were reported as the concentration capable of inhibiting 50% of the enzyme activity (IC_{50}), using the HIV-1 protease expressed and purified from *Escherichia coli* strain K12 [5]. All pharmacological data were obtained from the same laboratory, eliminating the potential noise that might have been introduced by the pooling of data sets from different sources. The IC_{50} (nM) values were converted into molar units, and, then, expressed in negative logarithmic units (pIC_{50}, M). The range of pIC_{50} values for the training and test set spans at least three orders of magnitude (6.44–9.52 M), and, in addition, the biological activity values show a regular distribution over the entire range (Fig. S1, Supplemental Material).

The structures of the 31 compounds (training and test sets) are listed in Table S1 (Supplemental Material) and HOE/BAY-793 (compound **31**) was used as a reference, since its crystallographic structure bound to HIV-PR (Fig. 1) is available under PDB ID: 1VIJ [14]. Since the inhibitors of this series are structurally related, the training and test set compounds were assigned as follows. The 31 compounds were assembled in four sub-groups according to their increasing pIC_{50} values (Group I: $pIC_{50} = 6.01–7.00$ M, Group II: $pIC_{50} = 7.01–8.00$ M, Group III: $pIC_{50} = 8.01–9.00$ M, and Group IV: $pIC_{50} = 9.01–10.00$ M) and one compound from each group was randomly selected as a test set compound. As a result, the remaining 27 compounds comprise the training set. Using this approach, the test set compounds are spread over the whole range of activity values, and there is, at least, one test set compound from each range of biological activity. HOE/BAY-793 (**31**, Group IV) was not considered in the random selection, since it was selected as the reference compound in the training set.

2.2. Energy minimization of the crystallographic complex

Hydrogen atoms were added to the crystallographic structure of the HOE/BAY-793 HIV-1 protease complex (PDB ID: 1VIJ) [14]. Energy minimization of the hydrogen atoms, followed by energy minimization of the side-chains, keeping the backbone fixed, was carried out, using simplex and conjugate gradient algorithms, until a gradient of 0.1 kcal mol^{-1} Å^{-1} was reached. All minimization steps were performed using the TRIPOS force field within the SYBYL software [13], applying a distance-dependent dielectric function, $\varepsilon_r = D \times r_{ij}$, which was set to $3 \times r_{ij}$ in order to model the solvent effect. In this expression, r_{ij} is the distance between atoms i and j, and D is the scale factor of the dielectric function.

2.3. Energy minimization of the model complexes

The 3D structures of the 31 inhibitors (Table S1) were built using as a template the HOE/BAY-793 (Fig. 1) crystal structure bound to the HIV-PR [14]. Each inhibitor was docked in the binding site of the energy-minimized structure of the crystallographic complex, replacing the HOE/BAY-793 molecule (Fig. 1). In order to find the most stable theoretical binding mode for each inhibitor, which was used as the theoretical active conformer to orient the compounds superposition and to afford a consistent alignment, the interaction energies (steric and electrostatic) between the protein and each compound were calculated. Subsequently, geometry optimization of the entire system (enzyme–inhibitor complex) was carried out, using conjugate gradient, until reaching a gradient of 0.1 kcal mol^{-1} Å^{-1}. The partial atomic charges, required to calculate the electrostatic interactions, were assigned based on the Gasteiger–Marsili formalism [15]. In addition, for the 3D-QSAR study the atomic partial charges derived from the Austin Model 1 (AM1) Hamiltonian [16], and those assigned by the Merck Molecular Force Field (MMFF94) [17] were tested.

2.4. CoMFA/CoMSIA alignment rules

The alignment is one of the most important steps in 3D-QSAR methodologies, such as CoMFA/CoMSIA. Thus, it was assumed that each molecule binds into the enzyme active site in a similar mode, since the compounds share a common scaffold with the template (HOE/BAY-793). If the aligned molecules are supposed to share a common global shape and location in the 3D lattice, the entropic contributions to the free energy are expected to be similar [18].

Therefore, due to the intricate structure of the peptidomimetics under study that makes it difficult to align them, only three different alignment strategies were evaluated, using the theoretical active conformers derived from modeled protease–inhibitor complexes. Alignment 1 (protein-based alignment, PBA) corresponds to the enzyme backbone superposition of the protease–inhibitor complexes. Alignment 2 (database alignment, DBA) corresponds to the superposition of the common structure shared by all molecules (Fig. 1). Alignment 3 (multifit alignment, MFA) corresponds to the superposition of four atoms (*a*, *b*, *c*, and *d*, Fig. 1) selected from the common structure, which is the alignment rule applied in a previous 4D-QSAR study on this series [12].

2.5. CoMFA/CoMSIA fields calculation

The aligned sets of molecules were positioned inside four grid boxes with grid spacing values of 1.5, 2.0 (default distance), 2.5, and 3.0 Å in all Cartesian directions and CoMFA fields were calculated using the QSAR module of SYBYL [13]. The interaction energies for each molecule were calculated at each grid point using two probe atoms: an sp^3 hybridized carbon atom with a van der Waals (vdW) radius of 1.52 Å and a +1.0 charge (default probe) and an sp^3 hybridized oxygen atom with a vdW radius of 1.38 Å and a –1.0 charge.

The steric (vdW interaction) and electrostatic (Coulombic values with a $1/r$ distance-dependent dielectric function) fields were calculated at each intersection on the regularly spaced grid. In order to reduce noise and improve efficiency, column filtering (minimum sigma) was set to 2.0 kcal mol^{-1}, excluding from the analysis those columns (lattice points) whose energy variance is less than 2.0 kcal mol^{-1}. Cutoffs for both steric and electrostatic fields were set to tentative values of 30 (default cutoff), 20 and 15 kcal mol^{-1}.

CoMSIA similarity index descriptors were derived using the same lattice boxes as those used in CoMFA calculations. Five properties, i.e., steric (s), electrostatic (e), hydrophobic (h), hydrogen bond donor (d) and acceptor (a), were evaluated using a probe atom of 1.0 Å radius and +1.0 charge. In CoMSIA, the steric indices are related to the third power of the atomic radii, the electrostatic descriptors are derived from atomic partial charges [19], the hydrophobic fields are derived from atom-based parameters developed by Viswanadhan and co-workers [20], and the hydrogen bond donor and acceptor indices are obtained from a rule-based method derived from experimental values [19]. A Gaussian function is used to evaluate the mutual distance between the probe atom and each molecule atom. Because of the different shape of the Gaussian function, the similarity indices (A) can be

calculated at all grid points of the molecular surface according to Equation (1).

$$A^{q}_{F,k}(j) = -\sum w_{\mathrm{probe},k} w_{ik} e^{\alpha r^{2}_{iq}} \quad (1)$$

In Equation (1), A is the similarity index at grid point q, summed over all atoms i of the molecule j; $W_{probe,k}$ is the probe atom with radius 1.0 Å, charge +1, hydrophobicity +1, hydrogen bond donating +1, hydrogen bond accepting +1; W_{ik} is the actual value of the physicochemical property k of atom i; r_{iq} is the mutual distance between the probe atom at grid point q and atom i of the test molecule. Alfa (α) is the attenuation factor, with a default value of 0.3, and an optimal value normally ranging from 0.2 to 0.4 [19–22].

2.6. 3D-QSAR models calculation, internal and external validation

In order to generate statistically significant 3D-QSAR models, PLS regression was used to analyze the training set by correlating the variation in their pIC_{50} values (the dependent variable) with variations in their CoMFA/CoMSIA interaction fields (the independent variables).

The best 3D-QSAR models generated from the PLS analysis, ranked according to their coefficient of determination (or squared correlation coefficient) (R^2) values, were submitted to a standard internal validation technique, named "leave-one-out" cross-validation (LOO-cv), that gives the LOO-cv R^2 (q^2) as a statistical index of predictive power. In the LOO-cv procedure, the coefficients of the independent variables of the original PLS model are recalculated, excluding one compound (i.e., activity values and calculated properties) from the original training set at once, and this "new" model is used to predict the activity of the excluded compound. This procedure is repeated through the whole data set, until all compounds have been excluded once, and then, q^2 values and standard error of predictions (SEP) are calculated.

The CoMFA/CoMSIA results were graphically represented by field contour maps, where the coefficients were generated using the field type "Stdev*Coeff". In order to select appropriate contour levels for each feature, the corresponding histograms of actual field values were analyzed, and the "best" contour levels was chosen interactively as those that produce the most interpretable contour map.

2.7. External validation

In order to test the real predictive ability of the best models generated by the 3D-QSAR CoMFA/CoMSIA approaches using the training set, the pIC_{50} values of the external validation set (i.e., test set compounds **16**, **19**, **25**, and **30** not included in the training set) were calculate using the same CoMFA/CoMSIA options which generated the best models, as described before. The quality of the external prediction is documented using the standard deviation of error prediction (r^2_{pred}), according to Equation (2).

$$r^{2}_{\mathrm{pred}} = 1 - \left(\frac{\mathrm{PRESS}}{\mathrm{SD}}\right) \quad (2)$$

In Equation (2), PRESS is the sum of the squared deviations between predict and actual pIC_{50} values for the test set compounds and SD is the sum of the squared deviation between the actual pIC_{50} values of the compounds in the test set and the mean pIC_{50} value of the training set compounds.

3. Results

3.1. Selection of the best CoMFA model

In order to choose the best alignment rule for the series of peptidomimetics studied, three tentative alignments were evaluated, namely, Alignment 1 (protein-based alignment, PBA), Alignment 2 (database alignment, DBA), and Alignment 3 (multifit alignment, MFA), combining different probe atoms, cutoff values, lattice space, and partial atomic charge calculation type. However, comparing all models (data not shown) obtained with the two types of probe atoms (carbon and oxygen) and the three types of partial atomic charge calculations (Gasteiger–Marsili, AM1, and MMFF94), in general, we can see that models derived from carbon atom as probe atom and AM1 partial atomic charges revealed q^2 values higher than those models derived from oxygen as probe atom and Gasteiger–Marsili or MMFF94 partial atomic charges for this data set. It is interesting to note that the AM1 scheme charge was the same used in a previous 4D-QSAR study in this series [12].

Therefore, from this point forward, Table S2 (Supplemental Material) summarizes only the statistical parameters of CoMFA models obtained with sp^3 carbon atom probe and AM1 partial atomic charges. Steric and electrostatic interaction energies between the carbon sp^3 probe atom and each one of the aligned molecules were calculated at predefined grid points with 1.5, 2.0, 2.5 and 3.0 Å spacing. The models were internal validated using the LOO-cv procedure. The influence of different steric and electrostatic cutoff values was investigated, where the SYBYL default energy cutoff value of 30 kcal mol^{-1} was used, as well as cutoffs of 20 and 15 kcal mol^{-1} for both steric and electrostatic fields.

It can be seen on Table S2 that using protein-based alignment (Alignment 1) and database alignment (Alignment 2), the quality of the CoMFA models was poorer than using multifit alignment (Alignment 3). Therefore, evaluating the statistical quality of the models in terms of q^2, Alignment 3 was selected as the best alignment rule. Again, it is interesting to note that the Alignment 3 was the same used in a previous 4D-QSAR study on this series [12].

According to Table S2, considering all variables, the best CoMFA model corresponds to the model derived from the multifit alignment (MFA), with a grid lattice space of 3.0 Å, steric and electrostatic cutoff values of 15 kcal mol^{-1}, using sp^3 carbon probe atom (radius = 1.52 Å and charge = +1.0), and AM1 derived partial atomic charges for the ligands.

Therefore, the best CoMFA model, with four principal components (PC) from the PLS analysis, has LOO-cv R^2 (q^2) equals to 0.637 and the Standard Error of Prediction (SEP) of 0.578, while R^2 is equal to 0.991 with a Standard Error of Estimate (SEE) equals to 0.096, and the F-Fisher value (statistical significance) equals to 360.342.

3.2. Selection of the best CoMSIA model

The same parameters used to generate the best CoMFA model were used to generate the CoMSIA models, i.e., the CoMSIA models were derived from Alignment 3 (multifit alignment, MFA), grid lattice space of 3.0 Å, sp^3 carbon probe atom (+1 charge), and AM1 derived partial atomic charges for the ligands. The CoMSIA analysis was performed using 18 combinations of field descriptors: steric (s), electrostatic (e), hydrophobicity (h), hydrogen bond donor (d) and acceptor (a), and using the default value for the attenuation factor ($\alpha = 0.3$).

Therefore, Table S3 (Supplemental Material) shows the results of the CoMSIA analysis, and how much each field combination contributes to each derived model. Using the steric and electrostatic field combination (in order to compare with CoMFA model), the predictive power of the derived CoMSIA model ($q^2 = 0.389$,

Table S3) was worse than that of the best CoMFA model ($q^2 = 0.637$, Table S2). On the other hand, the best CoMSIA model ($q^2 = 0.511$, Table S3) is derived from the combination of the electrostatic and hydrogen bond donor fields.

In order to verify if the default attenuation factor value ($\alpha = 0.3$, Table S3) is the optimum value, α values equal to 0.2 and 0.4 (Table S4, Supplemental Material) were tested, using the combination of the electrostatic and hydrogen bond donor fields in order to derive CoMSIA models. The q^2 values of the derived CoMSIA models when $\alpha = 0.2$ ($q^2 = 0.325$) or $\alpha = 0.4$ ($q^2 = 0.485$) were lower than $\alpha = 0.3$ ($q^2 = 0.511$). The reduction of the attenuation factor value ($\alpha = 0.2$) means that the probe atom placed at a particular lattice point recognizes molecular similarity in its neighborhood more globally, while the increase of the α value means ($\alpha = 0.4$) that the molecular similarity will be identified more closely [23].

Therefore, the best CoMSIA model, with six PC from the PLS analysis, has LOO-cv R^2 (q^2) equals to 0.511 and SEP equals to f 0.704, while R^2 is equal to 0.987 with SEE equals to 0.114, and the *F*-Fisher value (statistical significance) equals to 255.920.

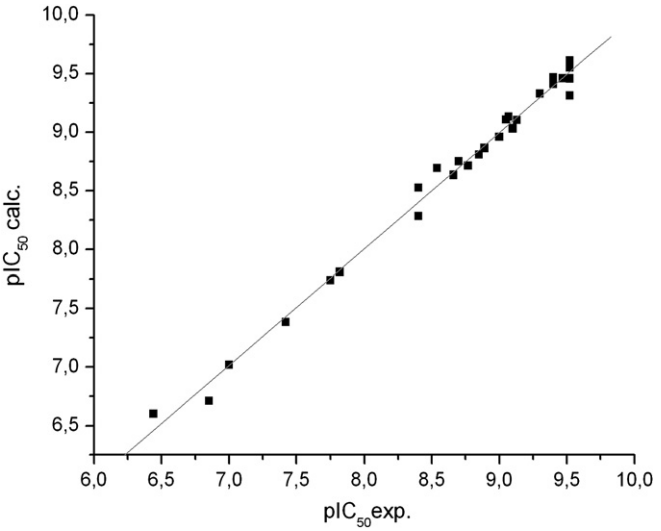

Fig. 2. Plot of the experimental versus the calculated pIC_{50} values of the BAY/HOE-793 analogs (training set), according to the best CoMFA model.

3.3. Outliers from CoMFA/CoMSIA best models

Table 1 shows the experimental (pIC_{50Exp}) and calculated (pIC_{50Calc}) potencies, using the best CoMFA/CoMSIA models, and the corresponding residual values ($pIC_{50Exp} - pIC_{50Calc}$) for both training and test sets. Figs. 2 and 3 show the plots of the pIC_{50Exp} *versus* the pIC_{50Calc} values for the training set of both CoMFA/CoMSIA models, respectively.

Outlier compounds were defined as those compounds whose residual values are higher than twice the standard deviation of the residuals (SDR). Analyses of the training set ($N = 27$) data (Table 1) showed that each one of the best CoMFA ($2 \times SDR_{CoMFA} = 0.17$) and CoMSIA ($2 \times SDR_{CoMSIA} = 0.20$) models has just one outlier, **31** (residual = 0.22) and **12** (residual = −0.21), respectively. In analogy

Table 1
Experimental and calculated (or predicted) pIC_{50} and residuals ($pIC_{50Exp} - pIC_{50Calc}$) for the training and test sets using the best CoMFA and CoMSIA models.

#[a]	pIC_{50Exp}	CoMFA pIC_{50Calc}	CoMFA Residual[b]	CoMSIA pIC_{50Calc}	CoMSIA Residual[c]
1	8.66	8.64	0.02	8.76	−0.10
2	6.44	6.60	−0.16	6.57	−0.13
3	8.54	8.70	−0.16	8.59	−0.05
4	9.40	9.41	−0.01	9.34	0.06
5	7.74	7.74	0.00	7.69	0.06
6	6.85	6.71	0.14	6.73	0.12
7	9.00	8.96	0.04	8.98	0.02
8	9.52	9.55	−0.03	9.47	0.05
9	9.30	9.33	−0.03	9.16	0.14
10	8.77	8.72	0.05	8.84	−0.07
11	9.52	9.46	0.06	9.52	0.00
12	9.40	9.47	−0.07	9.61	−0.21
13	7.00	7.02	−0.02	7.05	−0.05
14	9.52	9.61	−0.09	9.51	0.01
15	9.13	9.10	0.03	9.20	−0.07
16	6.47	8.74	−2.27	9.37	−2.90
17	9.07	9.13	−0.06	9.06	0.01
18	8.85	8.81	0.04	8.88	−0.03
19	7.66	7.95	−0.29	8.03	−0.37
20	8.70	8.75	−0.05	8.80	−0.10
21	8.89	8.86	0.03	8.83	0.06
22	7.82	7.81	0.01	7.80	0.02
23	7.42	7.38	0.04	7.41	0.01
24	9.05	9.11	−0.06	9.03	0.02
25	8.49	9.24	−0.75	9.12	−0.63
26	9.47	9.46	0.01	9.32	0.15
27	9.10	9.03	0.07	9.30	−0.20
28	8.40	8.53	−0.13	8.31	0.09
29	8.40	8.28	0.12	8.41	−0.01
30	9.10	8.85	0.25	8.74	0.36
31	9.52	9.31	0.21	9.32	0.20

[a] See Table S1 for compounds structures [5]. Training set compounds ($N = 27$). Test set compounds ($N = 4$) numbers are underlined. Outliers were defined as those compounds whose residual values are higher than twice the standard deviation of the residuals (SDR).
[b] CoMFA: Training set SDR = 0.085 and test set SDR = 1.083.
[c] CoMSIA: training set SDR = 0.099 and test set SDR = 1.407.

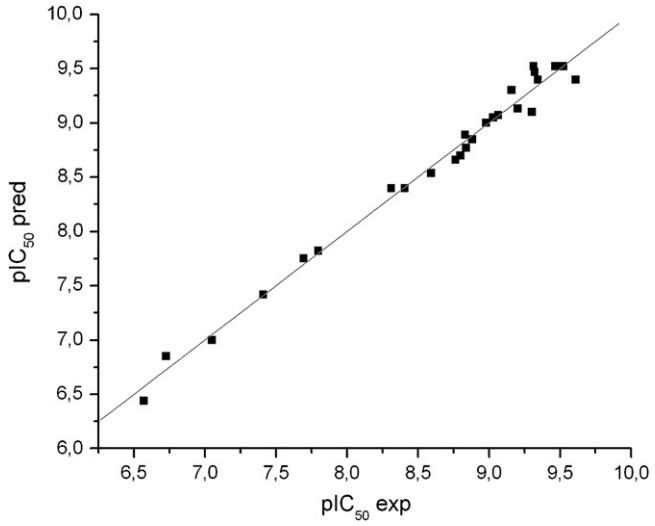

Fig. 3. Plot of the experimental versus the calculated pIC_{50} values of the BAY/HOE-793 analogs (training set), according to the best CoMSIA model.

to the training set, analysis of the test set ($N = 4$) data (Table 1) showed that **16** (CoMFA residual = −2.27 and CoMSIA residual = −2.90) is the only outlier of both best models, CoMFA ($2 \times SDR_{CoMFA} = 2.17$) and CoMSIA ($2 \times SDR_{CoMSIA} = 2.81$).

4. Discussion

4.1. CoMFA contour maps

The steric and electrostatic fields from the best CoMFA model are represented as 3D colored contour maps in Figs. 4 and 5, respectively, using compound **31** (HOE/BAY-793) as an example. The individual contributions from the steric and electrostatic fields were 45.7% and 54.3%, respectively. These contours were calculated as the product of the fields standard deviation (StDev) at each grid point and the coefficient (Coeff) from the PLS analysis (StDev*Coeff). These maps describe regions around the molecules where a substituent, with a particular steric and electrostatic characteristic, is able to increase or to decrease the biological potency (the pIC_{50} values of the HIV-1 protease inhibitors).

As stated by Hou and co-workers [24], it is important to notice that: "These maps show regions where differences in molecular

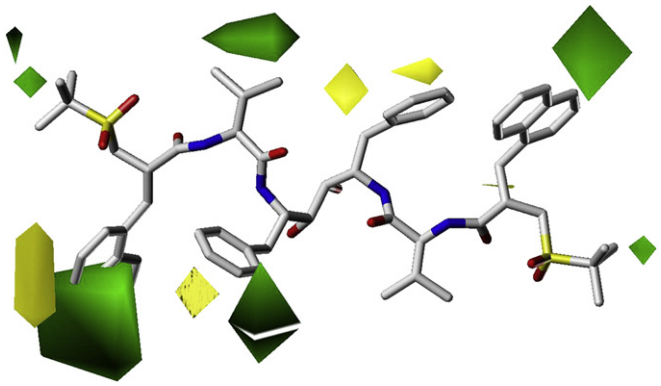

Fig. 4. The steric fields contour plots according to the best CoMFA model. Increasing HIV-PR inhibitory potency is associated with increasing (green) and decreasing (yellow) steric bulk. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

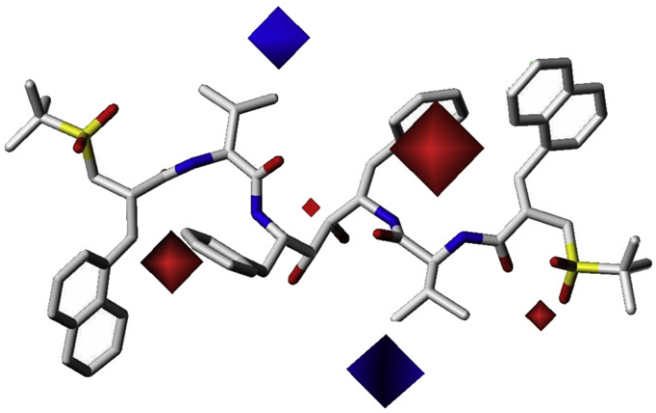

Fig. 5. The electrostatic fields contour plots according to the best CoMFA model. Increasing HIV-PR inhibitory potency is associated with increasing (red) and decreasing (blue) negative charge. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

fields are associated with differences in biological activity. The maps do not show what is common to all molecules of a set, and hence one cannot expect to generate a complete image of the receptor." Therefore, substituents, which have potency enhancing properties, but are constant to all data set, will not be detected in any kind of QSAR analysis as an important feature.

In the CoMFA steric field contour map (Fig. 4), 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. In the CoMFA electrostatic field contour map (Fig. 5), areas where electronegative substituents increase the potency are represented by red polyhedrons, while areas where electronegative substituents decrease the potency are represented by blue polyhedrons.

The contour plot of the green polyhedron (70% of contribution level) close to the isopropyl group on the P2 position of compound **31** ($pIC_{50Exp} = 9.52$) (Fig. 4) indicates a sterically favorable region, since bulky substituents, such as *sec*-butyl (**8**, $pIC_{50Exp} = 9.52$), *tert*-butyl (**9**, $pIC_{50Exp} = 9.30$) and cyclopentyl (**11**, $pIC_{50Exp} = 9.52$), in this position increase the potency. However, when the P2 substituent is the cyclohexyl group (**10**, $pIC_{50Exp} = 8.77$), also a bulky substituent, there is an unfavorable contribution to the potency, since the potency of this compound is decreased almost one logarithm unit, when compared with previous compounds **8**, **9** and **11**. This anomalous behavior could be explained by the fact that when compound **10** was originally built inside the HIV-1 protease active site, there was a steric hindrance for the cyclohexane "chair" conformation, so it was changed to the "boat" conformation. Since the "chair" is more stable (~7.1 kcal mol^{-1}) than the "boat" conformer [25], and if the "boat" (or an equivalent) is the preferential one in this particular system, this difference in energy could decrease the potency of compound **10**.

The large green polyhedron (Fig. 4) located around the 1-naphthylmethyl group on the P3 position of compound **31** ($pIC_{50Exp} = 9.52$) suggests favorable steric interactions of this structural moiety on the HIV-PR active site, but when the 1-naphthylmethyl group is replaced by the benzyl group, like for compound **25** ($pIC_{50Exp} = 8.49$), the potency decreases almost one logarithm unit. The same behavior was observed when there is not any aromatic group in the P3′ position, like for compound **28** ($pIC_{50Exp} = 8.40$).

Small green polyhedrons were also observed close to the methyl groups from the *tert*-butyl-sulfonyl group on the P3/P3′ position of compound **31** (Fig. 4) filling in the pocket created by the residues

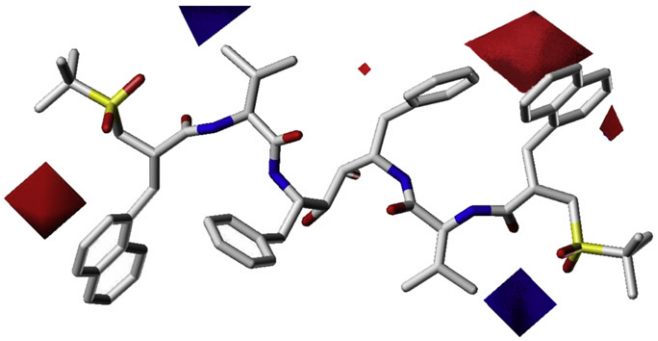

Fig. 6. The electrostatic fields contour plots according to the best CoMSIA model. Increasing HIV-PR inhibitory potency is associated with increasing (red) and decreasing (blue) negative charge. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Asp29, Asp30, Lys45, Ile47, and Ile76 of the HIV-PR. This bulky substituent (*tert*-butyl) in this position is allowed and correlates with improved biological activity profile, since from the 21 compounds presenting the *tert*-butyl group in this position (P3/P3′), 16 compounds show experimental pIC_{50} values larger than eight.

The contour plot of the yellow polyhedrons (contribution level of 30%) surround the benzyl group on the P1/P1′ positions of compound **31** (Fig. 4) indicates sterically unfavorable regions, since bulky substituents larger than the benzyl group in this position reduce the potency. This effect may be clearly observed by comparing the experimental potency values among compounds **2** (P1/P1′ = phenylethyl, $pIC_{50Exp} = 6.44$), **6** (P1/P1′ = 4-pyridylethyl, $pIC_{50Exp} = 6.85$), **5** (P1/P1′ = 2,4-dimethoxyl, $pIC_{50Exp} = 7.75$), and **31** (P1/P1′ = benzyl, $pIC_{50Exp} = 9.52$), suggesting limited bulk tolerance.

However, at the P1 position, there is also a medium size green polyhedron close to the benzyl group, indicating a sterically favorable region. This fact is due to the presence of bulky residues like Ile50, Ile84, Pro81, Val82, and Leu23 in the active pocket close to P1′ position, originating steric hindrance with the benzyl group, while in the P1 position, the benzyl group is outside this region. This asymmetric behavior is not unique, since symmetric inhibitors can bind to HIV-1 protease asymmetrically [14,26]. The higher potency of compound **4** ($pIC_{50Exp} = 9.40$), when compared with compounds **2** ($pIC_{50Exp} = 6.44$), **5** ($pIC_{50Exp} = 7.75$), and **6** ($pIC_{50Exp} = 6.85$), is probably because the 3,4-methylenedioxyl group of **4** occupies this medium size green polyhedron.

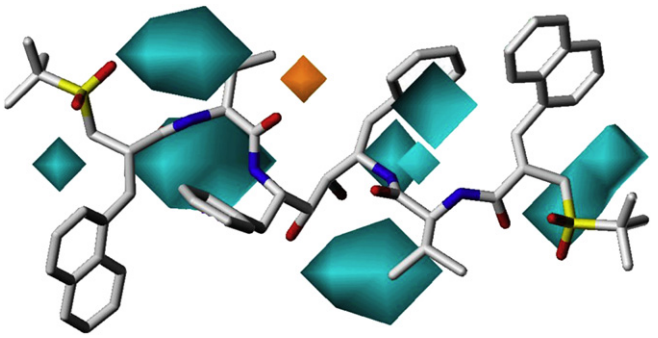

Fig. 7. The hydrogen bond donor field contour plots according to the best CoMSIA model. Increasing HIV-PR inhibitory potency is associated with increasing (cyan) and decreasing (orange) hydrogen bond donor ability. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Summary of the statistical results (SR) for the best CoMFA and CoMSIA models from this work comparing with the best 4D-QSAR (Model 2 and Model 7) models from a previous work [12].

Method/model	CoMFA	CoMSIA	4D-QSAR	4D-QSAR
SR[a]	Best model	Best model	Model 2	Model 7
q^2	0.637	0.511	0.855	0.796
R^2	0.991	0.987	0.893	0.834
N	27	27	28	28
Outliers	1 (31)	1 (12)	2 (8 and 28)[b]	0

[a] q^2 = cross-validated R^2; R^2 = coefficient of determination; N = number of training set compounds; Outliers = number of outliers (and compound number).
[b] Compound **28** from Ref. [12] was excluded from the CoMFA and CoMSIA analysis.

A medium to large yellow polyhedron was also observed close to 1-naphthylmethyl group on the P3 position of compound **31** (Fig. 4), however, this region is not associated to compounds with a 1-naphthylmethyl group. In fact, this polyhedron is located close to the aromatic rings of the compounds **20**, **21**, and **22**, reducing their observed potency.

The contour plot of the red polyhedrons (contribution level of 30%) is close to the following groups of compound **31** (Fig. 5): the hydroxyl group on the position P1 (a small red polyhedron close to the central diol vicinal group), the benzyl group on the P1/P1′ positions, and the sulfonyl group in the P3′ position. These red polyhedrons indicate active site favorable regions of high electron density, since electronegative substituents in these positions increase the potency. The most potent compounds have an aromatic ring (e.g., phenyl ring of the benzyl group) at P1/P1′ positions, which contribute to increase the compounds potencies due the electron density from the aromatic pi electron system. However, this aromatic ring should be localized at a precise distance as in the benzyl group of compound **31**, since just an additional CH_2 group, like in compounds **2** (P1/P1′ = phenylethyl, $pIC_{50Exp} = 6.44$) and **6** (P1/P1′ = 4-pyridylethyl, $pIC_{50Exp} = 6.85$), decreases the potency.

Finally, the contour plot of the blue polyhedron (contribution level of 70%) close to the isopropyl group on the P2/P2′ positions of compound **31** (Fig. 5) indicates active site unfavorable regions in order to accommodate high electron density groups, since electronegative substituents in these positions decrease the potency. This effect may be clearly observed comparing the potencies among compounds **13** ($pIC_{50Exp} = 7.00$), **16** ($pIC_{50Exp} = 6.47$), and **31** ($pIC_{50Exp} = 9.52$).

4.2. CoMSIA contour maps

The electrostatic and hydrogen bond donor fields from the best CoMSIA model are represented as 3D colored contour maps in Figs. 6 and 7, respectively, using compound **31** (HOE/BAY-793) as an example. The individual contributions from the electrostatic and hydrogen bond donor fields were 68.5% and 31.5%, respectively. These contours were calculated as described for CoMFA. These maps describe regions around the molecules where a substituent, with a particular characteristic (in these models, electrostatic charge and hydrogen bond donor ability), is able to increase or to decrease the biological potency (the experimental pIC_{50} values of the HIV-1 protease inhibitors).

The color scheme used in the CoMSIA electrostatic field contour map (Fig. 6) is the same described for the CoMFA electrostatic map. In the CoMSIA hydrogen bond donor field contour map (Fig. 7), the cyan and orange polyhedrons indicate areas where substituents with hydrogen bond donor ability increase or decrease, respectively, the compounds potency.

Table 3
P1/P1′, P2/P2, and P3/P3′ substituents characteristics according to the best CoMFA, CoMSIA and 4D-QSAR models.

QSAR models	P1/P1′	P2/P2′	P3/P3′
CoMFA steric field	Yellow polyhedrons: bulky substituents decrease the potency (tolerance limit = benzyl group).	Green polyhedrons: bulky substituents increase the potency (tolerance limit = cyclopentyl group).	Green polyhedrons: bulky substituents increase the potency (tolerance limit = 1-naphthylmethyl group).
CoMFA and CoMSIA electrostatic fields	Red polyhedrons: electronegative substituents increase the potency.	Blue polyhedrons: electronegative substituents decrease the potency.	Red polyhedrons: electronegative substituents increase the potency.
4D-QSAR	Grid cell occupied by non-polar atoms decreases the potency.	There is none grid cell corresponding to this position.	Grid cells occupied by aromatic atoms and non-polar atoms increase the potency.

The CoMSIA electrostatic contour plot map (Fig. 6) is similar to those of the CoMFA (Fig. 5). The contour plot of the red polyhedrons (contribution level of 30%) close to the benzyl and 1-naphthylmethyl groups on the P1′ and P3/P3′ positions, respectively, of compound **31** (Fig. 6) indicates active site favorable region of high electron density, since electronegative substituents in these positions increase the potency. The contour plot of the blue polyhedron (contribution level of 70%) close to the isopropyl group on the P2/P2′ of compound **31** (Fig. 6), indicates active site unfavorable region in order to accommodate high electron density groups, since electronegative substituents in this position decrease the potency.

In our previous docking studies [12], it was observed that compound **31** interacts with the HIV-PR active site through an extensive hydrogen bond network, including mainly the following residues: Asp25/Asp25′, Asp29/Asp29′, Gly27/Gly27′, and Gly48/Gly48′ [12]. Therefore, in order to understand the importance of such interactions in the series of C_2-symmetric diols studied as inhibitors of HIV-1 protease, it was performed a correlation between the compounds (substituents) hydrogen bond donor ability and the compounds potency.

The contour plot of the cyan polyhedrons (contribution level of 80%) close to almost all P/P′ positions of compound **31** (Fig. 7) indicates active site favorable regions in order to accommodate groups with hydrogen bond donor ability, since substituents such as –OH and –NH– groups in these positions increase the inhibitory potency. Cyan polyhedron, close to –NH– group on the P1/P1′ (contiguous amide group) and P2 positions of the compound **31** (Fig. 7), are probably correlated to Gly27/Gly27′and Gly48/Gly48′ residues in the HIV-1 protease active site. Cyan polyhedrons were also observed close to the sulfonyl group on the P3/P3′ positions of compound **31** (Fig. 7), however, they are not related to the sulfonyl ($-SO_2-$) group itself, since this group is not a hydrogen bond donor. In fact, these cyan polyhedrons corresponds to the P3/P3′ substituents of compounds **17** ($pIC_{50Exp} = 9.07$), **18** ($pIC_{50Exp} = 8.85$), **24** ($pIC_{50Exp} = 9.05$), and **30** ($pIC_{50Exp} = 9.10$). These compounds have hydrogen bond donor groups in these positions and the corresponding cyan polyhedrons are probably correlated to Asp29/Asp29′ residues in the HIV-1 protease active site.

Finally, the contour plot of the orange polyhedron (contribution level of 20%) close to the carbonyl group on the P2 position of compound **31** (Fig. 7) indicates an active site unfavorable region in order to accommodate groups with hydrogen bond donor ability. This effect may be explained from the fact that there is a steric hindrance in this region, diminishing the probability of hydrogen bond interaction between the ligands and the amino acid residues.

4.3. Comparison of the CoMFA/CoMSIA and 4D-QSAR models

3D-QSAR methods that can incorporate molecular flexibility, such as the Receptor Independent (RI) and the Receptor Dependent (RD) 4D-QSAR methods proposed by Hopfinger and co-workers [27–31], proved to be advantageous because they allow to explore large degrees of both conformational and alignment freedoms in the search for the "bioactive" conformation and ligand–receptor binding mode, respectively, being suited to study flexible molecules having multiple candidate pharmacophore sites [29].

The RI-4D-QSAR method was applied to the same series of HOE/BAY-793 analogs used in this CoMFA/CoMSIA study. The detailed methodology was describe elsewhere [12], but in short, the models were developed using 28 compounds, including HOE/BAY-793 and a compound which has been excluded from the CoMFA/CoMSIA analysis (pyridine-*N*-oxide derivative of compound **27**), and externally validated using the same test set. The HOE/BAY-793 crystal structure (PDB ID: 1VIJ) [14] was used for modeling all compounds inside the active site of HIV-PR. The ligand–enzyme complexes were geometry optimized (CVFF/InsightII), keeping the backbone fixed, and the final ligand structures were submitted to AM1 semi-empirical calculation (AMPAC-MOPAC/InsightII).

The RI-4D-QSAR analysis was carried out and the best 4D-QSAR model was graphically represented by plotting the significant grid cells in space along with their descriptor attributes. The postulated "bioactive" conformation of HOE/BAY-793, according to the best 4D-QSAR model, was overlapped to its crystal structure in complex with HIV-PR, showing a similar conformation. Moreover, the 4D-QSAR results showed that future inhibitor design studies should focus on the length and volume of the P3/P3′ substituents in order to avoid steric hindrance with the side chain of Pro81 and Val82 as predicted by the model.

A statistical summary comparison of the best CoMFA/CoMSIA and 4D-QSAR models for the series of HOE/BAY-793 analogs is presented in Table 2, where it can be observed that q^2 values of the best 4D-QSAR models are appreciably higher than those from the best CoMFA/CoMSIA models. On the other hand, R^2 values of the best 4D-QSAR models are lower than those presented for the best CoMFA/CoMSIA models. Since CoMFA/CoMSIA methods use only one conformation for each compound, the respective derived models are quite sensitive to training set conformation and alignment. In addition, Table 3 shows the characteristics summary of the CoMFA/CoMSIA and 4D-QSAR descriptors concerning positions P1/P1′, P2/P2′, and P3/P3′.

5. Conclusions

In order to obtain robust predictions, we realize that it is very important to use the insights from all the best 3D-QSAR models, i.e., CoMFA, CoMSIA, and 4D-QSAR models, in order to propose new derivatives of HOE/BAY-793. Taken together, the three methodologies suggest the following characteristics for the P1/P1′, P2/P2′, and P3/P3′ positions, considering HOE/BAY-793 as a reference compound:

(i) P1/P1′ positions — Substituents smaller than or with the same size and length of the benzyl group and with high electron density (e.g., aromatic pi system) should be preferential, indicating that there is a counterpart "receptor" unfavorable region for apolar substituents greater than benzyl group (steric hindrance), but, at the same time, a "receptor" favorable region for hydrogen bonding interactions.

(ii) P2/P2′ positions — Substituents smaller than or with the same size of the cyclopentyl group and with low electron density should be preferential, indicating that there is a counterpart "receptor" unfavorable region for polar groups and substituents greater than cyclopentyl group (steric hindrance).

(iii) P3/P3′ positions — Aromatic and apolar substituents should be preferential, indicating that there is a counterpart "receptor" favorable region for hydrophobic substituents.

Acknowledgement

We thank to the Brazilian agencies Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) for their support. E.F.F.C wants to thank to German Academic Exchange Service "Deutscher Akademischer Austausch Dienst, DAAD" for the graduate scholarship in Germany. R.B.A. and O.A.C.A. have research scholarships from CNPq, Brazil.

Appendix. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.ejmech.2009.05.016.

References

[1] A. Wlodawer, J.W. Erickson, Annu. Rev. Biochem. 62 (1993) 543–585.
[2] W.E. Harte Jr., S. Swaminathan, M.M. Mansuri, J.C. Martin, I.E. Rosenberg, D.L. Beveridge, Proc. Natl. Acad. Sci. U S A 87 (1990) 8864–8868.
[3] W.R.P. Scott, C.A. Schiffer, Structure 8 (2000) 1259–1265.
[4] E. de Clercq, J. Med. Chem. 48 (2005) 1297–1313.
[5] K.H. Budt, A. Peyman, J. Hansen, J. Knolle, C. Meichsner, A. Paessens, D. Ruppert, B. Stowasser, Bioorg. Med. Chem. 3 (1995) 559–571.
[6] R. Garg, S.P. Gupta, H. Gao, M.S. Babu, A.K. Debnath, C. Hansch, Chem. Rev. 99 (1999) 3525–3602.
[7] P. Gayathri, V. Pande, R. Sivakumar, S.P. Gupta, Bioorg. Med. Chem. 9 (2001) 3059–3063.
[8] R.D. Cramer III, D.E. Patterson, J.D. Bunce, J. Am. Chem. Soc. 110 (1988) 5959–5967.
[9] G. Klebe, U. Abraham, T. Meitzner, J. Med. Chem. 37 (1994) 4130–4146.
[10] W. Sippl, H.D. Holtje, J. Mol. Struct. (Theochem) 503 (2000) 31–50.
[11] W. Sippl, J. Comput. Aided Mol. Des. 16 (2002) 825–830.
[12] E.F.F. da Cunha, M.G. Albuquerque, O.A.C. Antunes, R.B. de Alencastro, QSAR Comb. Sci. 24 (2005) 240–253.
[13] SYBYL 6.22, TRIPOS International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA 2001.
[14] G. Lange-Savage, H. Berchtold, A. Liesum, K.H. Budt, A. Peyman, J. Knolle, J. Sedlacek, M. Fabry, R. Hilgenfeld, Eur. J. Biochem. 248 (1997) 313–322.
[15] J. Gasteiger, M. Marsili, Tetrahedron 36 (1980) 3219–3228.
[16] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc. 107 (1985) 3902–3909.
[17] T.A. Halgren, J. Comput. Chem. 17 (1996) 490–519.
[18] R. Perkins, H. Fang, W. Tong, W.J. Welsh, Environ. Toxicol. Chem. 22 (2003) 1666–1679.
[19] G. Folkers, A. Merz, D. Rognan, in: H. Kubinyi (Ed.), 3D-QSAR in Drug Design. Theory, Methods and Applications, ESCOM, Leiden, The Netherlands, 1993, pp. 583–618.
[20] V.N. Viswanadhan, A.K. Ghose, G.R. Revankar, R.K. Robins, J. Chem. Inf. Comput. Sci. 29 (1989) 163–172.
[21] S. Kamath, J.K. Buolamwini, J. Med. Chem. 46 (2003) 4657–4668.
[22] M. Bohm, J. Sturzebecher, G. Klebe, J. Med. Chem. 42 (1999) 458–477.
[23] T.J. Hou, Z.M. Li, Z. Li, J. Liu, X.J. Xu, J. Chem. Inf. Comput. Sci. 40 (2000) 1002–1009.
[24] T.J. Hou, L.L. Zhu, L.R. Chen, X.J. Xu, J. Chem. Inf. Comput. Sci. 43 (2003) 273–287.
[25] U. Burkert, N.L. Allinger, Molecular Mechanics, ACS Monograph 177, American Chemical Society, Washington, 1982.
[26] G.B. Dreyer, J.C. Boehm, B. Chenera, R.L. DesJarlais, A.M. Hassell, T.D. Meek, T.A. Tomaszek Jr., M. Lewis, Biochemistry 32 (1993) 937–947.
[27] A.J. Hopfinger, S. Wang, J.S. Tokarski, B. Jin, M.G. Albuquerque, P.J. Madhav, C. Duraiswami, J. Am. Chem. Soc. 119 (1997) 10509–10524.
[28] M.G. Albuquerque, A.J. Hopfinger, E.J. Barreiro, R.B. de Alencastro, J. Chem. Inf. Comput. Sci. 38 (1998) 925–938.
[29] M. Ravi, A.J. Hopfinger, R.E. Hormann, L. Dinan, J. Chem. Inf. Comput. Sci. 41 (2001) 1587–1604.
[30] D.H. Pan, Y.F. Tseng, A.J. Hopfinger, J. Chem. Inf. Comput. Sci. 43 (2003) 1591–1607.
[31] X. Hong, A.J. Hopfinger, J. Chem. Inf. Comput. Sci. 43 (2003) 324–336.