Design, Molecular Modeling, Synthesis, and Anti-HIV-1 Activity of New Indolyl Aryl Sulfones. Novel Derivatives of the Indole-2-carboxamide

14 AUTHORS, INCLUDING:

Rino Ragno
Sapienza University of Rome
109 PUBLICATIONS   2,078 CITATIONS

SEE PROFILE

Antonio Lavecchia
University of Naples Federico II
118 PUBLICATIONS   2,415 CITATIONS

SEE PROFILE

Alberto Bergamini
University of Rome Tor Vergata
68 PUBLICATIONS   1,067 CITATIONS

SEE PROFILE

Emmanuele Crespan
Institute of Molecular Genetics IGM
64 PUBLICATIONS   895 CITATIONS

SEE PROFILE

Design, Molecular Modeling, Synthesis, and Anti-HIV-1 Activity of New Indolyl Aryl Sulfones. Novel Derivatives of the Indole-2-carboxamide

Rino Ragno,*,† Antonio Coluccia,† Giuseppe La Regina,† Gabriella De Martino,† Francesco Piscitelli,† Antonio Lavecchia,‡ Ettore Novellino,‡ Alberto Bergamini,§ Chiara Ciaprini,§ Anna Sinistro,§ Giovanni Maga,|| Emanuele Crespan,|| Marino Artico,† and Romano Silvestri*,†

Dipartimento di Studi Farmaceutici, Istituto Pasteur−Fondazione Cenci Bolognetti, Università di Roma "La Sapienza", Piazzale Aldo Moro 5, I-00185 Roma, Italy, Dipartimento di Sanità Pubblica e Biologia Cellulare, Università di Roma "Tor Vergata", Via Tor Vergata 135, I-00133 Roma, Italy, Istituto di Genetica Molecolare−CNR, Via Abbiategrasso 207, I-27100 Pavia, Italy, and Dipartimento di Chimica Farmaceutica e Tossicologica, Università di Napoli "Federico II", Via Domenico Montesano 49, I-80131, Napoli, Italy

Received December 15, 2005

Molecular modeling studies and an updated highly predictive 3-D QSAR model led to the discovery of exceptionally potent indolyl aryl sulfones (IASs) characterized by the presence of either a pyrrolidyn-2-one nucleus at the indole-2-carboxamide or some substituents at the indole-2-carbohydrazide. Compounds **7** and **9** were found active in the sub-nanomolar range of concentration in both MT-4 and C8166 cell-based anti-HIV assays. These compounds, and in particular compound 9, also showed excellent inhibitory activity against both HIV-112 and HIV-AB1 primary isolates in lymphocytes and against HIV WT in macrophages.

Introduction

Human immunodeficiency virus (HIV) is the causative agent of acquired immune deficiency syndrome (AIDS), an infection characterized by loss of helper T lymphocytes and heavy damage to the lymphatic tissues. Global estimates by the World Health Organization showed that this viral infection is widespread, with 10 million HIV/AIDS infected people in 1990 and 40 million people by the end of 2005.[1]

There are currently 20 approved anti-retroviral agents (source FDA, May 2004), including some mixtures, falling in four classes: non-nucleoside reverse transcriptase inhibitors, nucleoside and nucleotide reverse transcriptase inhibitors (NRTI), protease inhibitors (PI), and the recently licensed fusion inhibitor enfuvirtide.[2] Highly active anti-retroviral therapy (HAART) based on combinations of these drugs allowed great progress in the clinical management of AIDS.[3] Nevertheless, HAART requires long-term or permanent treatments, because it cannot completely suppress viral multiplication. As a consequence, HAART leads to the emergence of drug-resistant mutant strains.

NNRTIs have received a great amount of attention due to low toxicity and favorable pharmacokinetics properties. The discovery of second-generation agents endowed with high potency and a broad spectrum of activity against drug resistant viral strains launched a new era for NNRTIs.[4,5]

Our extensive studies on sulfone NNRTIs led to the discovery of potent anti-HIV classes, such as pyrrolo[1,2-*b*][1,2,5]-benzothiadiazepine 5,5-dioxides (PBTDs),[6] pyrryl aryl sulfones (PASs),[7,8] and indolyl aryl sulfones (IASs).[9]

Recently, we reported that hybrid IAS−peptide derivatives bearing one to three glycine or alanine units as elongated 2-carboxamide chains (e.g. **1** and **2**) were more active against some clinically relevant mutant strains[10,11] than the parent indole

Chart 1
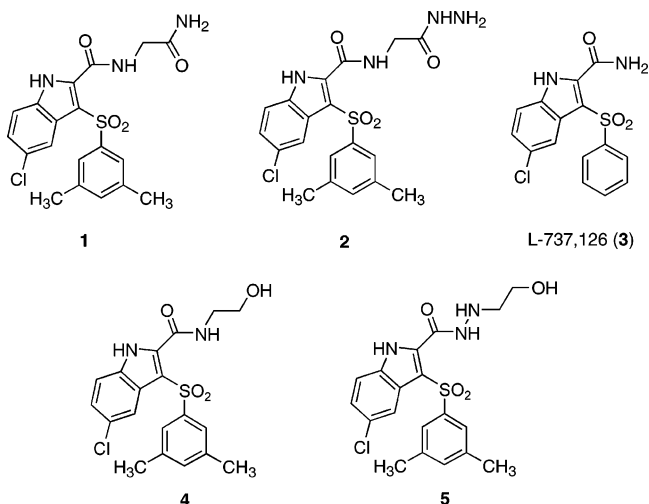
L-737,126 (**3**)[12] (Chart 1). These results supported our hypothesis that intramolecular transposition or replacement of bioisosteric groups would favor the remodeling of the inhibitor with the non-nucleoside binding site (NNBS) residues. These findings stimulated us to evaluate new substitutions at the indole-2-carboxamide function of compound **3**.

Previously, three-dimensional quantitative structure−activity relationship (3-D QSAR) studies and docking simulations led us to design novel *N*-(2-hydroxyethyl)carboxamide and *N*′-(2-hydroxyethyl)carbohydrazide derivatives. These experimental inhibitory activities were found to be in good agreement with the 3-D QSAR model predicted values. Notably, the new IASs **4** and **5** were more potent than compound **2** against the viral mutant strains.[13−15]

Pursuing this research line, we used refined molecular modeling studies and the application of an updated highly predictive 3-D QSAR model to design new IASs characterized by either the presence of a pyrrolidyn-2-one nucleus at the indole-2-carboxamide (compounds **6**, **7**) or some substituents at the indole-2-carbohydrazide (compounds **8**−**13**) (Table 1). Experimental data revealed that the new products were excep-
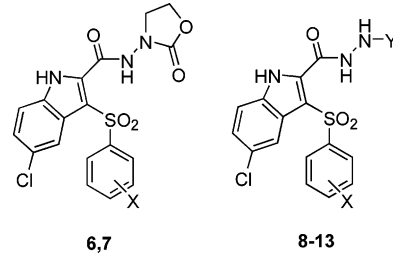
Table 1. Structures of the Designed Compounds **6**−**13**

compd	X	Y	compd	X	Y
6	H	−	**10**	3.5-Me$_2$	c-Hex
7	3.5-Me$_2$	−	**11**	3.5-Me$_2$	COMe
8	3.5-Me$_2$	Me	**12**	3.5-Me$_2$	COOEt
9	3.5-Me$_2$	i-Pr	**13**	3.5-Me$_2$	CONHNH$_2$

tionally potent against both WT and drug resistant mutants and gave us some insight for the development of a newer IAS generation.

3-D QSAR Studies. A widened 3-D QSAR model was developed both to understand the role of substitution at the 2-carboxamido position of L-737,126 and to design new and hopefully more potent anti-HIV derivatives.[14]
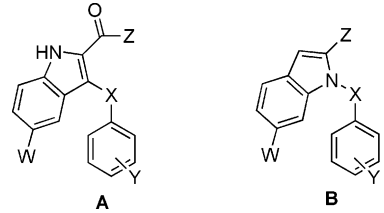
A training set of 101 IASs derivatives was compiled by adding to the previous training set (70 derivatives) test (eight *N*-benzenesulfonylindoles) and prediction (23 2-carboxamide/2-carboxhydrazide derivatives) data sets (Table 2). As reported[14] the structures of the training set were modeled starting from the closely structurally related diaryl sulfone 739W94[16] extracted from the corresponding complex with the HIV-1 RT (PDB entry code 1jlq). The training set was receptor-based aligned by means of docking experiments into the NNBS extracted from the 739W94/RT complex. The program GOLPE[17,18] was used to define the 3-D QSAR model using as a dependent variable the inverse of the experimental anti-HIV-1 activities (pEC$_{50}$s) logarithm.

After data pretreatment (see Experimental Section) a series of consecutive fractional factorial design (FFD) selections were conducted to refine the initial models. The FFD selections were carried out until no further statistical improvements were observed (Table 3). To measure the goodness of the model, the statistical indexes r^2, q^2, and SDEP were employed. The final 3-D QSAR model was achieved after five FFD selections. It was characterized by a conventional correlation coefficient value (r^2) of 0.90. The leave one out (LOO$_{CV}$), five group (L5O$_{CV}$), and leave half out (LHO$_{CV}$) cross-validation (CV) predictive correlation coefficients values (q^2_{LOO}, q^2_{L5O}, and q^2_{LHO}) corresponded to 0.79, 0.78, and 0.73. The LOO, L5O, and LHO cross-validated standard deviation errors of prediction (SDEP$_{CV-LOO}$, SDEP$_{CV-L5O}$, and SDEP$_{CV-LHO}$) values were 0.75, 0.77, and 0.86, respectively (Table 3). The group method for the FFD selections was used, setting to five the number of groups to be used. As with any 3-D QSAR model, an external validation should be compulsory. Thus, a test set (TS) of eight recent hybrid IAS−peptide derivatives[19] was used to evaluate the predictive ability of the model (Table 4). Such a test set was particularly suitable to check our hypothesis. In fact, the structures differ from those of the training set only in the substitution in the indole 2 position. The application of the final 3-D QSAR model to the TS led to a small standard deviation error of prediction (SDEP$_{TS}$ = 0.89), which was also comparable to the SDEP$_{CV}$ values. Although defined with three principal components (PCs), the actual updated 3-D QSAR model, compared to the previous one, proved to be statistically more

robust (compare r^2, q^2, SDEP$_{CV}$ in Table 5). Figure H in Supporting Information shows the $T-U$ scores plots for the first three PCs. From this plot, it is possible to have a clear idea of the correlation between the X's and the Y's obtained in the model for each PC. The first PCs plot is by far the most informative, containing the main relationship between activities (pEC$_{50}$s) and structural descriptors (field points). This information would be completely lost by only looking at the "predicted vs experimental" plot (Figure 1). The plot can be useful to identify influential objects or clusters of objects (outliers). Usually these objects do not correlate in the first component. Thus, the second or third PC is made to fit them. They appear as some (or few) objects completely distinct from the rest of the objects in the $T-U$ scores plot in either PC. Accordingly, the $T-U$ scores plots show this behavior for nonsignificant components. As seen from the $T-U$ scores plots in Figure H (see Supporting Information), there is a nice correlation between the independent variables X scores (T) and the dependent variable Y scores (U). In addition, no outliers are visible, thus confirming that the goodness of the 3-D QSAR model.

Regarding the predictive ability within an external test set, the latest model generally had a higher predictive ability, even though five out of the eight compounds in the test set had predicted pIC$_{50}$s comparable to those of the old model (compare predictions in Table 4 and SDEP$_{TS}$ values in Table 5). In fact, using the old model, derivatives **112** and **114** were predicted as false negative in the high micromolar range, while the new model was able to predict them as active compounds in the low or submicromolar range. Moreover, the new model correctly predicted the overall activity trend of the TS, while the old model did not. To have a clearer idea of the predictive ability, the size of the test set should be expanded. However, the aim of the present research was not just to build a respectable 3-D QSAR model but to evaluate the model itself as a tool to predict new and untested derivatives (see the next section).

One important feature of the 3-D QSAR studies is the graphical representation of the model, which allows an easier interpretation. In the GOLPE software several options are available to display the final model. Among these options, the PLS pseudocoefficients, the activity contribution, and the present fields plots are very useful. The PLS pseudocoefficients plot allows the visualization of the selected grid points at a determined molecules/probe interaction energy level, indicating the whole training set's graphical information. However, in this plot the signs of the coefficients can induce errors. This occurs due to the fact that the compounds produce positive or negative field values in a specific area, causing the coefficients to have opposite meanings. On the other hand, the activity contribution plot makes it possible to display spatial regions that are individually important for the selected molecule. The activity contribution plot, which is different for every molecule within the training set, is calculated by multiplying the values of the coefficients by the actual field values for every molecule. The present fields plot, which is also different for every molecule within the training set, allows the visualization of the actual GRID field coming out of interactions with the probe (OH2). To analyze a 3-D QSAR model, representative compounds were chosen from the highest and lowest active derivatives within the training set. Moreover, to better highlight the enhanced features of the updated model, representative compounds were selected among IAS derivatives containing a substitution on the amido/hydrazido group at position 2 of the indole. Figures 2−4 show the PLS coefficients, the activity contribution, and the

Table 2. Structures, ALOGpS Values, and Anti-HIV-1 Activities of Compounds 3−5 and 14−111, the Training Set Used for the 3-D QSARs

compd	structure	X	Y	Z	W	EC$_{50}$ (μM)	ALOGpS (mg/L)
3[a]	A	SO$_2$	H	NH$_2$	Cl	0.001	−5.6
14[a]	A	S	H	OEt	H	1.4	−7.49
15[a]	A	S	2-NH$_2$	OEt	H	>200	−5.23
16[a]	A	S	2-NH$_2$-5-Cl	OEt	H	≥200	−5.85
17[a]	A	S	2-NH$_2$	OEt	Cl	2.3	−4.93
18[a]	A	S	2-NH$_2$-5-Cl	OEt	Cl	2.5	−4.45
19[a]	A	SO$_2$	H	OEt	H	3.7	−4.68
20[a]	A	SO$_2$	2-NH$_2$	OEt	Cl	>200	−4.08
21[a]	A	SO$_2$	2-NH$_2$-5-Cl	OEt	H	2.5	−6.3
22[a]	A	SO$_2$	2-NH$_2$-5-Cl	OEt	Cl	1.9	−5.49
23[a]	A	S	H	NH$_2$	H	1.4	−5.06
24[a]	A	S	2-NH$_2$-5-Cl	NH$_2$	H	9	−8.01
25[a]	A	S	H	NH$_2$	Cl	0.02	−6.1
26[a]	A	S	2-Me	NH$_2$	Cl	0.3	−5.07
27[a]	A	S	4-Me	NH$_2$	Cl	>0.45	−5.71
28[a]	A	S	4-F	NH$_2$	Cl	1.4	−4.45
29[a]	A	S	4-Cl	NH$_2$	Cl	3.1	−6.24
30[a]	A	S	4-*i*-Pr	NH$_2$	Cl	1.9	−5.51
31[a]	A	S	4-*t*-Bu	NH$_2$	Cl	8	−5.73
32[a]	A	S	3,5-Me$_2$	NH$_2$	Cl	0.006	−7.32
33[a]	A	S	2,6-Cl$_2$	NH$_2$	Cl	1.2	−4.69
34[a]	A	S	2-NH$_2$-5-Cl	NH$_2$	Cl	1.6	−6.17
35[a]	A	SO$_2$	H	NH$_2$	H	0.18	−5.55
36[a]	A	SO$_2$	2-NH$_2$-5-Cl	NH$_2$	H	0.3	−5.11
37[a]	A	SO$_2$	2-Me	NH$_2$	Cl	0.001	−5.31
38[a]	A	SO$_2$	3-Me	NH$_2$	Cl	0.001	−5.59
39[a]	A	SO$_2$	4-Me	NH$_2$	Cl	0.003	−5.07
40[a]	A	SO$_2$	4-F	NH$_2$	Cl	0.014	−7.04
41[a]	A	SO$_2$	4-Cl	NH$_2$	Cl	0.011	−4.51
42[a]	A	SO$_2$	4-*i*-Pr	NH$_2$	Cl	0.08	−5.48
43[a]	A	SO$_2$	4-*t*-Bu	NH$_2$	Cl	0.13	−6.31
44[a]	A	SO$_2$	2,4-Me$_2$	NH$_2$	Cl	0.004	−6.42
45[a]	A	SO$_2$	3,5-Me$_2$	NH$_2$	Cl	0.004	−6.94
46[a]	A	SO$_2$	2,6-Cl$_2$	NH$_2$	Cl	0.1	−6.27
47[a]	A	SO$_2$	2-NH$_2$-5-Cl	NH$_2$	Cl	0.04	−4.45
48[a]	A	SO$_2$	3,5-Me$_2$	NH$_2$	Br	0.002	−5.11
49[a]	A	SO$_2$	3,5-Me$_2$	NH$_2$	COMe	0.015	−6.03
50[a]	A	SO$_2$	3,5-Me$_2$	NH$_2$	CH(OH)Me	0.025	−5.86
51[a]	A	S	H	NHNH$_2$	Cl	0.55	−5.74
52[a]	A	S	4-Me	NHNH$_2$	Cl	1.5	−7.47
53[a]	A	S	4-F	NHNH$_2$	Cl	5	−7.14
54[a]	A	S	4-Cl	NHNH$_2$	Cl	10	−6.35
55[a]	A	S	2-NH$_2$-5-Cl	NHNH$_2$	Cl	>13	−7.25
56[a]	A	SO$_2$	H	NHNH$_2$	H	0.53	−5.81
57[a]	A	SO$_2$	H	NHNH$_2$	Cl	0.01	−4.48
58[a]	A	SO$_2$	4-Me	NHNH$_2$	Cl	0.05	−7.17
59[a]	A	SO$_2$	4-F	NHNH$_2$	Cl	0.32	−7.12
60[a]	A	SO$_2$	4-Cl	NHNH$_2$	Cl	0.19	−6.72
61[a]	A	SO$_2$	3,5-Me$_2$	NHNH$_2$	Cl	0.13	−6.63
62[a]	A	SO$_2$	2-NH$_2$-5-Cl	NHNH$_2$	Cl	0.3	−6.75
63[a]	B	SO$_2$	H	H	H	>150	−6.24
64[a]	B	SO$_2$	4-Cl	COOEt	H	>100	−7.14
65[a]	B	SO$_2$	2-NO$_2$	COOEt	H	1.8	−6.24
66[a]	B	SO$_2$	2-NO$_2$-5-Cl	COOEt	H	>100	−4.38
67[a]	B	SO$_2$	2-NH$_2$-5-Cl	COOEt	H	1.8	−5.62
68[a]	B	SO$_2$	H	H	3-COOEt	>32	−5.89
69[a]	B	SO$_2$	H	H	3-COO-*i*-Pr	>54	−5.63
70[a]	B	SO$_2$	H	H	5-Cl	>75	−5.42
71[a]	B	SO$_2$	H	COOEt	5-Cl	>200	−5.54
72[a]	B	SO$_2$	4-Me	COOEt	5-Cl	>200	−5.61
73[a]	B	SO$_2$	4-Cl	COOEt	5-Cl	>200	−7.44
74[a]	B	SO$_2$	2-NO$_2$-5-Cl	COOEt	5-Cl	>31	−5.56
75[a]	B	SO$_2$	2-NH$_2$-5-Cl	COOEt	5-Cl	8.3	−7
76[a]	B	SO$_2$	H	H	5-Cl-3-COOEt	>42	−4.71
77[a]	B	SO$_2$	H	H	5-Cl-3-COO-*i*-Pr	>200	−4.08
78[a]	B	SO$_2$	H	CONH$_2$	H	15	−4.96
79[a]	B	SO$_2$	H	CONHNH$_2$	H	>200	−4.86

Table 2. (Continued)

compd	structure	X	Y	Z	W	EC$_{50}$ (μM)	ALOGpS (mg/L)
80[a]	B	SO$_2$	H	CONH$_2$	5-Cl	66.6	−5.52
81[a]	B	SO$_2$	H	H	3-CONH$_2$	>32	−4.68
82[a]	B	SO$_2$	H	H	5-Cl-CONH$_2$	≥200	−5.28
83[b]	B		2-NO$_2$			5.30	−4.66
84[b]	B		2-NH$_2$			4.96	−4.85
85[b]	B		2-NO$_2$-5-Cl			5.40	−5.48
86[b]	B		2-NH$_2$-5-Cl			6.00	−4.22
87[b]	B		2-NO$_2$-4-Cl			4.80	−6.92
88[b]	B		2-NH$_2$-4-Cl			4.13	−5.74
89[b]	B		2-Cl-5-NO$_2$			5.22	−5.66
90[b]	B		2-Cl-5-NH$_2$			3.82	−4.91
91[c]	A	S	H	NHCH$_2$CH$_2$OH	5-Cl	0.04	−5.29
92[c]	A	S	2-Me	NHCH$_2$CH$_2$OH	5-Cl	0.05	−4.44
93[c]	A	S	3-Me	NHCH$_2$CH$_2$OH	5-Cl	0.08	−5.97
94[c]	A	S	4-Me	NHCH$_2$CH$_2$OH	5-Cl	0.12	−5.48
95[c]	A	S	2,3-Me$_2$	NHCH$_2$CH$_2$OH	5-Cl	0.10	−5.34
96[c]	A	S	3,5-Me$_2$	NHCH$_2$CH$_2$OH	5-Cl	0.012	−4.44
97[c]	A	SO$_2$	H	NHCH$_2$CH$_2$OH	5-Cl	0.001	−4.47
98[c]	A	SO$_2$	2-Me	NHCH$_2$CH$_2$OH	5-Cl	0.031	−6.13
99[c]	A	SO$_2$	3-Me	NHCH$_2$CH$_2$OH	5-Cl	0.003	−4.24
100[c]	A	SO$_2$	4-Me	NHCH$_2$CH$_2$OH	5-Cl	0.043	−4.18
101[c]	A	SO$_2$	2,4-Me$_2$	NHCH$_2$CH$_2$OH	5-Cl	0.008	−5.72
4[c]	A	SO$_2$	3,5-Me$_2$	NHCH$_2$CH$_2$OH	5-Cl	0.0008	−5.52
102[c]	A	S	H	NHNHCH$_2$CH$_2$OH	5-Cl	0.2	−4.49
103[c]	A	S	2-Me	NHNHCH$_2$CH$_2$OH	5-Cl	0.1	−4.44
104[c]	A	S	4-Me	NHNHCH$_2$CH$_2$OH	5-Cl	0.6	−6.07
105[c]	A	S	2,4-Me$_2$	NHNHCH$_2$CH$_2$OH	5-Cl	1.2	−5.6
106[c]	A	S	3,5-Me$_2$	NHNHCH$_2$CH$_2$OH	5-Cl	0.03	−4.35
107[c]	A	SO$_2$	H	NHNHCH$_2$CH$_2$OH	5-Cl	0.01	−6.07
108[c]	A	SO$_2$	2-Me	NHNHCH$_2$CH$_2$OH	5-Cl	0.05	−4.32
109[c]	A	SO$_2$	3-Me	NHNHCH$_2$CH$_2$OH	5-Cl	0.007	−5.56
110[c]	A	SO$_2$	4-Me	NHNHCH$_2$CH$_2$OH	5-Cl	0.04	−5.91
111[c]	A	SO$_2$	2,4-Me$_2$	NHNHCH$_2$CH$_2$OH	5-Cl	0.08	−4.42
5[c]	A	SO$_2$	3,5-Me$_2$	NHNHCH$_2$CH$_2$OH	5-Cl	0.001	−5.31

Table 3. Statistical Values of the 3-D QSAR Model

				LOO		L5O		LHO	
FFD	Vars	PC	r^2	q^2	SDEP	q^2	SDEP	q^2	SDEP
0	3834	2	0.69	0.51	1.16	0.49	1.18	0.43	1.25
1	1690	2	0.77	0.66	0.96	0.65	0.97	0.60	1.05
2	1178	2	0.81	0.72	0.88	0.71	0.90	0.66	0.97
3	942	2	0.84	0.76	0.81	0.75	0.83	0.70	0.91
4	840	2	0.86	0.78	0.77	0.77	0.78	0.72	0.88
5	802	3	0.90	0.79	0.75	0.78	0.77	0.73	0.86
6	740	3	0.90	0.81	0.72	0.80	0.74	0.75	0.83
7	729	3	0.90	0.81	0.72	0.80	0.74	0.75	0.83

present fields plots for the subnanomolar active compound **4** (being the docked conformation fully superimposable on **4**, the nanomolar active derivative **5** could be used as well) and the submicromolar active IAS **94**.

The cyan polyhedrons (negative PLS coefficients) show areas where the activity increases after favorable interaction and decreases after unfavorable interaction. On the contrary, the yellow polyhedrons (positive PLS coefficients) show areas where the activity decreases after favorable interaction and increases after unfavorable interaction. Figure 2 shows that the cyan and yellow polyhedrons are properly arranged all around the structures of compounds **4** and **94** and they correctly describe the main features of the two IAS representative molecules. Focusing particularly on the molecular structures, the PLS-coefficient polyhedrons can be divided into three main groups: (i) a group surrounding the phenyl rings (top of Figure 2, blue circle); (ii) a group around the spatial area of the 5-chloroindole (left of Figure 2, magenta circle), and (iii) a group in the proximity of the amido/hydrazido substituent (right of Figure 2, green circle). Comparing the PLS-coefficient plot (Figure 2) with the present field plot (Figure 3A), the 3,5-dimethylbenzene of compound **4** makes favorable interactions with both methyl

Table 4. Test Set Structures, ALOGpS Values, and pEC$_{50}$ Prediction Using the Updated 3-D QSAR Model (new pred) and, for Comparison Purposes, the Prediction Using the Previous Model (old pred)

			pEC$_{50}$		
compd	Z	ALOGpS (mg/L)	expl	new pred	old pred
1	NHCH$_2$CONH$_2$	−4.72	8.222	9.093	7.584
2	NHCH$_2$CONHNH$_2$	−4.53	8.000	9.044	7.543
112	NHCH$_2$CONHCH$_2$CONH$_2$	−4.69	9.155	7.897	4.082
113	NHCH$_2$CONHCH$_2$CONHNH$_2$	−4.5	7.222	8.081	7.649
114	NHCH$_2$CONHCH(Me)CONHNH$_2$	−4.64	7.097	5.933	4.875
115	NHCH$_2$CONHCH(Me)CONH$_2$	−5.07	7.854	7.641	7.503
116	NHCH$_2$CONHCH$_2$CONHCH$_2$CONH$_2$	−4.69	6.921	7.346	6.388
117	NHCH$_2$CONHCH$_2$CONHCH$_2$CONHNH$_2$	−4.43	6.745	5.988	5.386

Table 5. Statistical Coefficients and Predictive Ability Comparison between Old and New 3-D QSAR Models

	PC	r^2	q^2_{L5O}	SDEP$_{CV-L5O}$	SDEP$_{TS}$[a]
old model[14]	2	0.85	0.68	0.95	2.05
new model	3	0.90	0.78	0.77	0.89

[a] The same test set reported in Table 4 was used for either models.

groups (yellow polyhedrons in Figure 2) and the benzene ring plane (cyan polyhedron in Figure 3A). On the other hand, the 4-methylphenyl ring of the weakly active compound **94** is slightly shifted and rotated (see Figure A of the Supporting Information) with respect to that of **4**. As a consequence, it makes unfavorable interactions in proximity to the unsubstituted
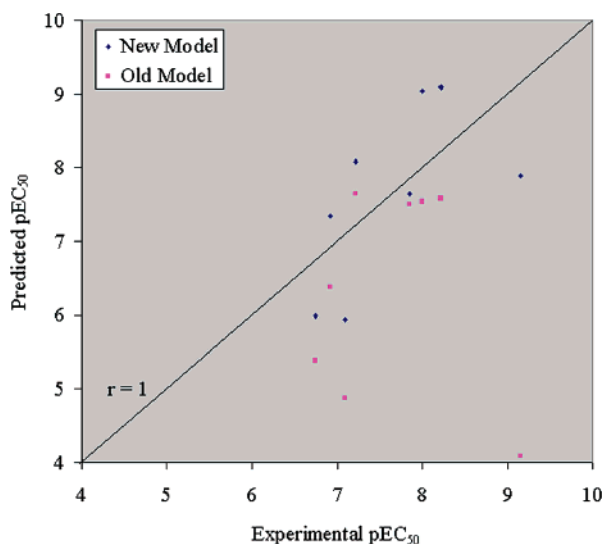
Figure 1. Plot of test set predictions as obtained by the new and the old 3-D QSAR models.
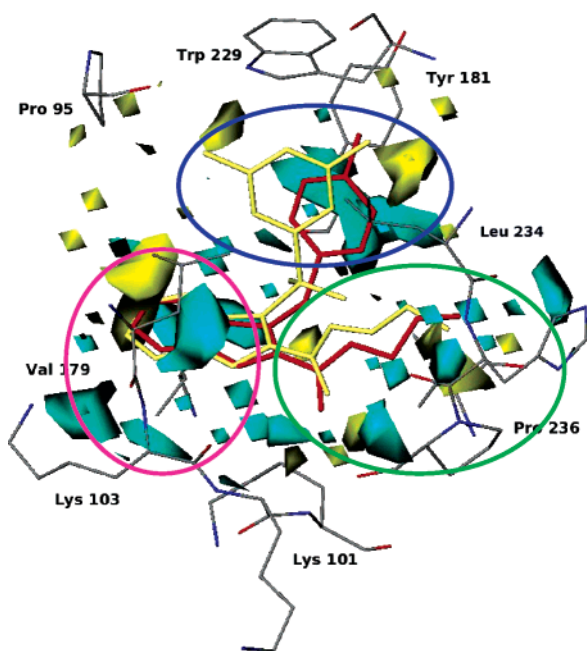

Figure 2. PLS−coefficient plot. Polyhedrons in yellow are negative coefficients (data level = −0.008); polyhedrons in cyan are positive coefficients (data level = +0.008). The highly active representative compound **4** (yellow), the less active compound **94** (red), and some relevant key residue are also reported for an easier interpretation (see the text).
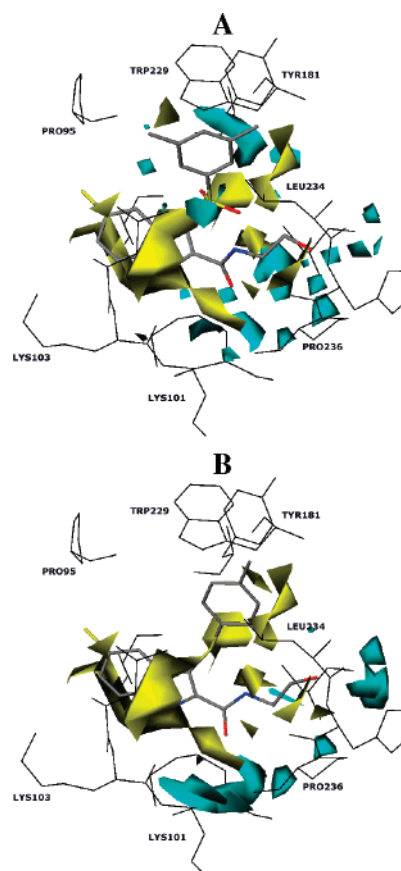

Figure 3. Present fields plots. Polyhedrons in yellow are negative fields (data level = −2.400); polyhedrons in cyan are positive fields (data level = +5.000). The highly active representative compound **4** is also reported in A, while the less active compound **94** is reported in B for an easier interpretation (see the text). The molecules are displayed by atom type.
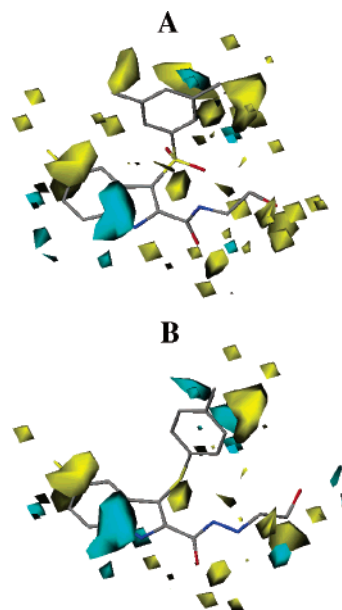
Figure 4. Activity contribution plots. Polyhedrons in yellow are negative contributions (data level = −0.035); polyhedrons in cyan are positive contributions (data level = +0.035). The highly active representative compound **4** is also reported in A, while the less active compound **94** is reported in B for an easier interpretation (see the text). The molecules are displayed by atom type.

benzene positions (see Figures 2 and 3B), thus highlighting the lack of adequate ligand/receptor contacts (compare right sides of Figures B-A and B-B of the Supporting Information). The activity contribution plots in Figure 4 confirmed the above observations. Three big yellow positive polyhedrons are visible around the 3,5-dimethylphenyl group (Supporting Information Figure C-A), while there is only one yellow polyhedron in compound **94** (Supporting Information Figure C-B). Regarding the 5-chloroindole moiety, no differences are observable between compounds **4** and **94**: the two indole rings are perfectly superimposable and this may be at least one of the reasons for the submicromolar activity of **94**. As a matter of fact, the PLS coefficient plots of the actual model substantially give the same graphical information as the previously reported model[14] (compare the left sides of A and B with Figure D in the

Supporting Information). In this area, the 5-chloroindole makes favorable interactions with both the chlorine atom and the indole
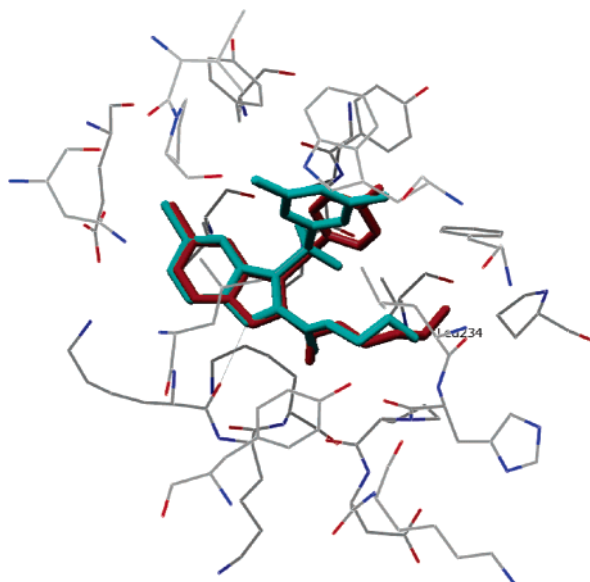
Figure 5. Binding mode of compounds **4** (cyan) and **94** (red).

ring (compare the PLS coefficient and activity contribution plots of Figures 2 and 4). Regarding the third group of polyhedrons (the green circle in Figure 2), the new model shows the greater differences once compared to the previous one[14] in this area (see Figures D−F of the Supporting Information). In the actual 3-D QSAR model, a cloud of cyan polyhedrons borders the hydroxyethyl groups of either **4** or **94**. Separate comparison of the PLS coefficient, the present fields, and the activity contribution plots for the two representative compounds makes it evident that the orientation of the hydroxyl group is of crucial importance. In fact, the hydroxyl group of the more active IAS **4**'s docked conformation makes a hydrogen bond with the highly conserved Leu234[20] (distance OH$_4$···O=C$_{Leu234}$ = 2.862 Å, Figure 5). The importance of the 4-OH orientation is correctly predicted by the PLS present fields as selected from the FFD procedure (see Figure F in the Supporting Information). On the contrary, the hydroxyl group of **94** does not make any hydrogen bond. In addition, the hydroxyl group also points toward Pro225 (not shown), making unfavorable interactions (see the small yellow polyhedron preceding the big cyan one on the right side of Figure F-B in the Supporting Information). Besides, a couple of small polyhedrons in the proximity of the sulfone bridge of only compound **4** confirmed our previous observations on the activity contribution plot (not shown).[14]

The New 3-D QSAR Model's Prediction Ability for Untested Compounds. The widening of the training set to 101 IAS structures led us to obtain a new 3-D QSAR model endowed with a highly internal predictive capability (see statistical values in Tables 3 and 4). To ascertain whether the new model would be powerful enough to predict new potent IAS derivatives, we designed a small series of compounds characterized by unreported structural modifications at the 2-hydrazido group of compound **5** (Chart 1). Apart from validating the model itself, the new compounds were also engineered to confirm the importance of the free hydroxyl groups of either compounds **4** or **5**. Inspection of the **4** and **5** binding modes (see above) indicated that the hydroxyethyl moiety lay in a hydrophobic pocket mainly formed by the side chains of Val106, Pro225, Phe227, Leu234, and Pro236 (Figure 5). This observation suggested the introduction of less hydrophilic substituents at the 2-hydrazido moiety of **5** (Table 1). Furthermore, taking into consideration the sterical interactions' contribution, the size of the substituents also varied from the
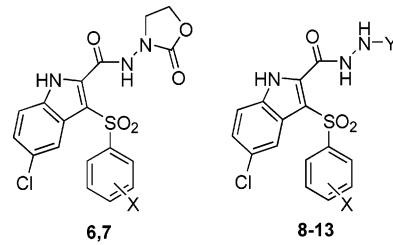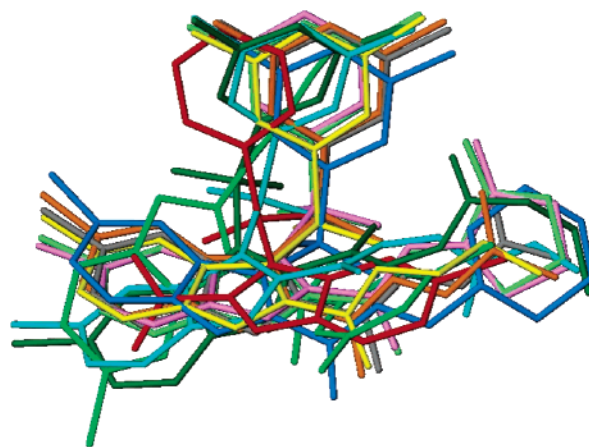

Figure 6. Docked conformations of the designed IASs **6**−**13** (**6**, violet; **7**, light green; **8**, green; **9**, orange; **10**, blue; **11**, gray; **12**, dark green; **13**, cyan). For comparison purposes, the conformation of reference compounds **3** and **4** are also reported (**3**, red; **4**, yellow).

small methyl to the bulky cyclohexyl groups. The designed compounds were structure-based aligned by docking, using the same protocol described for the training set (see the Experimental Section). Their antiviral activities were predicted using the above-described 3-D QSAR model. Autodock binding modes can give some insight on the potential anti-HIV-1 activities of the compounds **6**−**13**, prior to any quantitative model application inspections. Surprisingly, compound **8**, although having a similar active conformation, showed an alternative binding mode from those obtained for either L-737,126 (**3**) or **4**.[14] Compared to **4**, the indole ring in compound **8** was rotated by about 180° along a virtual axis (not shown) that expands from positions 6 to 3 of the indole (Figure 6). This different binding mode could lead to the loss of important favorable interactions, thus causing a lower anti-HIV-1 activity than those of the reference compounds **3** and **4**. As expected, the other seven designed IASs displayed a binding mode comparable to that of **4** (Figure 6), showing substantial differences in the mode of interaction with introduced modifications. Because of the unsubstituted benzene ring,[21] derivative **6** might be less active than **4** (the oxazolidinone ring seemed to fulfill the hydrophobic pocket above-described, but no particular interactions were visible). The 3,5-dimethyl derivative **7** roughly assumed the same binding mode as **6**; thus, this compound may exert some higher activity than **6**. The isopropyl, acetyl, and hydrazinocarbonyl groups of derivatives **9**, **11**, and **13**, respectively, seemed to optimally fulfill the pocket formed by Val106, Pro225, Phe227, Leu234, and Pro236 (not shown) and a higher activity may be expected. Regarding the designed compounds **10** and **12**, the respective cyclohexyl and carboxyethyl bulky groups made steric unfavorable interactions with Pro225 (not shown). In particular to accommodate the COOEt group, **12**'s indole ring is shifted away, thus causing the disruption of some favorable interactions, such as the π−π stacking interaction of the 3,5-dimethylbenzene with Tyr181 and Tyr188 (not shown) and the hydrogen bond between the indole NH and Lys101 (not shown). A similar shift, but to a lesser extent, was also observed for compound **10**.

The above qualitative interpretation of the binding modes of compounds **6**−**13** was quantitatively confirmed by the application of the final 3-D QSAR model. These compounds were predicted to have an average anti-HIV-1 activity of about 60 nM and thus deserved to be evaluated. In fact, after synthesis and assaying, the newly designed IASs proved to be effectively and highly active against acutely infected HIV-1 MT4 cells (Table 6). The new derivatives were more active than originally

Table 6. Structures, ALOGpS Values, and Predicted and Experimental pEC$_{50}$s of Designed Compounds **6–13**

| compd	X	Y	ALOGpS (mg/L)	pEC$_{50}$ pred	pEC$_{50}$ expl
6	H	–	−4.45	7.446	7.824
7	3.5-Me$_2$	–	−4.51	7.442	9.046
8	3.5-Me$_2$	Me	−4.97	6.540	7.699
9	3.5-Me$_2$	i-Pr	−4.31	8.455	9.155
10	3.5-Me$_2$	c-He	−5.29	6.578	7.301
11	3.5-Me$_2$	COMe	−4.79	8.139	8.301
12	3.5-Me$_2$	COOEt	−4.55	6.234	7.699
13	3.5-Me$_2$	CONHNH$_2$	−4.77	7.134	8.000

predicted, with an average pEC$_{50}$ of 8.12 (EC$_{50}$ about 10 nM). Among them, two compounds (**7** and **9**) showed sub-nanomolar anti-HIV-1 activity (see the SAR section).

In Silico Evaluation of Pharmacokinetic Properties. Along with the docking and 3-D QSAR studies, the aqueous solubility of the training set, test set, and prediction set was evaluated using the stand alone version of the ALOGPS 2.1 program developed by Tetko et al.[22] Analyzing the solubility data, it was interesting to note an increase of the compounds' calculated water solubility (ALOGpS values in Tables 2, 4, and 6 correspond to higher pEC$_{50}$ values). This is particularly evident in comparing similar compounds; for instance, a definite increase of aqueous solubility was observable between sulfur and sulfone derivatives for the diverse IAS series to which a definite augment of the pEC$_{50}$ values corresponded. In Table 2, this was seen by comparing the carboxamide ALOGpS values of sulfur compounds **25**, **26**, **27**, and **32** with those of sulfones **3**, **37**, **39**, and **45**; the hydroxyethylcarboxamide ALOGpS values of derivatives **91**, **92**, **94**, and **96** with those of **97**, **98**, **100**, and **4**; and the hydroxyethylcarbohydrazide ALOGpS values of **102**, **103**, **104**, and **106** with those of **107**, **108**, **110**, and **5**. It is also interesting to note that a general aqueous solubility enhancement was observable starting from carboxamide derivatives going to hydroxyethylcarboxamides and hydroxyethylcarbohydrazides (compare ALOGpS values of carboxamides **3**, **37**, **39**, and **45** with those of hydroxyethylcarboxamides **97**, **98**, **100**, and **4** and those of hydroxyethylcarbohydrazides **107**, **108**, **110**, and **5**). Nevertheless, the average pEC$_{50}$ values were comparable for these three series. These data suggested that to obtain highly active NNRTI compounds a better water solubility profile is necessary. Even so, some other parameters are very likely involved in the overall inhibition equilibrium, such as cell permeability. ALOGpS values were also calculated for the newly designed compounds **6–13**. Interestingly, the aqueous solubility values of the new IASs were comparable with those of the most active IASs listed in Tables 2 and 4 (compare the ALOGpS values of **3**, **4**, **5**, **37**, and **97** with those of **6–13**). These results fully agreed with the above-reported quantitative prediction and also supported the synthesis and biological evaluation of the designed derivatives.

Chemistry. Reaction of 3-(phenylsulfonyl)-5-chloro-1*H*-indole-2-carboxylic acid with 3-amino-2-oxazolidinone in the presence of EDCI and DMAP at room temperature afforded *N*-(2-oxazolidin-3-yl)-3-(phenylsulfonyl)-5-chloro-1*H*-indole-2-

Scheme 1[a]

6 X = H, **7** X = 3,5-Me$_2$
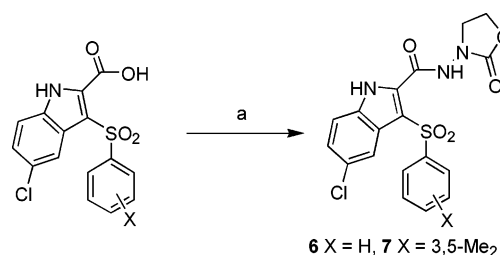
[a] Reagents and reaction conditions: (a) for **6**, 3-amino-2-oxazolidinone, EDCI, DMAP, THF, rt, 96 h; for **7**, 3-amino-2-oxazolidinone, BOP, Et$_3$N, DMF, rt, 48 h.

Scheme 2[a]

X = OMe (a), NHNH$_2$ (b,c)

8 Y = Me, **9** Y = *i*-Pr,
10 Y = *c*-He, **11** Y = COMe

12

13
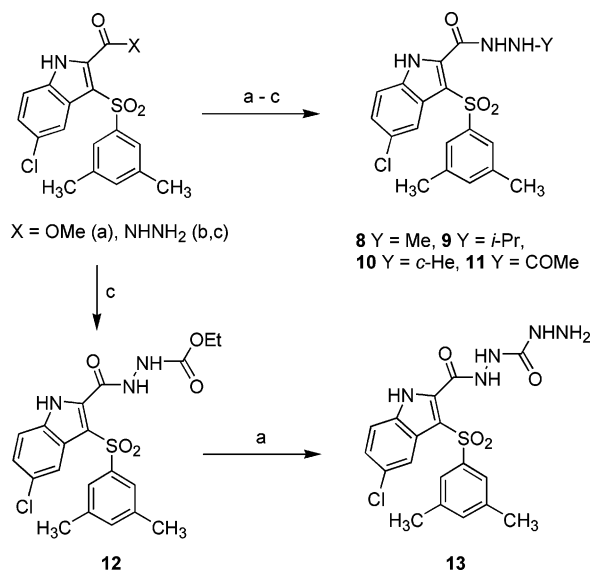
[a] Reagents and reaction conditions. (a) NHNHMe (**8**) or NH$_2$NH$_2$ (**13**), EtOH, rt overnight; (b) acetone (**9**) or cyclohexanone (**10**), NaBH$_3$CN, THF–MeOH–HCl. 0 °C to rt overnight; (c) acetyl (**11**) or ethyl chlorocarbonate (**12**), rt, overnight.

carboxamide (**6**) (Scheme 1). *N*-(2-Oxazolidinone-3-yl)-3-[(3.5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxamide (**7**) was prepared by reacting the corresponding acid with 3-amino-2-oxazolidinone in the presence of BOP reagent and triethylamine. *N*′-Methyl-3-[(3.5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxyhydrazide (**8**) was obtained by reacting methyl 3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxylate with methylhydrazine at room temperature (Scheme 2).

Reaction of 3-[(3.5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxyhydrazide with acetone or cyclohexanone in the presence of NaBH$_3$CN in acidic medium at room temperature afforded derivatives *N*′-isopropyl- (**9**) or *N*′-cyclohexyl-3-[(3.5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carbohydrazide (**10**), respectively. Treatment of 3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carbohydrazide with acetyl chloride or ethyl chlorocarbonate in anhydrous THF at room temperature furnished *N*′-acetyl-3-[(3.5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carbohydrazide (**11**) or *N*′-ethoxycarbonyl-3-[(3.5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carbohydrazide (**12**), respectively. Transformation of **12** to *N*′-hydrazinocarbonyl-3-[(3.5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carbohydrazide (**13**) was performed by treating with hydrazine hydrate at room temperature.

Structure–Activity Relationship Studies of the New designed IAS. Cytotoxicity and anti-retroviral activity of derivatives **6–13** were evaluated against HIV-1 WT and NNRTI

Table 7. Cytotoxicities and Anti-HIV-1 Activities of Derivatives 6−13 and Reference Compounds 1−5 in MT-4 and C8166 Cells

compd	MT-4			C8166		
	TC_{50}^{b}	EC_{50}^{c}	SI^{d}	TC_{50}^{b}	EC_{50}^{c}	SI^{d}
6	>100	0.015	>6666	>100	0.02	>5000
7	100	0.0009	111111	>100	0.003	>33333
8	>100	0.02	>5000	>100	0.03	>3333
9	>100	0.0007	>142857	>100	0.003	>33333
10	>100	0.05	>2000	>100	0.08	>1250
11	>100	0.005	>20000	>100	0.007	>14285
12	100	0.02	5000	>100	0.05	>2000
13	>100	0.01	>10000	>100	0.008	>12500
1	45	0.01	4500	>100	0.02	>5000
2	>100	0.02	>5000	>100	0.02	>5000
3[e]	84	0.002	42000	>100	0.003	>33333
4[e]	85	0.0008	106250	>100	0.002	>50000
5[e]	30	0.001	>30000	>100	0.004	>25000

[a] Data are mean values of two experiments performed in triplicate. [b] Cytotoxic concentration (μM) required to reduce the viability (MT-4) or to inhibit the proliferation (C8166) of mock-infected cells by 50% as monitored by the MTT method. [c] Compound concentration (μM) required to inhibit virus induced cell death by 50% as monitored by the MTT method. [d] Selectivity index: TC_{50}/EC_{50} ratio. [e] Data taken from ref 12.

Table 8. Activity of Derivatives 6−13 and Reference Compounds 1−5, Nevirapine, and Efavirenz against HIV-1 WT (IIIB) and the Y181C and K103N-Y181C Resistant Strains

compd	EC_{50}^{b}		
	WT (IIIB)	Y181C	K103N-Y181C
6	0.015	0.3	>20
7	0.001	0.01	0.9
8	0.05	0.13	>20
9	0.0007	0.04	8.6
10	0.1	1.2	>20
11	0.005	0.02	>20
12	0.02	0.04	9.2
13	0.04	0.02	5.7
1	0.01	0.05	1.2
2	0.02	0.06	2.6
3[c]	0.002	0.03	10
4[c]	0.0008	0.009	0.3
5[c]	0.001	0.04	3.2
NVP[c]	0.05	14	>20
EFV[c]	0.003	0.01	0.2

[a] Data are mean values of two experiments performed in triplicate. [b] Compound concentration (μM) required to achieve 50% protection of MT-4 cells from HIV-1 wild-type IIIB. Y181C and K103N-Y181C resistant strains cytopathogenicity as monitored by MTT method. [c] Data taken from ref 12.

resistant strains. Compounds **1−5**, nevirapine, and efavirenz were used as reference compounds. The activity of tested compounds against HIV-1 WT (EC_{50} values) was determined in MT-4 and C8166 cells by means of MTT assay (Table 7). The activity against the resistant mutant strains Y181C and K103N-Y181C (Table 8) was determined in MT-4 cells (MTT assay, EC_{50} values). The activity against the two primary isolates HIV-112 and HIV-AB1 highly resistant to HAART carrying K103N-V108I-M184V and L100I-V108I mutations, respectively, was carried out in lymphocytes by means of p24 assay, in parallel with the WT strain (EC_{50} values) (Table 4). The activity in macrophages infected by HIV-1 WT IIIB$_{Ba-L}$ was measured by means of a p24 assay (EC_{50} values) (Table 5).

IAS derivatives **6−13** were highly active at nanomolar concentration and selective against the HIV-1 WT strain. Three compounds (**7**, **9**, and **11**) inhibited the HIV-1 WT multiplication in low nanomolar range of concentration both in MT-4 and C8166 cells. In MT-4 cells the EC_{50} values ranged from 0.05 μM (**10**) to 0.0007 μM (**7**), and the selectivity indexes (SIs) from >5000 (**10**) to >142 857 (**9**). In C8166 cells the EC_{50}

Table 9. Activity of Derivatives 6−13 and Reference Compounds 4 and 5 against the HIV-1 WT (IIIB) and the Primary Isolates HIV-112 and HIV-AB1 in Lymphocytes[a]

compd	TC_{50}	EC_{50}^{b}			SI^{c}		
		IIIB	112	AB1	IIIB	112	AB1
6	>100	0.1	0.4	0.5	>1000	>250	>200
7	>100	0.01	0.03	0.02	>10000	>3333	>5000
8	>100	0.02	0.04	0.02	>5000	>2500	>5000
9	>100	0.008	0.006	0.008	>12500	>16666	>12500
10	>100	0.8	1.0	0.9	>125	>100	>111
11	>100	0.01	0.008	0.01	>10000	>12500	>10000
12	>100	0.02	0.03	0.02	>5000	>3333	>5000
13	>100	0.03	0.05	0.04	>3333	>2000	>2500
4[d]	>100	0.02	0.05	0.01	>5000	>2000	>10000
5[d]	>100	0.06	0.055	0.070	>1666	>1818	>1428
NVP[e]	>100	0.09	>20	>20	>1111	−	−
EF^{Vf}	>100	0.007	>20	>20	>14285	−	−

[a] Data are mean values of two experiments performed in triplicate. [b] Compound concentration (μM) required to reduce the amount of p24 by 50% in the indicated strain. HIV-p24 antigen production in control HIV-1-infected lymphocytes was 118 900 pg/mL and 247 200 pg/mL in HIV-112- and HIV-AB1-infected lymphocytes, respectively. [c] Selectivity index: TC_{50}/EC_{50} ratio. [d] Data taken from ref 12. [e] NVP, nevirapine. [f] EFV, efavirenz.

values ranged from 0.08 μM (**10**) to 0.003 μM (**7** and **9**), and the SIs from >1250 (**10**) to >33 333 (**7** and **9**). Noteworthy, all tested compounds were not found to be cytotoxic up to 100 μM.

Cyclization of the 2-hydroxyethylamino terminus of **5** produced *N*-(2-oxazolidone-3-yl)-3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxyamide (**7**) that retained substantially the antiviral activity of the parent compound. Due to its scarce cytotoxicity (CC_{50} = 100 μM) compound **7** showed SI values superior to those of the reference compounds **1−5** in MT-4 cells.

Similarly, replacement of the 2-hydroxyethyl chain of **5** with isopropyl (**9**) or acetyl (**11**) moieties led to noncytotoxic compounds active in the low nanomolar range of concentration. Other substitutions (**6**, **8**, **10**, **12**, and **13**) caused reductions in antiviral potency, but interestingly, these compounds were found deprived of cytotoxicity.

Against the Y181C mutant strain derivatives **7**, **9**, and **11−13** showed inhibitory activities EC_{50}s comparable to those of the reference compounds **1−5** and efavirenz. In particular, *N*-(2-oxazolidinone-3-yl)-3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxamide (**7**, EC_{50} = 0.01 μM), the most active derivative against the Y181C strain, was 4 times more potent than the parent compounds **5** and as potent as efavirenz. Against the K103N-Y181C mutant strain compounds **7**, **9**, **12**, and **13** showed interesting activities with EC_{50}s in the micromolar range. Compound **7** (EC_{50} = 0.9 μM) was 3.5 times superior to **5** but, respectively, 3 and 4.5 times less potent than **4** and efavirenz.

The most potent derivatives were also evaluated in lymphocytes against HIV-1 WT (IIIB) and the primary isolates HIV-112 and HIV-AB1 obtained from two HIV-1-Ab seropositive individuals who had experienced a therapeutic failure after treatment with HAART regimens (Table 9). The 112 strain was selected after treatment with NRTIs and NNRTIs and carried mutations at K103N-V108I-M184V positions. The AB1 strain was selected after treatment with NRTIs, NNRTIs, and at least one protease inhibitor and carried L100I-V108I mutations.

Tested compounds inhibited the 112 strain with concentrations ranging from 1.0 μM (**10**) to 0.006 μM (**9**) with SI values from >100 (**10**) to >16 666 (**9**) and inhibited the AB1 strain with concentrations ranging from 0.9 μM (**10**) to 0.08 μM (**9**) with

Table 10. Activity of Derivatives 6–13 and Refence Compounds 4 and 5 against the HIV-1 WT (IIIBBa-L) in Macrophages[a]

compd	TC_{50}	EC_{50}[b]	SI[c]	compd	TC_{50}	EC_{50}[b]	SI[c]
6	>100	0.06	>1666	**11**	>100	0.008	>12500
7	>100	0.002	>50000	**12**	>100	0.006	>16666
8	>100	0.03	>3333	**13**	>100	0.009	>11111
9	>100	0.005	>20000	**4**[d]	>100	0.01	>10000
10	>100	0.2	>500	**5**[d]	>100	0.02	>5000

[a] Data are mean values of two experiments performed in triplicate. [b] Compound concentration (μM) required to reduce the amount of p24 by 50% in the HIV-1 (IIIB$_{Ba-L}$) strain. HIV-p24 antigen production in control HIV-1-infected macrophages was 83 400 pg/mL. [c] Selectivity index: TC_{50}/EC_{50} ratio. [d] Data taken from ref 15.

Table 11. HIV-1 RT Inhibitory Activity of Compounds 7 and 9 against Wild-Type and Mutant Enzymes Carrying Single Amino Acid Substitutions[a]

compd	EC_{50}			ID_{50}[b]	
	WT	WT	K103N	Y181I	L100I
7	0.001	0.25	0.4	4	0.17
9	0.0007	0.01	0.8	8.2	2.2
NVP	0.37	0.4	7	36	9
EFV	0.004	0.08	25	0.4	nd[d]

[a] Data represent mean values of at least three separate experiments. [b] In MT-4 cells. [c] Compound dose (K_i, μM) required to inhibit by 50% the RT activity of the indicated strain. [d] nd, no data.

SI values from >111 (**10**) to >12 500 (**9**). The isopropyl carbohydrazide **9** was the most active against both 112 and AB1 strains.

In macrophages, four out of seven compounds (**7** and **11**–**13**) inhibited the multiplication of IIIB$_{Ba-L}$ strain at nanomolar concentration (Table 10). Compounds **7** (EC_{50} = 0.002 μM) and **9** (EC_{50} = 0.005 μM) were the most active and showed also the highest SI values (SI = >10 000 and >5000. respectively).

The most potent derivatives **7** and **9** showed a significant activity against RT WT and NNRTIs-resistant mutants. In particular, they were significantly superior to nevirapine in absolute terms against the Y181I mutant (9.0 and 4.4 times, respectively), the K103N mutant (17.5 and 8.75 times), and the L100I mutant (52.9 and 4.1 times), respectively. Even more interesting was their resistance profile with respect to the K103N mutant, when compared to the second-generation NNRTI efavirenz. The relative resistance values (ID_{50}mut/ID_{50}wt) of the K103N mutant toward **7** and **9** were 1.6- and 80-fold, respectively, whereas the same mutation showed a relative resistance of >300-fold toward efavirenz. In addition, **7** and **9** showed inhibition potencies of virus proliferation (EC_{50} values) even superior to that of efavirenz, confirming their potent anti-HIV activity.

Conclusions

In this study we describe the update of a previous 3-D QSAR model. The increase of the training set composition from 70 to 101 molecules expanded the model's application to a wider range of structures. This led to a higher predictive model and the validation against an external test set of recently published hybrid IAS−peptide derivatives (**1**, **2**, **112**−**117**). It also showed a SDEP value of only 0.89, which was comparable to the cross-validation standard deviation errors of prediction (Tables 3 and 4). With the aim to both validate the new 3-D QSAR model and design new anti-HIV-1 compounds, we engineered a prediction set of eight IASs by simply modifying the substituents at the 2-hydrazide position of the reference compound **5**. The combination of the structure-based alignment obtained by

docking and the 3-D QSAR model quantitative estimation of the pEC_{50} values proved to be highly effective and allowed the disclosure of new potent IAS derivatives.

Against the HIV-1 WT strain, compounds **7** and **9** were exceptionally potent in the sub-nanomolar range of concentration in both MT-4 and C8166 cells. These compounds, in particular compound **9**, also showed excellent inhibitory activity against the HIV-1 WT (IIIB) and the primary isolates HIV-112 and HIV-1-AB1 in limphocytes and against the HIV-1 WT (IIIB$_{Ba-L}$) in macrophages. Efavirenz is a highly effective NNRTI used as a first line drug in combination with two NRTIs in protease-sparing regimen. However, the K103N mutation is selected in the majority of patients (>90%) whose viral loads rebounded after an initial response to the drug. Furthermore, additional double mutations such as K103N-V108N and K103N-P225H slowly emerge in many patients.[23] Compounds **7** and **9** were found to be more active than efavirenz against the viral RT carrying the K103N mutation. Inspection of the binding modes suggested that the indolyl aryl sulfone scaffold has a high adaptability to the non-nucleoside binding site that makes many of the IASs capable of inhibiting not only the WT form of RT but also some of the most important mutant isoforms.

Further molecular modeling, docking simulation, 3-D QSAR studies, synthesis efforts, and biological characterization are in progress to search for other structural features that can help to modulate the anti HIV-1 activity against both WT and mutant strains.

Experimental Section

Chemistry. Melting points (mp) were determined on a Büchi 510 apparatus and are uncorrected. Infrared spectra (IR) were run on Perkin-Elmer 1310 and SpectrumOne spectrophotometers. Band position and absorption ranges are given in cm^{-1}. Proton nuclear magnetic resonance (^{1}H NMR) spectra were recorded on Bruker AM-200 (200 MHz) and Bruker Avance 400 MHz FT spectrometers in the indicated solvent. Chemical shifts are expressed in δ units (ppm) from tetramethylsilane. Column chromatographies were packed with alumina (Merck, 70−230 mesh) and silica gel (Merck, 70−230 mesh). Aluminum oxide TLC cards (Fluka, aluminum oxide precoated aluminum cards with fluorescent indicator at 254 nm) and silica gel TLC cards (Fluka, silica gel precoated aluminum cards with fluorescent indicator at 254 nm) were used for thin-layer chromatography (TLC). Developed plates were visualized with a Spectroline ENF 260C/F UV apparatus. Organic solutions were dried over anhydrous sodium sulfate. Concentration and evaporation of the solvent after reaction or extraction was carried out on a rotary evaporator (Büchi Rotavapor) operating at reduced pressure. Elemental analyses were found within ±0.4% of the theoretical values. 3-(Phenylsulfonyl)-5-chloro-1*H*-indole-2-carboxylic acid, 3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxylic acid, methyl 3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxylate, and 3-[(3.5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carbohydrazide were reported by us previously.[9]

***N*-(2-Oxazolidinone-3-yl)-3-(phenylsulfonyl)-5-chloro-1*H*-indole-2-carboxamide (6, RS 1580).** 3-Amino-2-oxazolidinone (0.17 g, 0.0016 mol), *N*-ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDCI, 0.31 g, 0.0016 mol), and 4-(dimethylamino)-pyridine (DMAP, 0.2 g, 0.0016 mol) were added in sequence to a solution of 5-chloro-3-(phenylsulfonyl)-1*H*-indole-2-carboxylic acid (0.55 g, 0.0016 mol) in anhydrous THF (40 mL). The reaction was stirred at room temperature for 96 h, concentrated to a small volume, and extracted with ethyl acetate. The organic layer was washed with water and dried. Removal of the solvent gave a residue which was purified on silica gel column (ethyl acetate as eluent) to afford **6** (0.27 g, 40%). Mp: >250 °C (from toluene/*n*-hexane). ^{1}H NMR (DMSO-*d₆*): δ 3.93 (t, *J* = 7.6 Hz, 2H), 4.56 (t, *J* = 7.6 Hz, 2H), 7.44 (dd, *J* = 1.9 and 8.7 Hz, 1H), 7.58−7.75 (m, 4H), 8.00 (d, *J* = 1.9 Hz, 1H), 8.12−8.22 (m, 2H), 11.28 (br s, 1H, disappeared

on treatment with D_2O), 13.47 (br s, 1H, disappeared on treatment with D_2O). IR (Nujol): ν 1650, 1770, 3240 cm^{-1}. Anal. Calcd ($C_{18}H_{14}ClN_3O_5S$ (419.83)) C, H, Cl, N, S.

***N*-(2-Oxazolidinone-3-yl)-3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxamide (7, RS 1893).** 3-Amino-2-oxazolidinone (0.30 g, 0.0028 mol), triethylamine (0.42 g, 0.58 mL, 0.0041 mol), and BOP ragent (0.61 g, 0.0014 mol) were added in sequence to a solution of 5-chloro-3-[(3.5-dimethylphenyl)sulfonyl]-1*H*-indole-2-carboxylic acid (0.5 g, 0.0014 mol) in anhydrous DMF (25 mL). The reaction mixture was stirred at room temperature for 48 h. Water was added and then the suspension was filtered to give **7** (0.31 g, 50%). Mp: 282−284 °C (dec). ^{1}H NMR (DMSO-d_6): δ 2.26 (s, 6H), 3.80 (m, 2H), 4.43 (m, 2H), 7.20 (s, 1H), 7.32 (m, 1H), 7.48 (d, J = 8.7 Hz, 1H), 7.64 (s, 2H), 7.86 (d, J = 2.1 Hz, 1H), 11.10 (br s, 1H, disappeared on treatment with D_2O). IR (Nujol): ν 1640, 1770, 3230 cm^{-1}. Anal. Calcd ($C_{20}H_{18}ClN_3O_5S$ (447.89)) C, H, Cl, N, S.

***N*′-Methyl-3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxyhydrazide (8, RS 1787).** A mixture of methyl 5-chloro-3-[(3,5-dimethylphenyl)sulfonyl]-1*H*-indole-2-carboxylate (0.70 g, 0.0019 mol) and methylhydrazine (4.88 g, 5.58 mL, 0.10 mol) in ethanol (96 °C, 5.58 mL) was stirred at room temperature overnight. The reaction mixture was poured on ice water and then the suspension was filtered to afford **8** (0.12 g, 17%). Mp: 284−287 °C (from aqueous DMF). ^{1}H NMR (DMSO-d_6): δ 2.33 (s, 6H), 3.98 (s, 3H), 7.28 (s, 1H), 7.44 (dd, J = 1.6 and 8.8 Hz, 1H), 7.55−7.67 (m, 3H), 8.25 (d, J = 1.6 Hz, 1H), 13.28 (br s, 1H, disappeared on treatment with D_2O). IR (Nujol): ν 1645, 3150, 3250 cm^{-1}. Anal. Calcd ($C_{18}H_{18}ClN_3O_3S$ (391.87)) C, H, Cl, N, S.

***N*′-Isopropyl-3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxyhydrazide (9, RS 1833).** Sodium cyanoborohydride (0.196 g, 0.0031 mol) was added to an ice-cooled mixture of acetone (0.15 g, 0.19 mL, 0.0026 mol) and 5-chloro-3-[(3,5-dimethylphenyl)sulfonyl]-1*H*-indole-2-carbohydrazide (0.98 g, 0.0026 mol) in methanol/THF (1:1, 32.5 mL) containing 6 N HCl/methanol (1: 1. 0.44 mL). The reaction was stirred at room temperature overnight. The mixture was reduced to a small volume, diluted with water, and extracted with ethyl acetate. The organic layer was washed with brine and dried. Removal of the solvent gave a crude product which was purified on silica gel column (chloroform/ethanol 95:5) to give **9** (0.69 g, 63%). Mp: 248−250 °C (from ethanol). ^{1}H NMR (DMSO-d_6): δ 1.09 (d, J = 6.1 Hz, 6H), 2.34 (s, 6H), 3.15 (m, 1H), 5.19 (br s, 1H, disappeared on treatment with D_2O), 7.25 (s, 1H), 7.33 (dd, J = 2.0 and 8.7 Hz, 1H), 7.53 (d, J = 8.7 Hz, 1H), 7.67 (s, 2H), 7.89 (d, J = 2.0 Hz, 1H), 10.25 (br s, 1H, disappeared on treatment with D_2O), 13.05 (br s, 1H, disappeared on treatment with D_2O). IR (Nujol): ν 1620, 3230, 3300 cm^{-1}. Anal. Calcd ($C_{20}H_{22}ClN_3O_3S$ (419.96)) C, H, Cl, N, S.

***N*′-Cyclohexyl-3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxyhydrazide (10, RS 1967)** was synthesized as **9** using cyclohexanone. Yield: 99%. Mp: 248−250 °C (from aqueous DMF). ^{1}H NMR (DMSO-d_6): δ 0.95−2.00 (m, 10H), 2.31 (s, 6H), 2.73 (m, 1H), 4.53 (br s, 1H, disappeared on treatment with D_2O), 7.24 (s, 1H), 7.33 (s, 1H), 7.53 (d, J = 9.0 Hz, 1H), 7.66 (s, 2H), 7.90 (s, 1H), 10.26 (br s, 1H, disappeared on treatment with D_2O), 13.00 ppm (br s, 1H, disappeared on treatment with D_2O). IR (Nujol): ν 1645, 3300 cm^{-1}. Anal. Calcd ($C_{23}H_{26}ClN_3O_3S$ (459.52)) C, H, Cl, N, S.

***N*′-Acetyl-3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carboxyhydrazide (11, RS 2042).** Acetyl chloride (0.04 g, 0.04 mL, 0.0005 mol) in anhydrous THF (1 mL) was added dropwise to a suspension of 5-chloro-3-[(3,5-dimethylphenyl)sulfonyl]-1*H*-indole-2-carbohydrazide (0.2 g, 0.0005 mol) in the same solvent (7.5 mL) while cooling at 0 °C. The reaction mixture was stirred at room temperature overnight. After evaporation of the solvent, water was added and then the suspension was filtered to give **11** (0.20 g, 100%). Mp: >300 °C (aqueous DMF). ^{1}H NMR (DMSO-d_6): δ 1.96 (s, 3H), 2.27 (s, 6H), 7.21 (s, 1H), 7.32 (dd, J = 1.9 and 8.7 Hz, 1H), 7.52 (d, J = 8.7 Hz, 1H), 7.73 (s, 2H), 7.94 (s,

1H), 10.40 (br s, 1H, disappeared on treatment with D_2O), 10.82 (br s, 1H, disappeared on treatment with D_2O), 13.10 (br s, 1H, disappeared on treatment with D_2O). Anal. Calcd ($C_{19}H_{18}ClN_3O_4S$ (419.07)) C, H, Cl, N, S.

***N*′-Ethoxycarbonyl-3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carbohydrazide (12, RS 2033)** was synthesized as **11** using ethylchloroformate. Yield: 93%. Mp: 255−257 °C (from DMF/H_2O). ^{1}H NMR (DMSO-d_6): δ 1.21 (t, J = 7.1 Hz, 3H), 1.28 (s, 6H), 4.11 (q, J = 7.1 Hz, 2H), 7.21 (s, 1H), 7.33 (dd, J = 1.9 and 8.8 Hz, 1H), 7.52 (d, J = 8.8 Hz, 1H), 7.71 (s, 2H), 7.95 (m, 1H), 9.70 (br s, 1H, disappeared on treatment with D_2O), 10.65 (br s, 1H, disappeared on treatment with D_2O), 13.10 (br s, 1H, disappeared on treatment with D_2O). IR (Nujol): ν 1655, 1716, 3250 cm^{-1}. Anal. Calcd ($C_{20}H_{20}ClN_3O_5S$ (449.91)) C, H, Cl, N, S.

***N*′-Hydrazinocarbonyl-3-[(3,5-dimethylphenyl)sulfonyl]-5-chloro-1*H*-indole-2-carbohydrazide (13, RS 2034)** was synthesized as **8** starting from hydrate hydrazine. Yield: 63%. Mp: 280−284 °C (from aqueous DMF). ^{1}H NMR (DMSO-d_6): δ 2.28 (s, 6H), 4.78 (br s, 1H, disappeared on treatment with D_2O). 7.32 (s, 1H), 7.29 (dd, J = 1.9 and 8.8 Hz, 1H), 7.23 (d, J = 8.8 Hz, 1H), 7.34 (s, 2H), 7.87 (d, J = 1.9 Hz, 1H), 7.91 (br s, 1H, disappeared on treatment with D_2O). 10.00 (br s, 1H, disappeared on treatment with D_2O), 13.12 (br s, 1H, disappeared on treatment with D_2O). IR (Nujol): ν 1652, 3214, 3306 cm^{-1}. Anal. Calcd ($C_{18}H_{18}ClN_5O_4S$ (435.88)) C, H, Cl, N, S.

Cell-Based Antiviral Assay Procedures. Cells. C8166 is a CD4+ T-cell line containing an HTLV-1 genome of which only the *tat* gene is expressed.[24] Normal human lymphocytes and macrophages were obtained as follows. Peripheral blood was obtained as follows. Peripheral blood obtained from healthy HIV-negative donors was enriched for peripheral blood mononuclear cells (PBMC) by centrifugation over Ficoll Hypaque gradient. PBMC were then further enriched for lymphocytes or monocytes by elutriation.[25]

Cells obtained by this method are >90% pure as determined by cytofluorimetric assay (FACS). Cells were grown at 37 °C in 5% CO_2 atm in RPMI 1640 medium and supplemented with 10% fetal calf serum (FCS), 100 IU/mL penicillin G, and 100 mg/mL streptomycin. Cell cultures were checked periodically for the absence of mycoplasma contamination with a MycoTect Kit (Gibco).

Viruses. A laboratory lymphocyte−tropic strain of HIV (HTLV-IIIB, also called HIV-IIIB) and two primary isolates (HIV-112 and HIV-AB1) were used to infect MT-4, C8166, and lymphocytes. A laboratory monocyte−tropic strain of HIV-1 (HTLV-III-Ba-L, also called HIV-1Ba-L) was used to infect macrophages. Both the laboratory strains are available through the AIDS Research and Reference Reagent Program (NIH, Bethesda, MD). The primary isolates of HIV-1 were obtained from two HIV-1-Ab seropositive individuals who had experienced a virologic failure (plasma HIV RNA levels >500 copies/mL on at least two consecutive measurements) during HAART including zidovudine, lamivudine, efavirenz, tenofovir, and at least one protease inhibitor. Isolation of these strains from the plasma of infected individuals was performed in PBMC cultures; the supernatants of these cultures were used as a source of virus.

The Y181C and the K103N-Y181C mutant strains were obtained from the AIDS Research and Reference Reagent Program (NIH, Bethesda, MD). The Y181C mutant (HIV-1 119) was obtained from CEM cells infected with AZT-sensitive HIV-1 patient isolate A018. Resistance was selected for in vitro by culturing the AZT-sensitive A018 isolate in the presence of nevirapine. The K103N-Y181C mutant (HIV-1 IIIB A17) highly resistant to inhibition by RT nonnucleoside inhibitors, including pyridinone derivatives, BI-RG-587, and TIBO, was obtained by repeated passages in H9 cells.

HIV Titration. Titration to determine the infectivity of laboratory viral strains and the primary isolates was performed in C8166 cells (HIV-IIIB), lymphocytes (HIV-112 and HIV-AB1), and macrophages (HIV-1Ba-L).[26,27] The titer of virus stocks, expressed as 50% tissue culture infectious dose (TCID$_{50}$) was determined as previously described.[28]

Anti-HIV Assays. The activity of test compounds against multiplication of HIV-1 in MT-4 and C8166 cells was based on inhibition of virus-induced cytopathogenicity. Briefly, 150 mL of culture medium containing 2×10^4 cells was added to each well of flat-bottom microtiter trays containing 50 mL of culture medium with or without various concentrations of test compounds. Then 50 mL of HIV suspension (100 $TCID_{50}$) was added. After 5 days of incubation at 37 °C, the viral activity of compounds was evaluated by quantifying the HIV-induced cytopathogenicity by means of the 3-(4,5-dimethylthiazol-1-yl)-2,5-diphenyltetrazolium bromide (MTT) method. [29] Protection (%) from the viral-induced cytopathogenicity in HIV-infected cells as measured by the MTT assay was calculated as follows: $[(ODT)HIV - (ODC)HIV]/[(ODC)mock - (ODC)HIV]$, where $(ODT)HIV$ is the optical density measured with a given concentration of cysteamine in HIV-infected cells, $(ODC)HIV$ is the optical density measured for the control untreated HIV-infected cells, $(ODC)mock$ is the optical density measured for the control untreated mock-infected cells. The optical density at 540/690 nm was measured using a plate reader. For evaluation of antiviral activity in acutely infected lymphocytes, PHA-stimulated lymphocytes were washed twice with PBS, counted, distributed into 15-mL polyethylene tubes at a concentration of 6×10^6 cells/mL, and exposed to 100 $TCID_{50}$ of HIV in the presence or the absence of antiviral compounds. After 2 h of incubation in 5% CO_2 at 37 °C, the cells were washed twice in PBS and inoculated in each well of the 48-well plates (Costar, Cambridge, MA) at a concentration of 10^6 cells/mL of complete medium supplement with 10 units/mL of recombinant interleukin-2 (Collaborative Research Inc., Bedford, MA). The cells were subsequently cultured in the presence of the same concentrations of drugs as before. Half the volume of supernatant in each well was replaced every 3−4 days. The antiviral activity of the compounds was assessed by measuring HIV-p24 antigen production in the supernatants of infected cultures as previously described by using a commercially available HIV-antigen kit (Abbott, Pomezia, Italy). The assay to evaluate anti-HIV drug efficacy in acutely infected mature macrophages has been previously described. [30]

Briefly, 10^5 macrophages were suspended in complete medium. The cells were exposed to 100 $TCID_{50}$ of HIV in the presence or in the absence of the compounds and incubated at 37 °C in a CO_2 incubator. Two hours after the viral exposure, the macrophages were extensively washed to remove excess virus and then cultured under the same conditions and drug concentrations before. Cell were washed and fed every 7 days. The antiviral activity of the compounds was assessed by measuring HIV-p24 antigen production in the supernatants of infected cultures.

Cytotoxicity. The cytotoxicity of test compounds was evaluated in parallel with their antiviral activity and was based on the viability of mock-infected cells, as monitored by the MTT method. The 50% effective dose (EC_{50}) and 50% cytotoxic dose (TC_{50}) values were calculated from pooled values in the effective dynamic range of the antiviral activity and cytotoxicity assays (5−95%) using the median effect equation as previously described. [31] The selectivity index (SI) was calculated as the TC_{50}:EC_{50} ratio.

Enzymatic Assay Procedures. Chemicals. [^{3}H]dTTP (40 Ci/mmol) was from Amersham and unlabeled dNTPs from Boehringer. Whatman was the supplier of the GF/C filters. All other reagents were of analytical grade and purchased from Merck or Fluka.

Nucleic Acid Substrates. The homopolymer poly(rA) (Pharmacia) was mixed at weight ratios in nucleotides of 10:1 to the oligomer oligo(dT)12−18 (Pharmacia) in 20 mM Tris-HCl (pH 8.0), containing 20 mM KCl and 1 mM EDTA, heated at 65 °C for 5 min, and then slowly cooled at room temperature.

Expression and Purification of Recombinant HIV-1 RT Forms. The coexpression vectors pUC12N/p66(His)/p51with the wild-type or the mutant forms of HIV-1 RT p66[32] were kindly provided by Dr. S. H. Hughes (NCI-Frederick Cancer Research and Development Center). Proteins were expressed in *Escherichia coli* and purified as described. [33]

HIV-1 RT RNA-Dependent DNA Polymerase Activity Assay. RNA-dependent DNA polymerase activity was assayed as fol-lows: a final volume of 25 μL contained reaction buffer (50 mM Tris-HCl pH 7.5, 1 mM DTT, 0.2 mg/mL BSA, 4% glycerol), 10 mM $MgCl_2$, 0.5 μg of poly(rA)/oligo(dT) 10:1 (0.3 μM 3′-OH ends), 10 μM [^{3}H]dTTP (1Ci/mmol), and 2−4 nM RT. Reactions were incubated at 37 °C for the indicated time. Aliquots (20 μL) were then spotted on glass fiber filters (GF/C) which were immediately immersed in 5% ice-cold TCA. Filters were washed twice in 5% ice-cold TCA and once in ethanol for 5 min and dried, and acid-precipitable radioactivity was quantitated by scintillation counting.

Inhibition Assays. Reactions were performed under the conditions described for the HIV-1 RT RNA-dependent DNA polymerase activity assay. Incorporation of radioactive dTTP into poly(rA)/oligo(dT) at different substrate (nucleic acid or dTTP) concentrations was monitored in the presence of increasing fixed amounts of inhibitor. Data were then plotted according to Lineweaver−Burke and Dixon. For K_i determination, an interval of inhibitor concentrations between 0.2 K_i and 5 K_i was used.

Molecular Modeling. Docking Studies and 3-D QSAR. All molecular modeling calculations and manipulations were performed using the software packages Macromodel 7.1[34] and Autodock 3.0.5[35] running on IBM-compatible AMD Athlon 3.2 GHz workstations with the Linux operating system SUSE 9.2. For the conformational analysis and for any minimization, the all-atom amber[36] force field was adopted as implemented in the Macromodel package. The crystal structure of 739W94[16] extracted from the corresponding RT complex filed in the Brookhaven Protein Data Bank[37] (entry code 1jlq) was used as a template to model all the IASs derivatives (training set, test set, and newly designed derivatives). The NNBS was defined by selecting all the residues within 20 Å from the inhibitor.

To obtain a 3-D QSAR model able to discriminate active versus inactive derivatives, the latter being a source of important structural information, inactive compounds were thus included in the training set, to which was arbitrarily assigned an EC_{50} value equal to 50% of the least active compound. [38,39]

For the receptor-based alignment, the docking procedure within Autodock was used. For the docking a grid spacing of 0.375 Å and $60 \times 80 \times 60$ number of points were used. The grid was centered on the mass center of the experimental bound 739W94 coordinates. The GA-LS method was adopted using the default settings. Amber united atoms were assigned to the protein using the program ADT (Auto Dock Tools). Autodock generated 100 possible binding conformations grouped in clusters, the lowest energy conformation of the most populated cluster (Best Cluster conformation) was selected for the subsequent 3-D QSAR study. [14] Attempts to use the first ranked docked conformation (Best Docked conformation) led to less predictive models. [14]

The starting conformations for all the docking studies were obtained using a molecular dynamic run with simulated annealing procedure as implemented in Macromodel version 7.1 and conducted as previously reported. [14,40]

The interaction energies were calculated by using GRID (version 21). [41] A grid spacing of 1 Å was used setting the grid dimensions (Å) to X_{min}/X_{max}, 6.0/28.0; Y_{min}/Y_{max}, −46.0/−14.0; and Z_{min}/Z_{max}, −12.0/10.0, centered on the experimental bound conformation of 739W94. The water probe (OH2) was used to evaluate the molecular interaction fields of the modeled molecules.

PLS models were calculated with GOLPE 4.5.12[17] running on a O2 SGI workstation equipped with the operating system 6.5.11. The resulting probe−target interaction energies for each compound were unfolded to produce one-dimensional vector variables for each compound, which were assembled in the so-called **X** matrix. This matrix was pretreated by first using a cutoff of 5 kcal/mol to produce a more symmetrical distribution of energy values, zeroing small (<0.01) variable values, and finally removing variables with small standard deviation, using appropriate cutoffs. In addition, variables taking only two distributions were also removed. The smart region definition (SRD)[42] algorithm as implemented in the GOLPE program was also used. A number of seeds (1000) were selected using a D-optimal design criterion in the *weight space*. Structural

differences between different molecules in the series will be reflected in groups of variables, and therefore, groups were generated around each seed in the 3D-space. Variables with a distance of no more than 2 Å to the seeds were included in the groups. If two neighboring groups (with a distance smaller than 10 Å) contained the same information, the groups were collapsed. The groups were used in the variable selection procedure by replacing the original variables. The effect of the groups on the predictability was evaluated, and groups instead of individual variables were removed from the data file. The effect of the grouped variables on the predictability was evaluated using a fractional factorial design (FFD) procedure. A number of reduced models (twice the number of variables) were built by removing some of the variables according to the FFD design. The effect of dummy variables (20%) on the predictability was calculated, and only if a variable had a positive effect on the predictability larger than the effect of the average dummy variable was the variable included in the final model. The FFD selection was repeated until the r^2 and q^2 value did not increase significantly. In the FFD selection the cross-validation was conducted using five random groups 20 times and a maximum of three principal components. The models were validated using random groups. Molecules were assigned in a random way to five groups of equal size. Reduced models were built by keeping out one group at a time. The formation of the groups was repeated 20 times using a maximum model dimensionality of three components.

To ascertain the robustness of the model the cross-validation was also carried out using the less restrictive leave one out method (LOO) and the most restrictive leave half out method (LHO).

ALOGpS values were calculated using the standalone version of the ALOGPS program running under the Linux operating system.

Supporting Information Available: Elemental analysis results for compounds **6**−**13**, PLS coefficients, present fields, activity contribution, and T−U scores plots for the interpretation of the 3-D QSAR model. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) WHO/UNAIDS, AIDS Epidemic Update, 2005.

(2) Williams, I. G. Enfuvirtide (Fuzeon): The first fusion inhibitor. *Int. J. Clin. Pract.* **2003**, *57*, 890−7.

(3) Vandamme, A. M.; Van Vaerenbergh, K.; De Clercq, E. Anti-human immunodeficiency virus drug combination strategies. *Antivir. Chem. Chemother.* **1998**, *9*, 187−203.

(4) Tachedjian, G.; Goff Stephen, P. The effect of NNRTIs on HIV reverse transcriptase dimerization. *Curr. Opin. Invest. Drugs* **2003**, *4*, 966−73.

(5) De Clercq, E. Non-nucleoside reverse transcriptase inhibitors (NNRTIs): Past, present and future. *Chem. Biodiversity* **2004**, *1*, 44−64.

(6) Artico, M.; Silvestri, R.; Pagnozzi, E.; Stefancich, G.; Massa, S.; Loi, A. G.; Putzolu, M.; Corrias, S.; Spiga, M. G.; La Colla, P. 5*H*-Pyrrolo[1,2-*b*] [1,2,5]benzothiadiazepines (PBTDs): A novel class of non-nucleoside reverse transcriptase inhibitors. *Bioorg. Med. Chem.* **1996**, *4*, 837−50.

(7) Artico, M.; Silvestri, R.; Pagnozzi, E.; Bruno, B.; Novellino, E.; Greco, G.; Massa, S.; Ettorre, A.; Loi, A. G.; Scintu, F.; La Colla, P. Structure-based design, synthesis, and biological evaluation of novel pyrrolyl aryl sulfones: HIV-1 non-nucleoside reverse transcriptase inhibitors active at nanomolar concentrations. *J. Med. Chem.* **2000**, *43*, 1886−91.

(8) Artico, M.; Silvestri, R.; Massa, S.; Loi, A. G.; Corrias, S.; Piras, G.; La Colla, P. 2-Sulfonyl-4-chloroanilino moiety: A potent pharmacophore for the anti-human immunodeficiency virus type 1 activity of pyrrolyl aryl sulfones. *J. Med. Chem.* **1996**, *39*, 522−30.

(9) Silvestri, R.; De Martino, G.; La Regina, G.; Artico, M.; Massa, S.; Vargiu, L.; Mura, M.; Loi, A. G.; Marceddu, T.; La Colla, P. Novel indolyl aryl sulfones active against HIV-1 carrying NNRTI resistance mutations: Synthesis and SAR studies. *J. Med. Chem.* **2003**, *46*, 2482−93.

(10) Silvestri, R.; Artico, M.; De Martino, G.; La Regina, G.; Loddo, R.; La Colla, M.; La Colla, P. Simple, short peptide derivatives of a sulfonylindolecarboxamide (L-737,126) active in vitro against HIV-1 wild-type and variants carrying non-nucleoside reverse transcriptase inhibitor resistance mutations. *J. Med. Chem.* **2004**, *47*, 3892−3896.

(11) Silvestri, R.; Artico, M.; De Martino, G.; La Regina, G.; Loddo, R.; La Colla, M.; Mura, M.; La Colla, P. Simple, short peptide derivatives of a sulfonylindolecarboxamide (L-737,126) active in vitro against HIV-1 wild type and variants carrying non-nucleoside reverse transcriptase inhibitor resistance mutations. *J. Med. Chem.* **2004**, *47*, 3892−6.

(12) Williams, T. M.; Ciccarone, T. M.; MacTough, S. C.; Rooney, C. S.; Balani, S. K.; Condra, J. H.; Emini, E. A.; Goldman, M. E.; Greenlee, W. J.; Kauffman, L. R.; et al. 5-Chloro-3-(phenylsulfonyl)-indole-2-carboxamide: A novel, non-nucleoside inhibitor of HIV-1 reverse transcriptase. *J. Med. Chem.* **1993**, *36*, 1291−4.

(13) Young, S. D.; Amblard, M. C.; Brichter, S. F.; Grey, V. E.; Tran, L. O.; Lumma, W. C. J.; Huff, J. R.; Schleif, W. A.; Emini, E. E.; O'Brien, J. A.; Pettibone, D. J. 2-Heterocyclic indole-3-sulfones as inhibitors of HIV-1 reverse transcriptase. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 491−496.

(14) Ragno, R.; Artico, M.; De Martino, G.; La Regina, G.; Coluccia, A.; Di Pasquali, A.; Silvestri, R. Docking and 3-D QSAR Studies on indolyl aryl sulfones. Binding mode exploration at the HIV-1 reverse transcriptase non-nucleoside binding site and design of highly active *N*-(2-Hydroxyethyl)carboxamide and *N*-(2-Hydroxyethyl)-carbohydrazide derivatives. *J. Med. Chem.* **2005**, *48*, 213−223.

(15) De Martino, G.; La Regina, G.; Ragno, R.; Coluccia, A.; Bergamini, A.; Ciaprini, C.; Sinistro, A.; Maga, G.; Crespan, E.; Artico, M.; Silvestri, R. Indolyl aryl sulfones (IASs). Indolyl aryl sulfones (IASs) as HIV-1 non-nucleoside reverse transcriptase inhibitors. Synthesis, biological evaluation and binding mode studies of new derivatives at indole-2-carboxamide. *Antiviral Chem. Chemother.* **2006**, in press.

(16) Chan, J. H.; Hong, J. S.; Hunter, R. N. r.; Orr, G. F.; Cowan, J. R.; Sherman, D. B.; Sparks, S. M.; Reitter, B. E.; Andrews, C. W., III; Hazen, R. J.; St. Clair, M.; Boone, L. R.; Ferris, R. G.; Creech, K. L.; Roberts, G. B.; Short, S. A.; Weaver, K.; Ott, R. J.; Ren, J.; Hopkins, A.; Stuart, D. I.; Stammers, D. K. 2-Amino-6-arylsulfo-nylbenzonitriles as non-nucleoside reverse transcriptase inhibitors of HIV-1. *J. Med. Chem.* **2001**, *44*, 1866−82.

(17) Baroni, M.; Costantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementi, S. Generating optimal linear PLS estimations (GOLPE): An advanced chemometric tool for handling 3D-QSAR problems. *Quant. Struct.-Act. Relat.* **1993**, *12*, 9−20.

(18) Pastor, M.; Cruciani, G.; Watson, K. A. A strategy for the incorporation of water molecules present in a ligand binding site into a three-dimensional quantitative structure−activity relationship analysis. *J. Med. Chem.* **1997**, *40*, 4089−102.

(19) Silvestri, R.; Artico, M.; De Martino, G.; La Regina, G.; Loddo, R.; La Colla, M.; Mura, M.; La Colla, P. Simple, short peptide derivatives of a sulfonylindolecarboxamide (L-737,126) active in vitro against HIV-1 wild type and variants carrying non-nucleoside reverse transcriptase inhibitor resistance mutations. *J. Med. Chem.* **2004**, *47*, 3892−6.

(20) Hopkins, A. L.; Ren, J.; Milton, J.; Hazen, R. J.; Chan, J. H.; Stuart, D. I.; Stammers, D. K. Design of non-nucleoside inhibitors of HIV-1 reverse transcriptase with improved drug resistance properties. 1. *J. Med. Chem.* **2004**, *47*, 5912−22.

(21) Silvestri, R.; De Le Martino, G.; La Regina, G.; Artico, M.; Massa, S.; Vargiu, L.; Mura, M.; Loi, A. G.; Marceddu, T.; La Colla, P. Novel indolyl aryl sulfones active against HIV-1 carrying NNRTI resistance mutations: Synthesis and SAR studies. *J. Med. Chem.* **2003**, *46*, 2482−2493.

(22) Tetko, I. V.; Tanchuk, V. Y.; Kasheva, T. N.; Villa, A. E. Estimation of aqueous solubility of chemical compounds using E-state indices. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1488−93.

(23) Corbett, J. W.; Rodgers, J. D. Discovery of second generation quinazolinone non-nucleoside reverse transcriptase inhibitors of HIV-1. *Prog. Med. Chem.* **2002**, *40*, 63−105.

(24) Sodroski, J. G.; Rosen, C. A.; Haseltine, W. A. Trans-acting transcriptional activation of the long terminal repeat of human T lymphotropic viruses in infected cells. *Science* **1984**, *225*, 381−5.

(25) Placido, R.; Mancino, G.; Amendola, A.; Mariani, F.; Vendetti, S.; Piacentini, M.; Sanduzzi, A.; Bocchino, M. L.; Zembala, M.; Colizzi, V. Apoptosis of human monocytes/macrophages in *Mycobacterium tuberculosis* infection. *J. Pathol.* **1997**, *181*, 31−8.

(26) Andreoni, M.; Sarmati, L.; Parisi, S. G.; Ercoli, L.; Rocchi, G. Efficient and reproducible new semimicromethod for the detection and titration of HIV in human plasma. *J. Med. Virol.* **1992**, *38*, 207−13.

(27) Bergamini, A.; Capozzi, M.; Ghibelli, L.; Dini, L.; Salanitro, A.; Milanese, G.; Wagner, T.; Beninati, S.; Pesce, C. D.; Amici, C.; et al. Cystamine potently suppresses in vitro HIV replication in acutely and chronically infected human cells. *J. Clin. Invest.* **1994**, *93*, 2251−7.

(28) Karber, G. Beitrag zur kollectiven behandlung pharmakologisher reihenversuche. *Arch. Exp. Pharmakol.* **1931**, *162*, 480−483.

(29) Pauwels, R.; Balzarini, J.; Baba, M.; Snoeck, R.; Schols, D.; Herdewijn, P.; Desmyter, J.; De Clercq, E. Rapid and automated tetrazolium-based colorimetric assay for the detection of anti-HIV compounds. *J. Virol. Methods* **1988**, *20*, 309−21.

(30) Perno, C. F.; Yarchoan, R.; Cooney, D. A.; Hartman, N. R.; Gartner, S.; Popovic, M.; Hao, Z.; Gerrard, T. L.; Wilson, Y. A.; et al. Inhibition of human immunodeficiency virus (HIV-1/HTLV-IIIBa-L) replication in fresh and cultured human peripheral blood monocytes/macrophages by azidothymidine and related 2′,3′-dideoxynucleosides. *J. Exp. Med.* **1988**, *168*, 1111−25.

(31) Chou, T. C. Derivation and properties of Michaelis−Menten type and Hill type equations for reference ligands. *J. Theor. Biol.* **1976**, *59*, 253−76.

(32) Boyer, P. L.; Tantillo, C.; Jacobo-Molina, A.; Nanni, R. G.; Ding, J.; Arnold, E.; Hughes, S. H. Sensitivity of wild-type human immunodeficiency virus type 1 reverse transcriptase to dideoxynucleotides depends on template length; the sensitivity of drug-resistant mutants does not. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4882−6.

(33) Maga, G.; Amacker, M.; Ruel, N.; Hubscher, U.; Spadari, S. Resistance to nevirapine of HIV-1 reverse transcriptase mutants: Loss of stabilizing interactions and thermodynamic or steric barriers are induced by different single amino acid substitutions. *J. Mol. Biol.* **1997**, *274*, 738−47.

(34) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. MacroModel−An integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.* **1990**, *11*, 440−467.

(35) Goodsell, D. S.; Morris, G. M.; Olson, A. J. Automated docking of flexible ligands: Applications of AutoDock. *J. Mol. Recog.* **1996**, *9*, 1−5.

(36) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, I. Thomas E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a package of computer programs for applying molecular mechanics, normal-mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Comm.* **1995**, *91*, 1−41.

(37) Berman, H.; Henrick, K.; Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat. Struct. Biol.* **2003**, *10*, 980.

(38) Ragno, R.; Marshall, G. R.; Di Santo, R.; Costi, R.; Massa, S.; Rompei, R.; Artico, M. Antimycobacterial pyrroles: Synthesis, anti-*Mycobacterium tuberculosis* activity and QSAR studies. *Bioorg. Med. Chem.* **2000**, *8*, 1423−32.

(39) Tafi, A.; Anastassopoulou, J.; Theophanides, T.; Botta, M.; Corelli, F.; Massa, S.; Artico, M.; Costi, R.; Di Santo, R.; Ragno, R. Molecular modeling of azole antifungal agents active against *Candida albicans*. 1. A comparative molecular field analysis study. *J. Med. Chem.* **1996**, *39*, 1227−35.

(40) Ragno, R.; Frasca, S.; Manetti, F.; Brizzi, A.; Massa, S. HIV-reverse transcriptase inhibition: Inclusion of ligand-induced fit by cross-docking studies. *J. Med. Chem.* **2005**, *48*, 200−212.

(41) Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *28*, 849−57.

(42) Pastor, M.; Cruciani, G.; Clementi, S. Smart region definition: A new way to improve the predictive ability and interpretability of three-dimensional quantitative structure−activity relationships. *J. Med. Chem.* **1997**, *40*, 1455−64.