3D-QSAR studies of substituted 1-(3, 3-diphenylpropyl)-piperidinyl amides and ureas as CCR5 receptor antagonists

ARTICLE *in* JOURNAL OF MOLECULAR MODELING · MAY 2007

Impact Factor: 1.74 · DOI: 10.1007/s00894-007-0173-z · Source: PubMed

4 AUTHORS, INCLUDING:

Prasad V Bharatam
National Institute of Pharmaceutical Educa…
240 PUBLICATIONS **2,227** CITATIONS
SEE PROFILE

Prabha Garg
National Institute of Pharmaceutical Educa…
45 PUBLICATIONS **249** CITATIONS
SEE PROFILE

J Mol Model (2007) 13:519–529
DOI 10.1007/s00894-007-0173-z

ORIGINAL PAPER

3D-QSAR studies of substituted 1-(3, 3-diphenylpropyl)-piperidinyl amides and ureas as CCR5 receptor antagonists

Yogesh D. Aher · Avantika Agrawal · Prasad V. Bharatam · Prabha Garg

Received: 13 October 2006 / Accepted: 11 January 2007 / Published online: 16 February 2007

Abstract 3D-QSAR studies on the derivatives of 1-(3,3-diphenylpropyl)-piperidinyl amide and urea as CCR5 receptor antagonists were performed by comparative molecular field analysis (CoMFA) and comparative molecular similarity indices (CoMSIA) methods to rationalize the structural requirements responsible for the inhibitory activity of these compounds. The global minimum energy conformer of the template molecule, the most active and pharmacokinetically stable molecule of the series, was obtained by systematic search and used to build structures of the molecules in the dataset. The best predictions for the CCR5-receptor were obtained with the CoMFA standard model (q^2=0.787, r^2=0.962) and CoMSIA model combined steric, electrostatic and hydrophobic fields (q^2=0.809, r^2=0.951). The predictive ability of CoMFA and CoMSIA were determined using a test set of 12 compounds giving predictive correlation coefficients of 0.855 and 0.83, respectively, indicating good predictive power. Further, the robustness of the model was verified by bootstrapping analysis. The contour maps produced by the CoMFA and CoMSIA models were used to identify the structural features relevant to the biological activity in this series. Based on the CoMFA and CoMSIA analysis, we have identified some key features in the series that are responsible for CCR5 antagonistic activity which may be used to design more potent 1-(3,3-diphenylpropyl)-piperidinyl derivatives and predict their activity prior to synthesis.

Keywords CCR5 · CoMFA · CoMSIA · 3D-QSAR · HIV-1

Introduction

Acquired immunodeficiency syndrome (AIDS), a fatal pandemic caused by infection with human immunodeficiency virus subtype 1 (HIV-1), a retrovirus of the lentivirus family, is one of the greatest challenges to humankind and a leading cause of death worldwide. HIV can only replicate inside human cells. The replicative cycle of HIV can be divided into entry and post entry steps [1, 2]. Entry of the HIV into a target cell (the cell of a macrophage or T-cell) consists of three vital steps: (1) attachment of the surface subunit (gp120) to the receptor (CD4); (2) gp120 engagement with co-receptor (CXCR4 or CCR5) resulting in the exposure of a surface co-receptor-binding domain; (3) membrane fusion mediated by the transmembrane subunit (gp41). Each of the stages can serve as a target for the HIV entry. Post entry steps [3] require the reverse transcriptase (RT), integrase (IN), and protease (PR) enzymes to complete the viral replication cycle. Once internalized, the viral capsid core is disassembled and the exposed viral RNA is reverse-transcribed, with the help of RT, into a double-stranded viral cDNA that is integrated into the host chromosome with the help of the retroviral integrase enzyme. The integrated DNA, called a provirus, serves as a template to form viral mRNA which subsequently yields the new functional proteins of the virus. In the last stage of the life cycle, long strands of HIV proteins are chopped into smaller pieces by protease to construct a mature viral core

Electronic supplementary material The online version of this article (doi:10.1007/s00894-007-0173-z) contains supplementary material, which is available to authorized users.

Y. D. Aher · A. Agrawal · P. V. Bharatam · P. Garg (✉)
Centre of Pharmacoinformatics, National Institute of Pharmaceutical Education and Research (NIPER), Sector 67, S. A. S. Nagar, Mohali 160 062, India
e-mail: prabhagarg@niper.ac.in

which re-assembles into a new viral envelope and then buds from the surface of cell membrane to become a new infective viral agent. The immune system gradually scrambles down and patients are potentially vulnerable to opportunistic infections.

Among the available 20 FDA approved anti-HIV drugs, most belong to two classes of inhibitors that target either HIV reverse transcription or proteolytic maturation [4]. Triple therapy, commonly referred to as Highly Active Anti-Retroviral Therapy (HAART), is presently the standard for treatment of HIV infections [5]. It consists of a protease inhibitor (PI) or a non-nucleoside reverse transcriptase inhibitor (NNRTI) in combination with two nucleoside reverse transcriptase inhibitors. However, the success of current therapy is limited by the emergence of multi drug-resistant mutant strains, necessitating sustained disciplined adherence to complex regimens, the potential for toxic effects and high costs [6, 7]. Also, while these allow suppression of viral replication, the virus is not eradicated and the immune system eventually succumbs to infection [8]. Therefore, additional therapeutic approaches are warranted.

The chemokine receptor CCR5, a primary co-receptor essential for HIV recognition and entry into cell, has been identified as a potential new target to devise newer anti-HIV therapeutics [9]. CCR5 receptor is characterized by seven trans-membrane motifs and belongs to the rhodopsin family of G-protein coupled receptors (GPCR) [10]. Macrophage Inflammatory Proteins (MIP-1α and MIP-1β) and RANTES are the natural ligands of CCR5 [11]. It has been found that blocking the function of CCR5 strongly reduces HIV activity while exhibiting few side effects [12]. Furthermore, experimentation on CCR5-deficient mice indicate a novel role of CCR5 in down-regulation of mammalian immune function [13]. These observations strengthened the belief that functional inhibition of CCR5 receptor can be highly protective against macrophage–tropic HIV-1 infection.

Drugs that target the CCR5 receptor are not available in the market. Hence, novel antagonistic agents with high

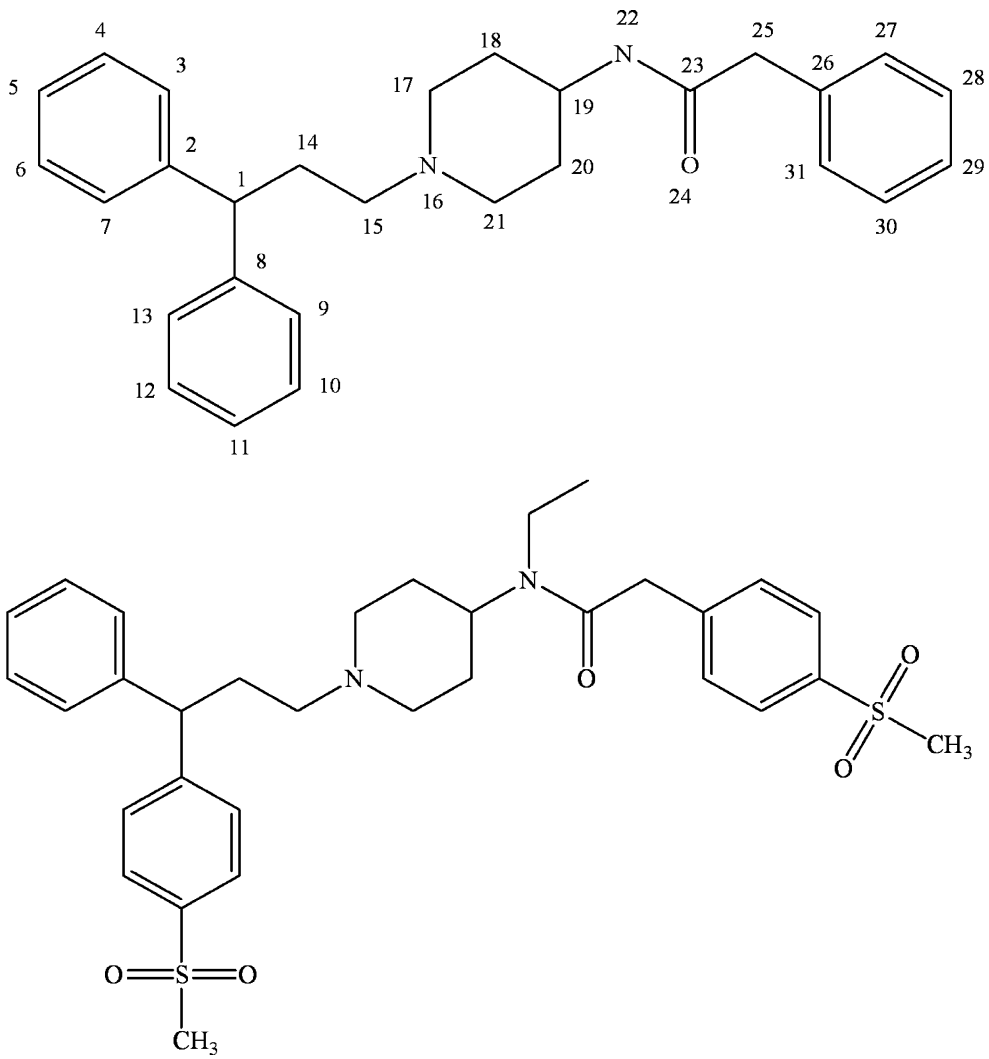


Fig. 1 Common fragment used for alignment and template molecule (compound **67**)

potency and selectivity are still in demand, the development of which requires a more detailed investigation about how the structural features of these antagonists influence their biological activity toward the CCR5 receptor. Since the crystal structure of CCR5 or its complex with any ligand is unavailable, correlating the physicochemical properties or structural features of compounds with their biological activities is believed to gain an insight into the interaction mechanism of the CCR5 receptor to antagonists, providing useful clues for designing new anti-HIV-1 drugs.

Previously, a few SAR and QSAR studies using 3D-QSAR methods, multiple linear regression analysis and the linear free energy-related (LFER) model of Hansch have been carried out on CCR5 receptor antagonists [14–22]. 2D QSAR analysis has already been reported on the series of 1-(3,3-diphenylpropyl)-piperidinyl amide and urea derivatives [22]. In the 2D study, the 3D structural information has not been employed, which is very crucial for the varieties of conformations possible at C-25 substitution, which are indeed responsible for the observed variations. Thus, in order to understand the antagonistic mechanism and guide the discovery and synthesis of more potent ligands, this series was chosen to perform 3D-QSAR studies with both CoMFA and CoMSIA methods.

Methods

Dataset for analysis

The affinity (50% inhibitory concentration, IC_{50}) data of 75 compounds was collected from the published work by Cumming et al. [23, 24]. It is imperative to evaluate the predictivity of the 3D-QSAR models generated. Selection of the training set and test set molecules was done such that the test set molecules represent a range of biological activity similar to that of the training set; thus, the test set is truly representative of the training set. The IC_{50} values were converted to pIC_{50} (M) to get the linear relationship in the equation, using the following formula:

$$pIC_{50} = -\log IC_{50}$$

where IC_{50} is the concentration of the antagonist producing 50% inhibition of CCR5 receptor.

Molecular modeling

All computations and molecular modeling studies were carried out on a Silicon Graphics Fuel work station running the molecular modeling software package SYBYL 7.1 [25]. Since the crystal structure of the CCR5 is not known, a systematic search was carried out for the most active molecule **67** with an interval of 15° on rotatable bonds to obtain the lowest energy conformation which was then minimized using PM3 Hamiltonian using MOPAC interfaced with SYBYL7.1. The rest of the molecules were built by changing the required substitution of the template and were minimized similarly. Finally, Gasteiger-Hückel charges were assigned to all the molecules.

Molecular alignment

Structural alignment is one of the most sensitive parameters in 3D-QSAR analyses. The accuracy of the prediction of CoMFA and CoMSIA models and the reliability of the contour models depend strongly on the structural alignment

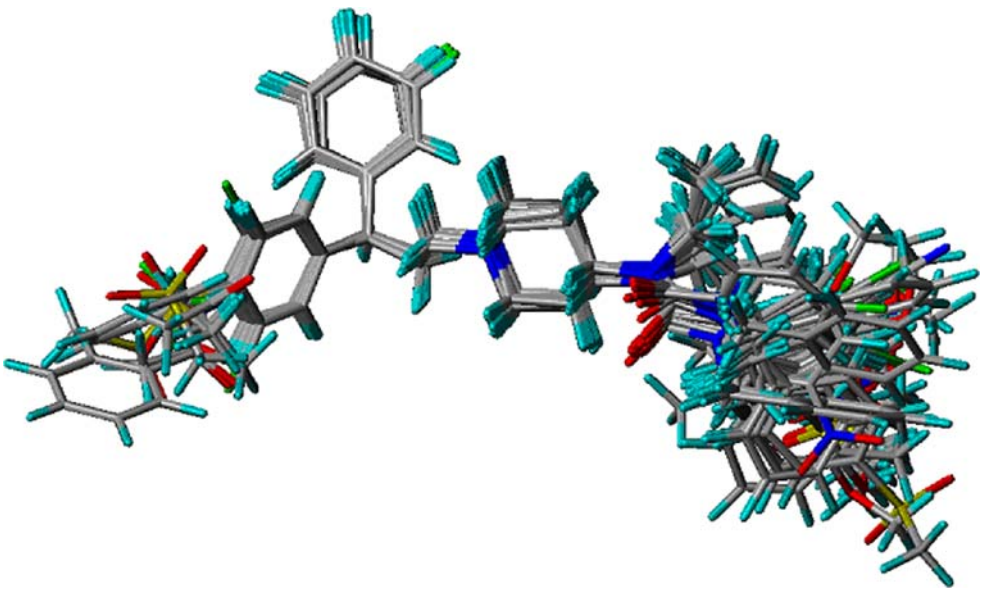
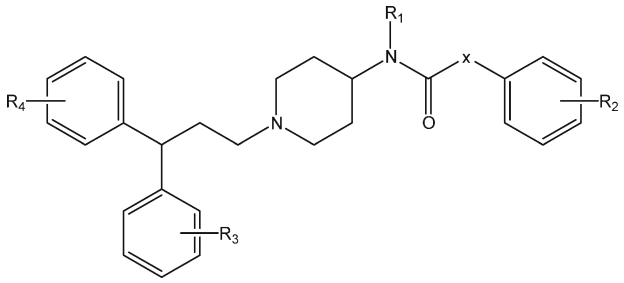


Fig. 2 Superposition of 75 molecules including compounds in the training set and test set on the template molecule **67**

Table 1 Structures and CCR5 binding affinities of 1-(3,3-diphenylpropyl)-piperidinyl amides and ureas in the training set and test set



Compound	X	R_1	R_2	R_3	R_4	pIC_{50} (M)
Training set						
1	–	CH_3	4-F	–	–	5.14
3	$CH(CH_3)$	CH_3	–	–	–	5.17
4	CH_2	CH_3	–	–	–	5.23
5	CH_2	CH_3	2-Cl	–	–	5.44
6	CH_2	CH_3	3-Cl	–	–	5.66
7	CH_2	CH_3	4-Cl	–	–	6.1
8	CH_2	CH_3	3,4-Cl_2	–	–	6.11
9	CH_2	CH_3	2,4-Cl_2	–	–	5.59
10	CH_2	CH_3	2-F	–	–	5.72
11	CH_2	CH_3	3-F	–	–	5.85
12	CH_2	CH_3	4-F	–	–	6.18
13	CH_2	CH_3	3,4-F_2	–	–	6.16
15	CH_2	CH_3	4-OMe	–	–	6.24
16	CH_2	CH_3	3,4-OMe	–	–	6.19
17	CH_2	CH_3	3,5-OMe	–	–	5.57
18	CH_2	CH_3	2,4,5-OMe	–	–	5.96
19	CH_2	CH_3	4-Br	–	–	6.24
20	CH_2	CH_3	4-Benzyloxy	–	–	5.46
21	CH_2	CH_3	4-Phenyl	–	–	5.64
22	CH_2	CH_3	4-CF_3	–	–	6.43
23	CH_2	CH_3	4-OCF_3	–	–	6.54
24	CH_2	CH_3	4-NHCOMe	–	–	6.17
25	CH_2	CH_3	4-CN	–	–	7.22
26	CH_2	CH_3	4-SO_2NH_2	–	–	7.04
28	CH_2	CH_3	4-SMe	–	–	6.25
29	CH_2	CH_3	4-CO_2Me	–	–	6.2
30	CH_2	CH_3	4-OH	–	–	6.33
32	CH_2	Ethyl	4-OCF_3	–	–	6.51
33	CH_2	Ethyl	4-CN	–	–	7.18
34	CH_2	Ethyl	4-SO_2NH_2	–	–	7.42
35	CH_2	Ethyl	4-SO_2Me	–	–	7.12
36	CH_2	Ethyl	4-NO_2	–	–	6.96
38	CH_2	c-propyl	4-SO_2Me	–	–	7.29
39	CH_2	c-propyl	4-NO_2	–	–	6.51
40	CH_2	Allyl	4-OCF_3	–	–	6.46
41	CH_2	Allyl	4-SO_2Me	–	–	7.43
42	CH_2	Allyl	4-NO_2	–	–	6.74
43	NH	CH_3	3-CN	–	–	5.31
44	NH	CH_3	3-CH_3	–	–	5.23
45	NH	Allyl	–	–	–	5.68
46	NH	CH_3	3,4-Cl_2	–	–	6.43
48	NH	Ethyl	4-CH_3	–	–	6.49
49	NH-CH_2	CH_3	–	–	–	7.00
51	NH-$CH(CH_3)$	Ethyl	–	–	–	5.27
52	NH-CH_2	Allyl	3-CH_3	–	–	6.62

Table 1 (continued)

Compound	X	R_1	R_2	R_3	R_4	pIC_{50} (M)
53	$NH\text{-}CH_2$	Allyl	$4\text{-}OCH_3$	–	–	6.96
54	$NH\text{-}CH_2$	Ethyl	$3\text{-}CH_3$	–	–	7.46
55	$NH\text{-}CH_2$	Ethyl	4-OMe	–	–	7.38
56	$NH\text{-}CH_2$	Ethyl	$4\text{-}SO_2Me$	–	–	7.31
58	CH_2	Ethyl	$4\text{-}SO_2Me$	4-Cl	–	8.07
59	CH_2	Ethyl	$4\text{-}SO_2Me$	3-Cl	–	7.24
60	CH_2	Ethyl	$4\text{-}SO_2Me$	$3,4\text{-}Cl_2$	–	7.11
61	CH_2	Ethyl	$4\text{-}SO_2Me$	$4\text{-}CH_3$	–	7.55
63	CH_2	Ethyl	$4\text{-}SO_2Me$	$4\text{-}CO_2CH_3$	–	8.15
64	CH_2	Ethyl	$4\text{-}SO_2Me$	4-OMe	–	8.2
66	CH_2	Ethyl	$4\text{-}SO_2Me$	4-SMe	–	7.92
67	CH_2	Ethyl	$4\text{-}SO_2Me$	$4\text{-}SO_2Me$	–	8.77
69	CH_2	Ethyl	$4\text{-}SO_2Me$	4-NHCOMe	–	7.59
71	CH_2	Ethyl	$4\text{-}SO_2Me$	$4\text{-}NHSO_2Me$	–	7.74
72	CH_2	Ethyl	$4\text{-}SO_2Me$	4-Cl	3-F	7.72
73	CH_2	Ethyl	$4\text{-}SO_2Me$	$4\text{-}NH_2$	3-F	6.98
74	CH_2	Ethyl	$4\text{-}SO_2Me$	4-NHCOMe	3-F	8.10
75	CH_2	Ethyl	$4\text{-}SO_2Me$	$4\text{-}NHSO_2Me$	3-F	8.04
Test set						
2	–	CH_3	$3\text{-}NO_2$	–	–	5.29
14	CH_2	CH_3	3-OMe	–	–	6.17
27	CH_2	CH_3	$4\text{-}SO_2N(Me)_2$	–	–	7.34
31	CH_2	CH_3	$4\text{-}NO_2$	–	–	6.82
37	CH_2	c-propyl	$4\text{-}SO_2NH_2$	–	–	7.48
47	NH	CH_3	4-F	–	–	6.72
50	$NH\text{-}CH_2$	Ethyl	–	–	–	7.21
57	CH_2	Ethyl	$4\text{-}SO_2Me$	4-F	–	6.51
62	CH_2	Ethyl	$4\text{-}SO_2Me$	$4\text{-}CF_3$	–	8.64
65	CH_2	Ethyl	$4\text{-}SO_2Me$	4-Ph	–	7.07
68	CH_2	Ethyl	$4\text{-}SO_2Me$	$4\text{-}NH_2$	–	6.70
70	CH_2	Ethyl	$4\text{-}SO_2Me$	4-NHCOPh	–	6.89

of the molecules [26]. The most active and pharmacokinetically stable compound **67** was used as a template for superimposition, assuming that its conformation represents the most bioactive conformation at receptor active site level. The common fragment shown in Fig. 1 was selected for DATABASE ALIGNMENT method in SYBYL and the rest of the molecules were aligned on it. The aligned compounds are shown in Fig. 2.

PLS analysis and validation of QSAR models

To derive 3D-QSAR models, the CoMFA and CoMSIA descriptors were used as independent variables and the pIC_{50} as the dependent variable. PLS method was used to linearly correlate these CoMFA and CoMSIA descriptors to the inhibitory activity values.

The CoMFA cutoff values were set to 30 kcal mol^{-1} for both steric and electrostatic fields, and all fields were scaled by the default options in SYBYL. The cross validation analysis was performed using the leave one out (LOO) method in which one compound is removed from the data set and its activity is predicted using the model derived from the rest of the dataset. The cross-validated correlation coefficient (q^2) that resulted in optimum number of components and lowest standard error of prediction were considered for further analysis and calculated using following formula:

$$q^2 = 1 - \frac{\sum_{\gamma} \left(\gamma_{pred} - \gamma_{actual}\right)^2}{\sum_{\gamma} \left(\gamma_{actual} - \gamma_{mean}\right)^2}$$

$$PRESS = \sum_{\gamma} \left(\gamma_{pred} - \gamma_{actual}\right)^2$$

where, γ_{pred}, γ_{actual} and γ_{mean} are predicted, actual, and mean values of the target property (pIC_{50}), respectively, and PRESS is the sum of predictive sum of squares. The non-cross-validated PLS analyses were performed with a

Table 2 Summary of CoMFA and CoMSIA results for various models

Model	q^2	n	r^2	SEE	F	r^2–bs	SD–bs
CoMFA (std)	0.787	3	0.962	0.185	239.4	0.98	0.004
CoMSIA (SE)	0.797	3	0.947	0.222	162.8	0.975	0.008
CoMSIA (SEH)	0.809	3	0.951	0.210	182.8	0.976	0.006
CoMSIA (SEHA)	0.801	3	0.948	0.217	170.1	0.972	0.007
CoMSIA (SEHD)	0.791	4	0.938	0.237	141.9	0.961	0.011
CoMSIA (SEA)	0.77	3	0.939	0.236	142.7	0.963	0.008
CoMSIA (SED)	0.767	5	0.925	0.261	114.7	0.949	0.008
CoMSIA (SEAD)	0.763	6	0.928	0.255	120.5	0.957	0.012
CoMSIA (ALL)	0.783	5	0.942	0.264	164.5	0.967	0.009

q^2 Leave one out (LOO) cross-validated correlation coefficient
n Optimum number of components
r^2 Non cross-validated correlation coefficient
SEE Standard error of estimate
FF-test value
r^2–bs Mean *r*-squared of boot strapping analysis (ten runs)
SD–bs Standard deviation by boot strapping analysis
S, *E*, *H*, *A*, *D* Steric, electrostatic, hydrophobic, as well as hydrogen-bond acceptor and donor fields, respectively
ALL The combination of all fields

column filter value of 2.0, to reduce analysis time with small effect on the q^2 values. To further assess the robustness and statistical confidence of the derived models, bootstrapping analysis for 10 runs was performed.

To assess the predictive power of the 3D-QSAR models derived using the training set, biological activities of an external test set of twelve molecules were predicted. The predictive ability of the models is expressed by the predictive r^2 value, which is analogous to cross-validated r^2 (q^2) and is calculated using the formula:

$$r^2_{pred} = \frac{SD - PRESS}{SD}$$

where SD is the sum of the squared deviations between the biological activities of the test set and mean activities of the training molecules and PRESS is the sum of squared

Table 3 Experimental pIC_{50}, predicted pIC_{50} and residual values of molecules used in the training set and test set for CoMFA and CoMSIA

Compound	Actual pIC_{50}	Predicted pIC_{50}		Residual values	
		CoMFA	CoMSIA	CoMFA	CoMSIA
Training set					
1	5.14	5.21	4.95	−0.07	0.19
3	5.17	5.33	5.42	−0.16	−0.25
4	5.23	5.43	5.53	−0.20	−0.30
5	5.44	5.38	5.31	0.06	0.13
6	5.66	5.57	5.70	0.09	−0.04
7	6.10	6.09	5.97	0.01	0.13
8	6.11	6.10	6.06	0.01	0.05
9	5.59	5.73	5.64	−0.14	−0.05
10	5.72	5.56	5.54	0.16	0.18
11	5.85	5.65	5.79	0.20	0.06
12	6.18	6.19	6.23	−0.01	−0.05
13	6.16	6.22	6.49	−0.06	−0.33
15	6.24	6.22	6.15	0.02	0.09
16	6.19	6.17	6.16	0.02	0.03
17	5.57	5.48	5.57	0.09	0.00
18	5.96	5.96	6.24	0.00	−0.28
19	6.24	6.10	6.00	0.14	0.24
20	5.46	5.35	5.45	0.11	0.01
21	5.64	5.73	5.73	−0.09	−0.09
22	6.43	6.68	6.71	−0.25	−0.28
23	6.54	6.55	6.26	−0.01	0.28
24	6.17	6.22	6.40	−0.05	−0.23

Table 3 (continued)

Compound	Actual pIC_{50}	Predicted pIC_{50}		Residual values	
		CoMFA	CoMSIA	CoMFA	CoMSIA
25	7.22	7.13	6.89	0.09	0.33
26	7.04	6.83	7.07	0.21	−0.03
28	6.25	6.45	6.13	−0.20	0.12
29	6.20	6.24	6.39	−0.04	−0.19
30	6.33	6.10	6.19	0.23	0.14
32	6.51	6.66	6.67	−0.15	−0.16
33	7.18	7.31	7.19	−0.13	−0.01
34	7.42	7.29	7.43	0.13	−0.01
35	7.12	7.21	7.31	−0.09	−0.19
36	6.96	6.81	6.92	0.15	0.04
38	7.29	7.19	6.96	0.10	0.33
39	6.51	6.49	6.50	0.02	0.01
40	6.46	6.41	6.56	0.05	−0.10
41	7.43	7.30	7.34	0.13	0.09
42	6.74	6.63	6.66	0.11	0.08
43	5.31	5.45	5.62	−0.14	−0.31
44	5.23	5.35	5.41	−0.12	−0.18
45	5.68	5.87	5.74	−0.19	−0.06
46	6.43	6.09	6.13	0.34	0.30
48	6.49	6.34	6.24	0.15	0.25
49	7.00	7.14	7.03	−0.14	−0.03
51	5.27	5.19	5.14	0.08	0.13
52	6.62	6.62	6.64	0.00	−0.02
53	6.96	6.87	6.85	0.09	0.11
54	7.46	7.48	7.44	−0.02	0.02
55	7.38	7.53	7.51	−0.15	−0.13
56	7.31	7.29	7.40	0.02	−0.09
58	8.07	7.74	7.74	0.33	0.33
59	7.24	7.29	7.29	−0.05	−0.05
60	7.11	7.35	7.40	−0.24	−0.29
61	7.55	7.41	7.51	0.14	0.04
63	8.15	8.35	8.33	−0.20	−0.18
64	8.20	8.19	8.15	0.01	0.05
66	7.92	8.04	8.09	−0.12	−0.17
67	8.77	8.59	8.55	0.18	0.22
69	7.59	7.67	7.60	−0.08	−0.01
71	7.74	7.84	7.71	−0.10	0.03
72	7.72	7.56	7.46	0.16	0.26
73	6.98	7.14	7.27	−0.16	−0.29
74	8.10	8.00	7.94	0.10	0.16
75	8.04	7.93	8.11	0.11	−0.07
Test set					
2	5.29	5.44	5.67	−0.15	−0.38
14	6.17	6.04	5.97	0.13	0.20
27	7.34	6.99	7.08	0.35	0.26
31	6.82	6.52	6.43	0.30	0.39
37	7.48	7.24	7.26	0.24	0.22
47	6.72	6.53	6.38	0.19	0.34
50	7.21	6.77	6.79	0.44	0.42
57	6.51	6.19	6.08	0.32	0.43
62	8.64	8.24	8.35	0.40	0.29
65	7.07	7.45	7.51	−0.38	−0.44
68	6.70	7.03	6.99	−0.33	−0.29
70	6.89	7.21	7.15	−0.32	−0.26

deviation between predicted and actual activities of the test set molecules.

Results and discussion

CoMFA and CoMSIA analyses

CoMFA and CoMSIA 3D-QSAR models were derived using a series of substituted 1-(3, 3-diphenylpropyl)-piperidinyl amides and ureas, possessing CCR5 receptor antagonistic activity. The chemical structures of the molecules and their actual pIC_{50} values are shown in Table 1. The data set was divided into a training set of 63 and a test set of 12 molecules (Table 1). The statistical parameters obtained from CoMFA and CoMSIA analysis are listed in Table 2. The best predictions were obtained by the CoMFA standard model (q^2=0.787, r^2=0.962) and CoMSIA combined steric, electrostatic and hydrophobic fields (q^2=0.809, r^2=0.951) which indicated that the built 3D-QSAR models are reliable and able to predict binding affinities of new derivatives accurately.

In CoMFA, the steric and electrostatic contributions were found to be 55% and 45%, respectively. Therefore, the steric field had a greater influence than the electrostatic field, indicating that the steric interactions of the molecules with the receptor could be an important factor for CCR5 antagonistic activity. CoMSIA models with different combinations of steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields were also evaluated by LOO cross validation and test set evaluation. These models showed good correlative and predictive values which indicated that the CoMSIA models were less dependant on the fields employed. In most of the models, hydrophobic field was a common factor indicating the importance of lipophilicity for the present series of molecules. To check whether the addition of hydrogen bond donor or acceptor descriptors affect the model, each descriptor was considered along with steric, electrostatic and hydrophobic descriptors for generating the model. It was observed that inclusion of hydrogen bond donor descriptor caused reduction in q^2 (0.791) whereas the addition of hydrogen bond acceptor caused an increase in q^2 (0.801). This indicates the importance of hydrogen bond acceptor functional group for the biological activity.

The observed and calculated (predicted) activity values of training (test) set molecules for CoMFA and CoMSIA are given in Table 3. The plots of observed versus calculated activity values for training set molecules and predicted versus observed activity values for the test set molecules for CoMFA and CoMSIA analyses are shown in Figs. 3a,b and 4a,b respectively.

The statistical significances of the 3D-QSAR models are much more than those of QSAR models built using the 2D approach, as indicated by the better q^2 values and smaller estimated errors [22]. Further, the QSAR models built using 3D methods provide an opportunity to explain the observed variance in the activity in terms of contour maps as discussed below.

Graphical interpretation of the results

The results of the CoMFA and CoMSIA models were generated by interpolating the products between the 3D-QSAR coefficients and their associated standard deviations. The CoMFA steric and electrostatic fields as well as

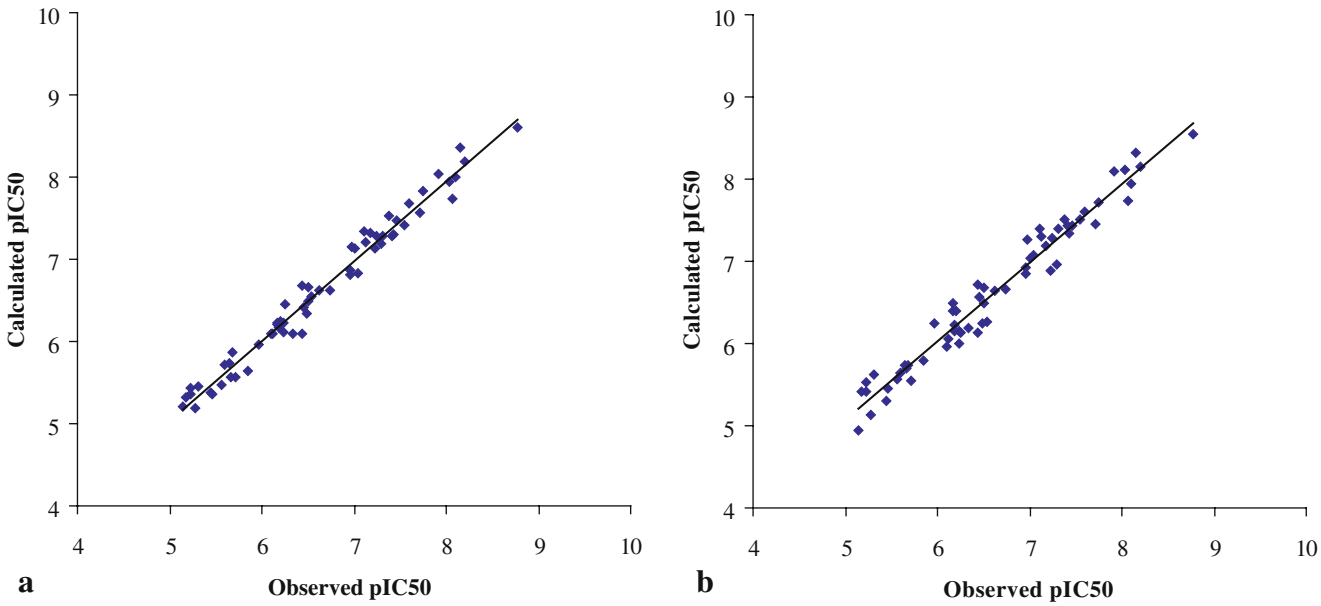


Fig. 3 Plot of observed versus calculated activity from CoMFA (**a**) and CoMSIA (**b**) analyses of the training set

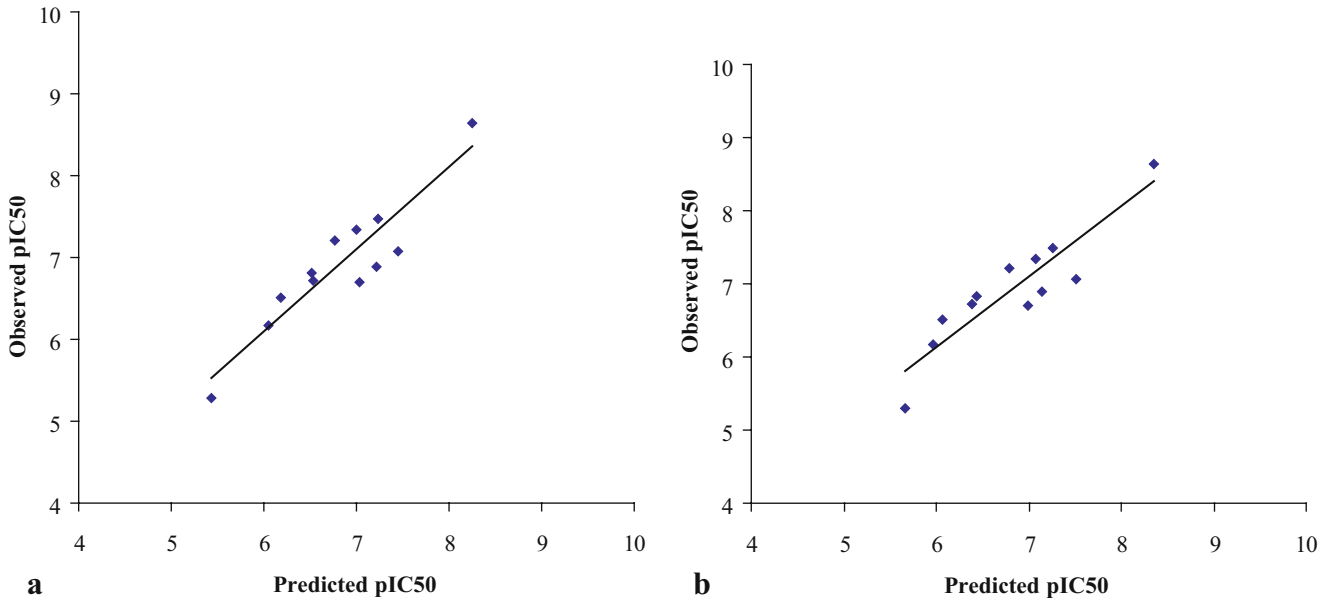


Fig. 4 Plot of predicted versus observed activity from CoMFA (**a**) and CoMSIA (**b**) analyses of the test set

CoMSIA different fields from the final non-cross-validated analyses were plotted as 3-D colored contour maps using compound **67** as a template structure.

In these figures, the green contours represent regions of high steric tolerance (80% contribution), while the yellow contours represent regions of low steric bulk tolerance (20% contribution). The increase in positive charge is favored in blue regions (80% contribution) while increase in negative charge is favored in red regions (20% contribution). Yellow contours represent hydrophobically favored regions (80% contribution) and white contours hydrophobically disfavored regions (20% contribution). Cyan contours represent regions where hydrogen bond donor substituents are favorable (80% contribution) while purple contours indicate regions where hydrogen bond donor substituents are not favorable (20% contribution).

The two special regions mapped to the template compound were C-29-substituted groups in the parent diphenylpropyl piperidinyl amide or urea structure and additional C-11-substituted groups. The influence of these substituent groups on their binding to the CCR5 receptor is discussed below from QSAR contour plots:

The steric contour of CoMFA (Fig. 5a) shows a large green contour enclosing the methanesulfonyl (R_2 substituent) of the template structure. This indicates that bulky R_2 substituents on the sulfonyl will enhance the CCR5 inhibitory activity. The good inhibitory potencies of **58**, **63**, **64** and **74** are due to orientation of their 4-methanesulfonyl group towards sterically favored green contour. Compounds **3** and **4** lacking bulky groups at the fourth position of the phenyl ring exhibited low inhibitory activity. Replacing the methyl with ethyl, cyclopropyl or allyl had

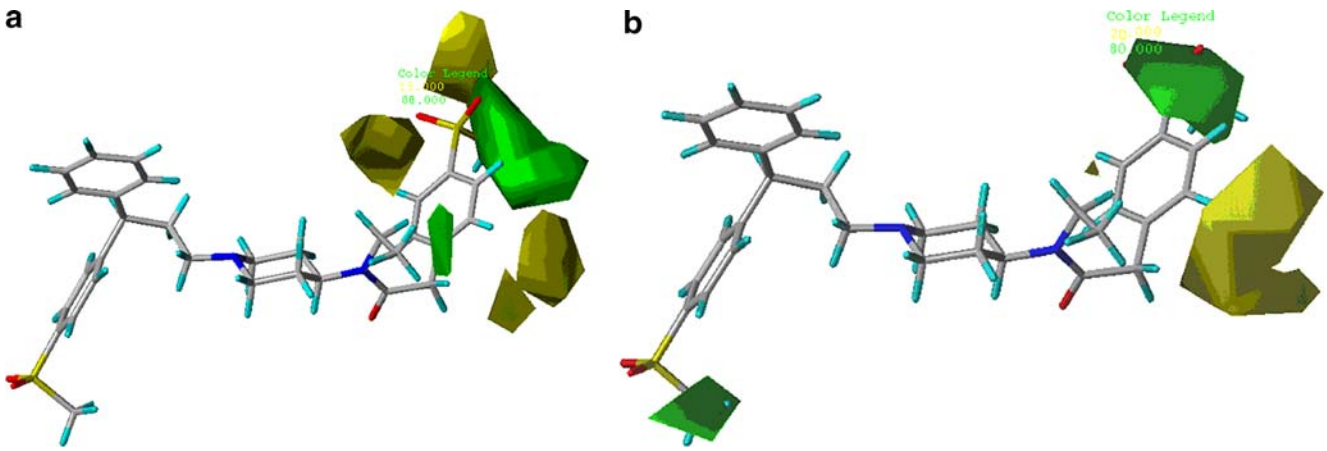


Fig. 5 CoMFA std (**a**) and CoMSIA combined (**b**) steric stdev* coeff contour plots: *green contours* indicate bulky (sterically favored) areas for enhancing binding affinity and increasing activity, whereas *yellow contours* indicate less bulky (sterically disfavored) areas for binding and decreasing activity

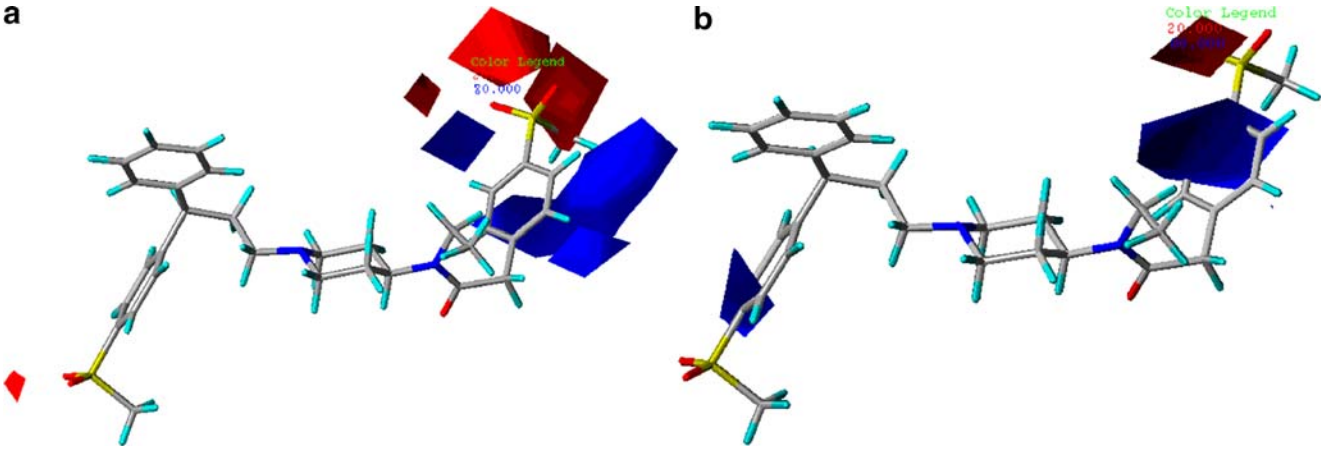


Fig. 6 CoMFA std (**a**) and CoMSIA combined (**b**) electrostatic stdev* coeff contour plots: *red contours* indicate negative charge increase binding affinity, whereas *blue contours* indicate positive charge favoring binding affinity

little effect on the binding potency when compounds with the same phenyl substituent were compared (e.g., the 4-nitrophenyl compounds **31**, **36**, **39** and **42** all have pIC_{50} values between 6.5 and 6.96M). Also small yellow contours at phenyl ring indicate that phenyl ring can be replaced with less bulkier 5-membered rings to enhance the activity. In addition, CoMSIA (Fig. 5b) shows small green contour at methyl group of methanesulfonyl (R_3 substituent), which indicates that more bulky group will enhance the activity.

The electrostatic contour of CoMFA (Fig. 6a) shows the red contour enclosing the sulfonyl oxygens (R_2 and R_3 substituents) of the template molecule where high electron density is expected to increase the activity. Hence, compounds having strong electron withdrawing groups at both these positions (**69**, **71**, **72**, **75**) exhibit good activity covered by the red contour. Increases in potency were also seen with electron-withdrawing groups at C-29 of phenyl ring (R_2 substituent), such as trifluoromethyl (**22**) and trifluoromethoxy (**23, 32**), and particularly with polar electron-withdrawing groups, such as nitro (**36**), sulfonamido (**26, 34**) and cyano (**33**). This is because these groups increase electron density, which is favorable for inhibitory activity. The electrostatic contours also show blue polyhedra in the vicinity of phenyl ring where low electron density is expected to increase activity. Substitution of halo group at *ortho* position (R_2 substituent) resulted in decrease of activity as seen in compounds **5**, **9** and **10**. The CoMSIA contour map for electrostatic field (Fig. 6b) also has similar a blue polyhedron indicating importance of electron density at the same region. In addition, it has a small blue contour at the C-11 position of the phenyl ring (R_3 substituent) indicating substitution of electropositive group at this position would increase the activity.

The hydrophobic contour (Fig. 7a) shows the presence of a large yellow contour surrounding the methanesulfonyl (R_3 substituent) of the template molecule indicating that hydrophobic substituents are well tolerated in that region. The white hydrophobic contour near the substituent X indicates that groups with hydrophilic characters are preferred in this position. The hydrogen bond donor contours (Fig. 7b) show presence of two purple contours close to the methanesulfonyl (R_3 substituent) and C-29 of

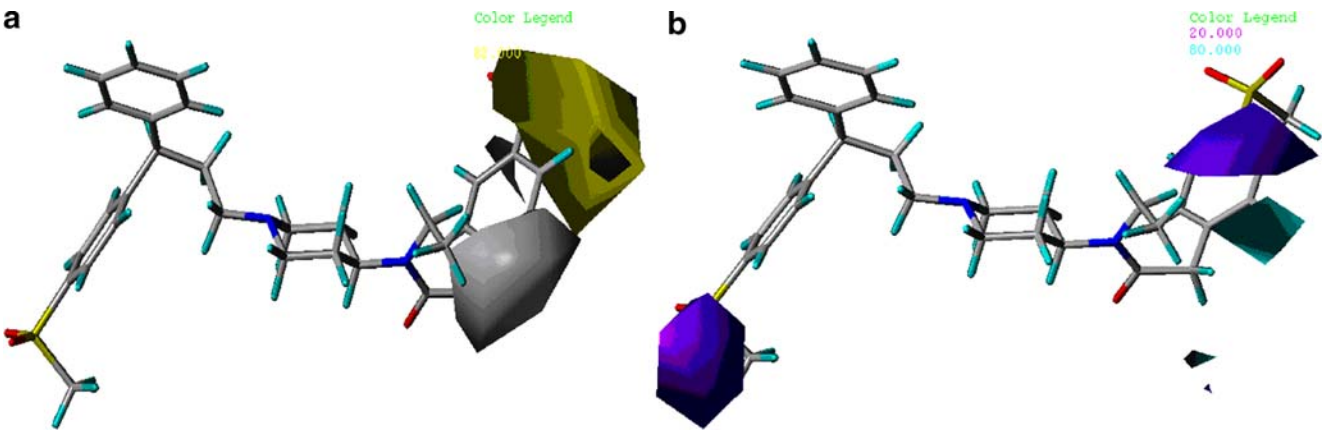


Fig. 7 CoMSIA combined hydrophobic (**a**) and H-bond donor (**b**) stdev*coeff contour plots: *yellow contours* predict hydrophobic groups enhance affinity, whereas *white contours* predict hydrophilic group enhance affinity; *cyan contours* predict H-bond acceptors on the receptor enhance binding, whereas *purple contours* reduce binding

phenyl ring (R_2 substituent) of the template molecule, where hydrogen bond donor substituents will decrease the activity. Two cyan contours are seen; one close to the phenyl ring and a small one enclosing the substituent X indicate regions in space where hydrogen bond donor groups are well tolerated.

Conclusions

In this study, CoMFA and CoMSIA 3D-QSAR models were derived to rationalize the CCR5 antagonistic activity of 75 compounds. The QSAR models showed good correlative and predictive capabilities in terms of q^2 and r^2 values. High bootstrapped r^2 values and small standard deviations indicate that a similar relationship exists between all compounds. The statistical significance and robustness of the 3D-QSAR models generated were confirmed using a test set. The effects of the steric, electrostatic, hydrophobic, hydrogen bond acceptor and hydrogen bond donor fields around the aligned molecules on their activities were clarified by analyzing the CoMFA and CoMSIA contour maps. The influence of C-29 and C-11 substitutions on binding affinities was also highlighted by these QSAR studies. The information from this study suggests that incorporating steric bulk, higher degree of electropositivity, and hydrophobicity on C-29 substitution at the phenyl ring might be favorable for better CCR5 antagonists. The structural requirements of the antagonists identified through the CoMFA and CoMSIA contour plots will help in designing new CCR5 receptor antagonists with enhanced activity.

References

1. Jiang S, Zhao Q, Debnath AK (2002) Curr Pharm Des 8:563–580
2. Sanders RW, Dankers MM, Busser E, Caffrey M, Moore JP, Berkhout B (2004) Retrovirology 1:3
3. Mager PP (2001) Med Res Rev 21:348–353
4. Xie L, Zhao CH, Zhou T, Chen HF, Fan BT, Chen XH, Ma JZ, Li JY, Bao ZY, Lo Z, Yu D, Lee KH (2005) Bioorg Med Chem 29:29
5. De Clercq E (1995) J Med Chem 38:2491–2517
6. Finzi D, Blankson J, Siliciano JD, Margolick JB, Chadwick K, Pierson T, Smith K, Lisziewicz J, Lori F, Flexner C, Quinn TC, Chaisson RE, Rosenberg E, Walker B, Gange S, Gallant J, Siliciano RF (1999) Nat Med 5:512–517
7. Richman DD (2001) Nature 410:995–1001
8. Moore JP, Stevenson M (2000) Nat Rev Mol Cell Biol 1:40–49
9. Wells TNC, Proudfoot AE, Power CA, Marsh M (1996) Chem Bio 3:603–609
10. Flower DR (1999) Biochim Biophys Acta 1422:207–234
11. Olson WC, Rabut GE, Nagashima KA, Tran DN, Anselma DJ, Monard SP, Segal JP, Thompson DA, Kajumo F, Guo Y, Moore JP, Maddon PJ, Dragic T (1999) J Virol 73:4145–4155
12. Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, MacDonald ME, Stuhlmann H, Koup RA, Landau NR (1996) Cell 86:367–377
13. Zhou Y, Kurihara T, Ryseck RP, Yang Y, Ryan C, Loy J, Warr G, Bravo R (1998) J Immunol 160:4018–4025
14. Dorn CP, Finke PE, Oates B, Budhu RJ, Mills SG, MacCoss M, Malkowitz L, Springer MS, Daugherty BL, Gould SL, DeMartino JA, Siciliano SJ, Carella A, Carver G, Holmes K, Danzeisen R, Hazuda D, Kessler J, Lineberger J, Miller M, Schleif WA, Emini EA (2001) Bioorg Med Chem Lett 11:259–264
15. Kim D, Wang L, Caldwell CG, Chen P, Finke PE, Oates B, MacCoss M, Mills SG, Malkowitz L, Gould SL, DeMartino JA, Springer MS, Hazuda D, Miller M, Kessler J, Danzeisen R, Carver G, Carella A, Holmes K, Lineberger J, Schleif WA, Emini EA (2001) Bioorg Med Chem Lett 11:3103–3106
16. Palani A, Shapiro S, Clader JW, Greenlee WJ, Vice S, McCombie S, Cox K, Strizki J, Baroudy BM (2003) Bioorg Med Chem Lett 13:709–712
17. Shankaran K, Donnelly KL, Shah SK, Guthikonda RN, MacCoss M, Mills SG, Gould SL, Malkowitz L, Siciliano SJ, Springer MS, Carella A, Carver G, Hazuda D, Holmes K, Kessler J, Lineberger J, Miller MD, Emini EA, Schleif WA (2004) Bioorg Med Chem Lett 14:3419–3424
18. Song M, Breneman CM, Sukumar N (2004) Bioorg Med Chem 12:489–499
19. Xu Y, Liu H, Niu C, Luo C, Luo X, Shen J, Chen K, Jiang H (2004) Bioorg Med Chem 12:6193–6208
20. Roy K, Leonard JT (2005) J Chem Inf Model 45:1352–1368
21. Afantitis A, Melagraki G, Sarimveis H, Koutentis PA, Markopoulos J, Igglessi-Markopoulou O (2006) J Comput Aided Mol Des 20:83–95
22. Leonard TJ, Roy K (2006) Bioorg Med Chem Lett 16:4467–4474
23. Burrows JN, Cumming JG, Fillery SM, Hamlin GA, Hudson JA, Jackson RJ, McLaughlin S, Shaw JS (2005) Bioorg Med Chem Lett 15:25–28
24. Cumming JG, Cooper AE, Grime K, Logan CJ, McLaughlin S, Oldfield J, Shaw JS, Tucker H, Winter J, Whittaker D (2005) Bioorg Med Chem Lett 15:5012–5015
25. SYBYL, version 7.1. (2005) Molecular Modeling Software, 474 Tripos Associates Inc, 1669, South Hanley Road, Suite 303, St. 475 Louis, Missouri, MO 63144–2913, USA
26. Cho SJ, Tropsha A (1995) J Med Chem 38:1060–1066