International Journal of Current Research in Chemistry and Pharmaceutical Sciences
www.ijcrcps.com Volume 1 Issue: 6 2014 Pages: 64-70
(pISSN: 2348-5213; eISSN: 2348-5221)

RESEARCH ARTICLE

3D QSAR STUDY OF PYRIMIDINES AS HIV ENTRY INHIBITORS

UTTAM K. TRIPATHI[1], SANGHMITRA SINGH[2], JAINENDRA JAIN[3] AND JAYA DWIVEDI[2]

[1]Teva API (I) Ltd. Greater Noida, U. P.-201306, India
[2]Bansthali vidhyapeeth, Tonk- , India
[3]Ram Eesh Institute of Vocational and Technical Education, Greater Noida, U. P.-201310, India

*Corresponding author e-mail: uktripathi@gmail.com

Abstract

Computation based QSAR study has been becoming a powerful tool in understanding the structural requirements for chemicals to bind with HIV entry receptors. A study was performed on a series of pyrimidienes derivatives as anti-HIV agents against HIV-1 (RP) by using molecular design suite - Vlife MDS. The study was performed on a series of 14 compounds (data set) using using random selection method for training and test data set. In this study multiple regression approach was applied to generate models. A significant model was obtained with r^2 = 0.9322, q^2 = 0.8261, r^2pred = 0.6759 and with electrostatic descriptor E_594 and E907. QSAR Equation is concluded that less electronegative groups are required in region of E594 * E907. The molecular field analysis (MFA) contour plots provided further understanding of the relationship between structural features of substituted pyrimidines darivatives and their activities which should be applicable to design newer potential anti-HIV agents.

Keywords: HIV entry inhibitors, pyrimidines, 3D QSAR, docking, multiple regression.

Introduction

According to UNAIDS 2012 global report, people living with HIV in the world in year 2011 is 34 millions. Once anybody got infection never possible to get rid off. Hence better protection is prevention not cure. Cure is usable to increase life time of peoples. There ma be different stages like viral entry, uncoating of capsid core, reverse transcription, nuclear import, integration, transcription and translation, production of new viron(Schols et al., 1997). No doubt that the reverse transcription is most interested target for researcher but cell fusion/entry inhibitors are also attracting researchers now a days.

Enveloped viruses, as a rule, enter their host cells by fusion between the viral envelope and cellular plasma membrane. This fusion process is basically similar for several enveloped virus families (that is, retro-, paramyxo- and herpesviruses), but HIV is preceded by the interaction of gp120 with its co-receptor on the host cell - the chemokine (C-X-C) motif receptor 4 (CXCR4) for T-tropic or X4 HIV strains, or the chemokine (C-C) motif receptor 5 (CCR5) for M-tropic or R5 HIV strains. CXCR4 and CCR5 normally act as the receptors for the C-X-C chemokine, SDF1 (stromalcell-derived factor 1), and the C-C chemokines RANTES (regulated upon activation, normal T-cell expressed and secreted) and MIP1 (macrophage inflammatory protein 1), respectively. The coincidental use of both CXCR4 and CCR5 by HIV as co-receptors to enter cells has prompted the search for CXCR4 and CCR5 antagonists, which, through blockade of the corresponding co-receptor, might be able to block HIV entry into the cells. This has now been shown with several compounds, the most prominent among the CXCR4 antagonists being the bicyclam AMD3100 (Schols et al., 1997; De Clercq, 2000),

and the best documented among the CCR5 antagonists being TAK779 9 (Baba et al., 1999; Dragic et al., 2000). The site of interaction of TAK779 with the transmembrane helices of CCR5 has been mapped (Dragic et al., 2000), and, likewise, crucial amino-acid residues involved in the binding of AMD3100 to CXCR4 have been identified (Hatse et al., 2001*)*. Recently, a new CCR5 antagonist, SCH-C (SCH351125), was announced as an orally bioavailable inhibitor ofM-tropic R5 strains that is capable of suppressing R5 HIV-1 infection both *in vitro* and *in vivo* (SCID-hu Thy/Liv mice) (Strizki et al., 2007). The clinical potential of the CXCR4 and CCR5 antagonists in the management of HIV infections remains to be proved. To ensure maximal coverage of both X4 and R5 strains, dual CXCR4/CCR5 antagonists should be developed, or single CCR5 and CXCR4 antagonists should be combined. The interaction of gp120 with its co-receptor (CCR5 or CXCR4) triggers a series of conformational changes in the gp120–gp41 complex that ultimately lead to the formation of a 'trimer-of-hairpins' structure in gp41 — a bundle of six α-helices: three α-helices formed by the carboxy-terminal regions packed in an antiparallel manner with three α-helices formed by the aminoterminal regions. The fusion-peptide region, located at the extreme amino terminus,will insert into the cellular membrane, whereas the carboxy-terminal region remains anchored in the viral envelope. In this sense, the trimer-of-hairpins motif brings the two membranes together, so agents that interfere with the formation of the gp41 trimer-of-hairpins structure might be expected to inhibit the fusion process.

Several constructs have been designed to interfere with the gp41-mediated fusion process: the so-called '5-helix', which binds the carboxy-terminal region of gp41 (Root et al., 2001); D-peptide inhibitors, which dock into the pocket formed by the α-helices of gp41 (Eckert et al., 1999); and T20 (pentafuside, previously called DP178, a synthetic 36-amino-acid peptide that corresponds to residues 127–162 of the ectodomain of gp41). T20 has proved effective in reducing plasma HIV levels in humans, providing the proof of concept that viral entry can be successfully blocked *in vivo* (Kilby et al., 1998).

Insight into the HIV fusion process should help in designing fusion inhibitors for other viruses as well, as trimer-of-hairpins motifs could also be predicted for other virus families (Root et al., 2001), including paramyxoviridae, such as parainfluenza virus, measles and respiratory syncytial virus. In fact, for each of these paramyxoviruses, peptides similar to T20 have been shown to block viral fusion (Lambert et al., 1996). Also, a cobalt-chelating complex (CTC96) that inhibits infection by HSV-1 through blocking fusion (Schwartz et al., 2001) might have an extended antiviral activity spectrum, given the premise that enveloped viruses belonging to different families share an analogous process of membrane fusion.

Infect, no work is reported on pyrimidines QSAR as viral fusion inhibitors but some derivatives have been evaluated as entry inhibitors and 3D QSAR study of these compounds may open a new field for future research.

Experimental

Material and methods

Set of molecules under study & corresponding EC_{50} value against HIV-1 (RP) was taken from website http://chemdb.niaid.nih.gov/ by selecting chemical class - pyrimidines and target - HIV, entry (assed in January, 2013) (Stiziki et al., 2005; Palani et al, 2003 and Tagat et al. 2004). The structure for all compounds is shown as figure 1 – 14. The biological activity data used in this study are *IC50*, the half maximal inhibitory concentration. Corresponding value of *IC50* for all compounds are listed in table 2.

3D QSAR

This is a module for generation of 3D QSAR equation using various statistical regression methods. The facilities for molecular alignment and and generation of steric and electrostatic interactive energies are also provided. This module helps in designing the novel ligand based on generated SAR model for variable selection.

It consist the following steps -
1. Building of 3D QSAR equation using multiple correlation statistical methods with stepwise forward analysis method.
2.
Here 3D QSAR was performed by using Vlife 3D QSAR (Vlife, 2002) installed on Core 2 Duo workstation. 3D structure were drawn for each molecule and molecular geomety were optimezed by using montecalo conformational search, Merck molecular force field (MMFF) and charges.

Figure1. structure of compounds under study

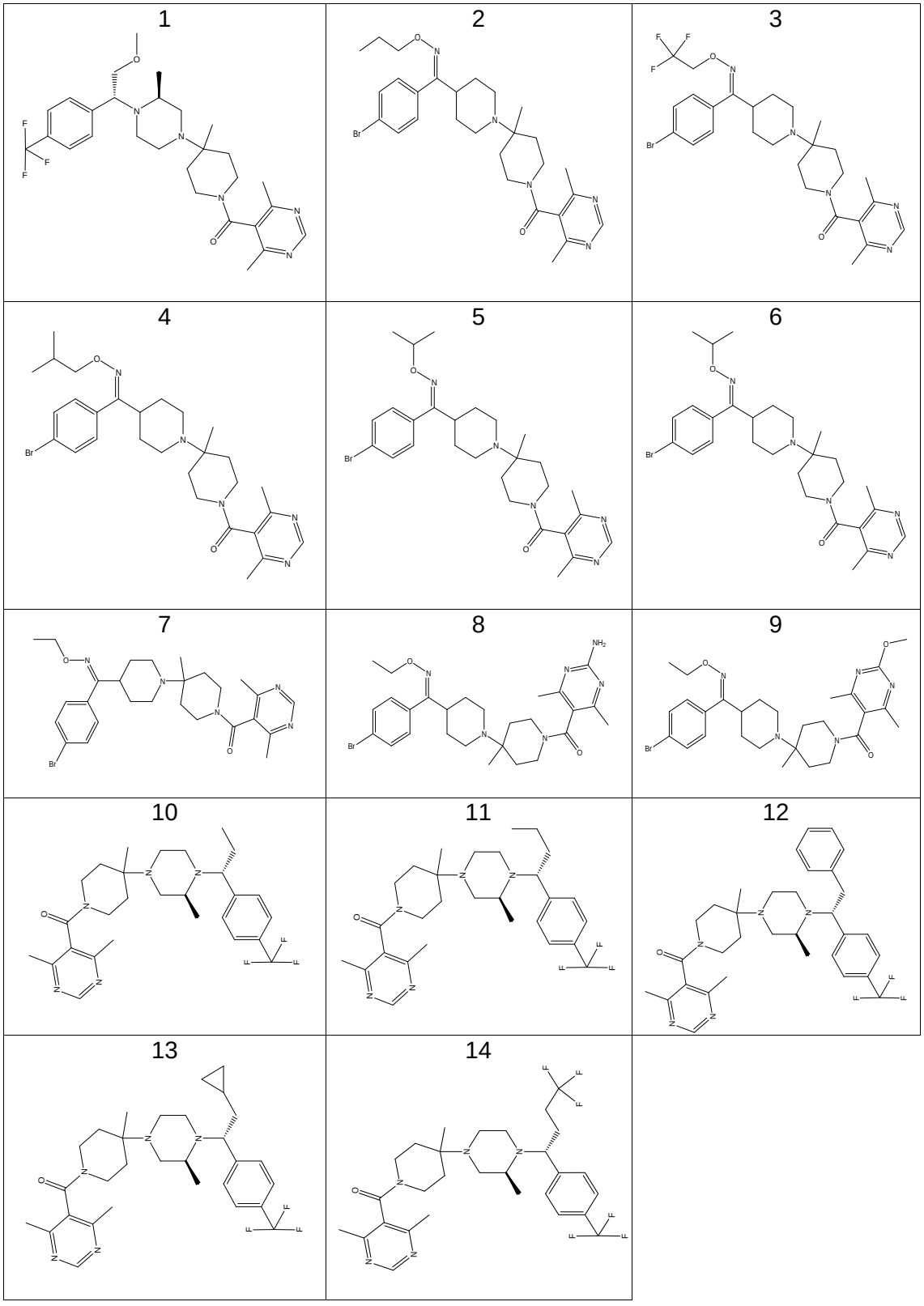


Table 2. for biological activity, descriptors value and calculated biolodical activity

Comp. No.	AIDS No.	EC_{50} (µm)	Mol. Wt.	EC_{50} (mg/ml)	-logEC_{50} (BA_{obs})	E_594	E_907	(BA_{calc})
1	165850	0.0001	533.64	5.3364E-08	7.27	0.185	-3.501	6.421
2	165851	0.00089	556.55	4.953295E-07	6.31	2.918	-10.00	6.429
3*	165852	0.0038	596.49	2.266662E-06	5.64	0.952	-2.236	6.637
4*	165857	0.0017	570.57	9.69969E-07	6.01	0.742	-10.00	6.075
5*	210504	0.00028	556.55	1.55834E-07	6.81	-1.611	-3.362	7.452
6*	210501	0.00093	556.55	5.175915E-07	6.29	0.055	-1.286	6.857
7	210502	0.00026	610.52	1.587352E-07	6.80	0.366	-0.945	6.748
8	210503	0.00035	557.53	1.951355E-07	6.71	-0.742	-0.111	7.037
9	210505	0.00013	572.54	7.44302E-08	7.13	2.772	-7.806	6.336
10	210506	0.00018	517.64	9.31752E-08	7.03	1.986	1.474	6.150
11	165845	0.00035	531.66	1.86081E-07	6.73	1.794	-3.049	6.426
12	165846	0.00013	579.71	7.53623E-08	7.12	3.47	-2.245	5.890
13*	165847	0.0008	543.67	4.34936E-07	6.36	1.637	-3.123	6.476
14	165849	0.0014	585.63	8.19882E-07	6.09	2.064	-3.108	6.349

*Compounds are in test set.

Monte carlo simulation explores the conformational search scale of molecule using random moves are accepted such that a different region of search space is sampled at each step. The most commonly used output from a conformational search in the set of torsion angel values, each of which produces a visible conformation, this together with the defnition of all rotatable bonds and starting cordinates for the molecule allow all valid conformation to be generated. In all its uses the metropolicies conditions and RMS deveation to acept or descard generaed conformers, dielectric properties were keeping constant at 1.0 using distance dipendant functions. Optimized molecule were aligned by template based method using the most active molecule.

Descriptors calculation

A common rectangular grid around molecule was generated. The steric and electrostatic energies were computed at lattice points of the grid using the methyl probe using charge +1. The term descriptor is utilized to indicate field value at lattice point 2080. The 3D descriptors were calculated setting charge type as Gasteiger-Marseli (GM) and dielectric constant value at 1.0. The descriptors with no variation in values were rejected. Descriptor with constant values will not contribute to QSAR, selection of training and test set and variable selection and model building Optimal training and test set was generated using RS algorithm. A training set of 14 molecules were generated. Forward method with PCR variable selection method was employed for selection of variable to obtain QSAR model. The step by step procedure begins by developing a trial model with single independent variable and adds independent variable, one step a time examining the fit of the model at each step, the method continues until there are no more significant variables remaining outside the model.

Cross validation

The standard leave-one-out (LOO) procedure was implemented that is a molecule in the training set was eliminated and its biological activity was predicted as the weighted average activity of the k most similar molecules (eq. 1)

$$y_i = \sum w_i y_i \quad \text{.....(1)}$$

The similarities were evaluated as the inverse of Euclidean distance between molecules (eq. 2) using only the subset of descriptors

$$d_{ij} = [\sum_{k=1}^{\neg n} (x_{i,k} - x_{j,k})]^{1/2} \quad \text{.....(2)}$$

This step was repeated until every molecule in the training set has been eliminated and its activity predicted once. The cross validated r^2 (q^2) values was calculated using equation 3 where yi and ŷi are actual and predicted of the ith molecule, respectively and ymean is the mean of observed activity of all molecule in the training set

$$q^2 = 1 - \sum (y_i - Y_i)^2 / \sum (y_i - y_{mean})^2 \quad \text{.....(3)}$$

Since the calculation of the pairwise molecules similarities enhance the prediction was based upon current training set, the q^2 values obtained (0.69) is the indicative power of the PCR model.

External validation

The external validation i.e. predicted r^2 (pred_r^2) value was calculated using the following equation , where y_i and ŷi are the actual and predicted biological activity values of the ith molecules in the test set respectively and ymean is the mean of observed biological activity values of all molecules in the training set

$$Pred\_r^2 = 1 - \sum (y_i - Y_i)^2 / \sum (y_i - Y_{mean})^2 \quad \text{.....(4)}$$

Both the simulations are overall molecules in the test set, the pred_r^2 value is indicative of all the predicted power of current MR model for external test set. For the given test set, this value was found to be 0.6759. The relative position of electrostatic field around aligned molecules, these are important for ClogP variations in the model shown by the Fig. 16.

Results and Discussion

Electronic parameters are important for Anti HIV activity and of pyrimidines deraivaties. As shown in both equation obtained.

Different training and test set of prymidines derivatives were constructed using auto data selection with 30% in test set.

Multiple regression (MR) in conjunction with stepwise forward method and cross correlation 0.5 was applied for building QSAR model. Results of model developed by MR method was satisfactorily and shown below.

Equation (Obtained by stepwise, forward,)

BA = -0.2969 (± 0.0330) E_594+ -0.0484(± 0.0116) E_907 + 6.8114

Test set molecule = 9, Training set molecule = 5, Degree of freedom = 6, r^2=0.9322, q^2=0.8261, F=41.2207, r^2Se = 0.1264, q^2Se=0.2024, Pred r^2= 0.6759, Pred r^2Se = 0.4023.

Data fitness plot were shown for model in figure 15. Results for observed and predicted biological activity were shown in table 2 for training and test data set (marked by *). This is clear from fitness plot that both models are able to predict activity of training set quite well as well as test set (external).

Interpretation of Model

The equation explains 93% (r^2=0.9322) of the total variance in the training set and has internal (q^2) and external (pred-r^2) predictive ability ~82% and 67% respectively. The F test show the statistical significance of 99.99% of the model which means that ability of failure ht model is 1 in 10000. In addition the randomisation test shows confidence of 99.9 (Alpha rad r2=-0.10994) that generated model is not random and hence chosen QSAR model.

From fitness plot it can be seen that model is able to predict the activity of training set quite well (all points are close to the regression line) as well as external test set providing confidence in the predictive ability of the model (Figure 1)

Model obtained by multiple regression method shows electrostatic interaction (both two) plays major role in determining anti-HIV activity. Statistically model is better with respect to squared correlation coefficient (pred r^2). It uses two electrostatc field descriptors contribution chart indicate that contribution E_594 86% and E_907 14%. There is no contribution of any steric and hydrophobic parameters according to model. Results plot in which 3D alignment of molecule with the important electrostatic field in the model provide guideline for new molecule design (Figure 2)..

Fig 15 fitness plot of training and test set molecules between actual and predicted biological activity

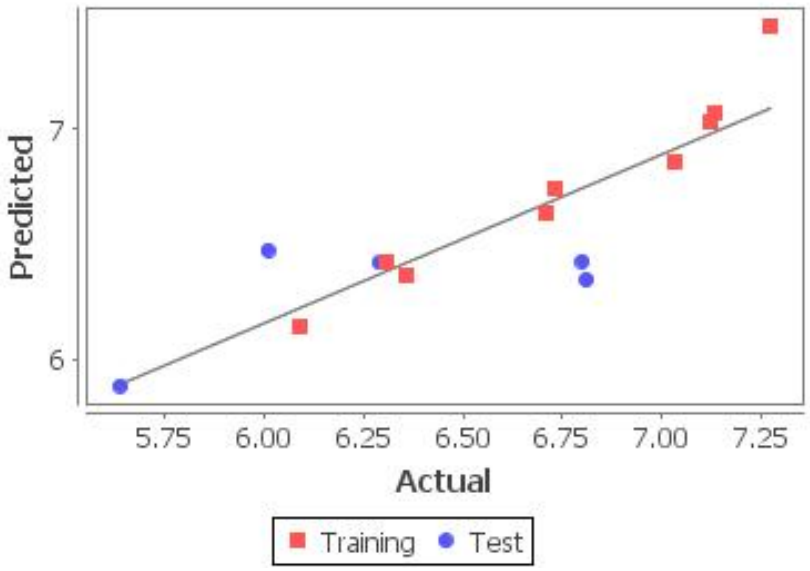
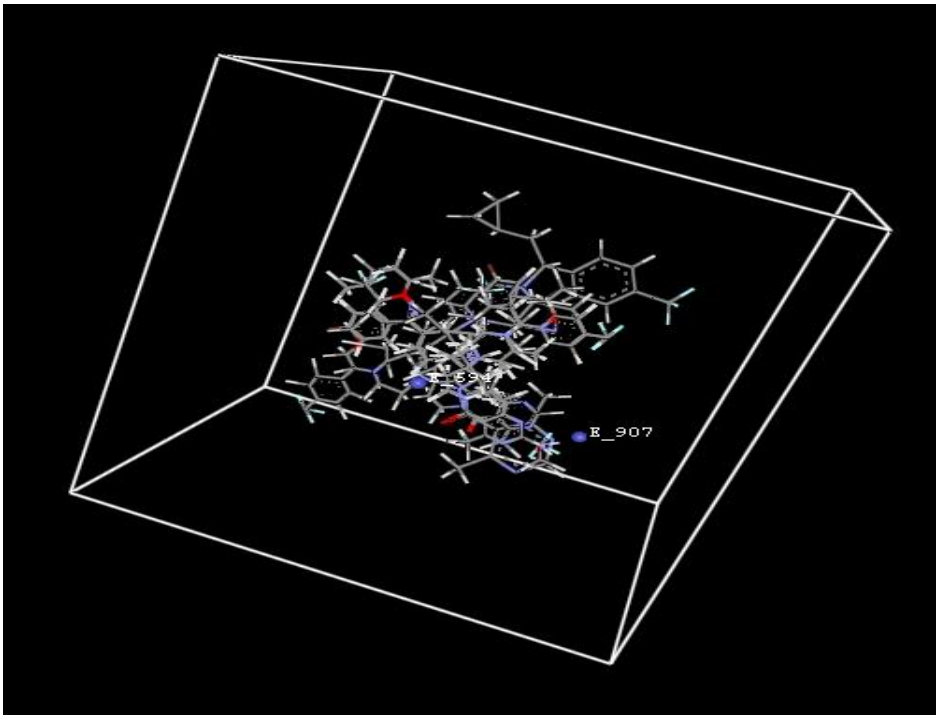


Fig. 16 3D alignment of molecule



(a) Electrostatic field E_594 has negative range indicates that positive electrostatic potential is unfavorable for increase in the activity and hence less electronegative substituent group preferred in that region.

(b) Electrostatic field E_907 has negative range indicates that negative electrostatic potential is favourable to increase in activity and hence less electronegatie substituent group is preffered in that region.

Finally, it is hopped that the work presented here will play an important role in understanding the relationship of physiochemical parameters with structure and anti-HIV activity . By studying the QSAR model, one can select the suitable substituens for active compound with maximum potency.

References

Baba, M. *et al.* A small-molecule, nonpeptide CCR5 antagonist with highly potent and selective anti-HIV-1 activity. *Proc. Natl Acad. Sci. USA* 96, 5698–5703 (1999).

De Clercq, E. Inhibition of HIV infection by bicyclams, highly potent and specific CXCR4 antagonists. *Mol. Pharmacol.* 57, 833–839 (2000)

Dragic, T. *et al.* A binding pocket for a small molecule inhibitor of HIV-1 entry within the transmembrane helices of CCR5. *Proc. Natl Acad. Sci. USA* 97, 5639–5644 (2000).

Eckert, D. M., Malashkevich, V. N., Hong, L. H., Carr, P. A. & Kim, P. S. Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled–coil pocket. *Cell* 99, 103–115 (1999).

Hatse, S. *et al.* Mutation of Asp171 and Asp262 of the chemokine receptor CXCR4 impairs its coreceptor function for human immunodeficiency virus-1 entry and abrogates the antagonistic activity of AMD3100. *Mol. Pharmacol.* 60, 164–173 (2001).

Kilby, J. M. *et al.* Potent suppression of HIV-1 replication in humans by T-20, a peptide nhibitor of gp41-mediated virus entry. *Nature Med.* 4, 1302–1307 (1998).

Lambert, D. M. *et al.* Peptides from conserved regions of paramyxovirus fusion (F) proteins are potent inhibitors of viral fusion. *Proc. Natl Acad. Sci. USA* 93, 2186–2191 (1996).

Palani A, Shapiro S, Clader JW, Greenlee WJ, Vice S, McCombie S, Cox K, Strizki J, Baroudy BM., Oximino-piperidino-piperidine-based CCR5 antagonists. Part 2: synthesis, SAR and biological evaluation of symmetrical heteroaryl carboxamides. *Bioorg Med Chem Lett.* 24;13(4):709-12 (2003).

Root, M. J., Kay, M. S. & Kim, P. S. Protein design of an HIV-1 entry inhibitor. *Science* 291, 884–888 (2001).

Schols, D. *et al.* Inhibition of T-tropic HIV strains by selective antagonization of the chemokine receptor CXCR4. *J. Exp. Med.* 186, 1383–1388 (1997).

Schwartz, J. A., Lium, E. K. & Silverstein, S. J. Herpes simplex virus type 1 entry is inhibited by the cobalt chelate complex CTC-96. *J. Virol.* 75, 4117–4128 (2001).

Strizki JM1, Tremblay C, Xu S, Wojcik L, Wagner N, Gonsiorek W, Hipkin RW, Chou CC, Pugliese-Sivo C, Xiao Y, Tagat JR, Cox K, Priestley T, Sorota S, Huang W, Hirsch M, Reyes GR, Baroudy BM., Discovery and characterization of vicriviroc (SCH 417690), a CCR5 antagonist with potent activity against human immunodeficiency virus type 1, *Antimicrob Agents Chemother.* Dec;49(12):4911-9 (2005).

Strizki, J. M. *et al.* SCH-C (SCH 351125), an orally bioavailable, small molecule antagonist of the chemokine receptor CCR5, is a potent inhibitor of HIV-1 infection *in vitro* and *in vivo*. *Proc. Natl Acad. Sci. USA* 98, 12718–12723 (2001).

Tagat JR, McCombie SW, Nazareno D, Labroli MA, Xiao Y, Steensma RW, Strizki JM, Baroudy BM, Cox K, Lachowicz J, Varty G, Watkins R, Piperazine-based CCR5 antagonists as HIV-1 inhibitors. IV. Discovery of 1-[(4,6- dimethyl -5-pyrimidinyl) carbonyl]- 4-[4-[2-methoxy-1(R)-4-(trifluoromethyl) phenyl] ethyl-3(S)-methyl -1-piperazinyl]- 4-methylpiperidine (Sch-417690/Sch-D), a potent, highly selective, and orally bioavailable CCR5 antagonist, *J Med Chem.*; 6;47(10):2405-8 (2004).

