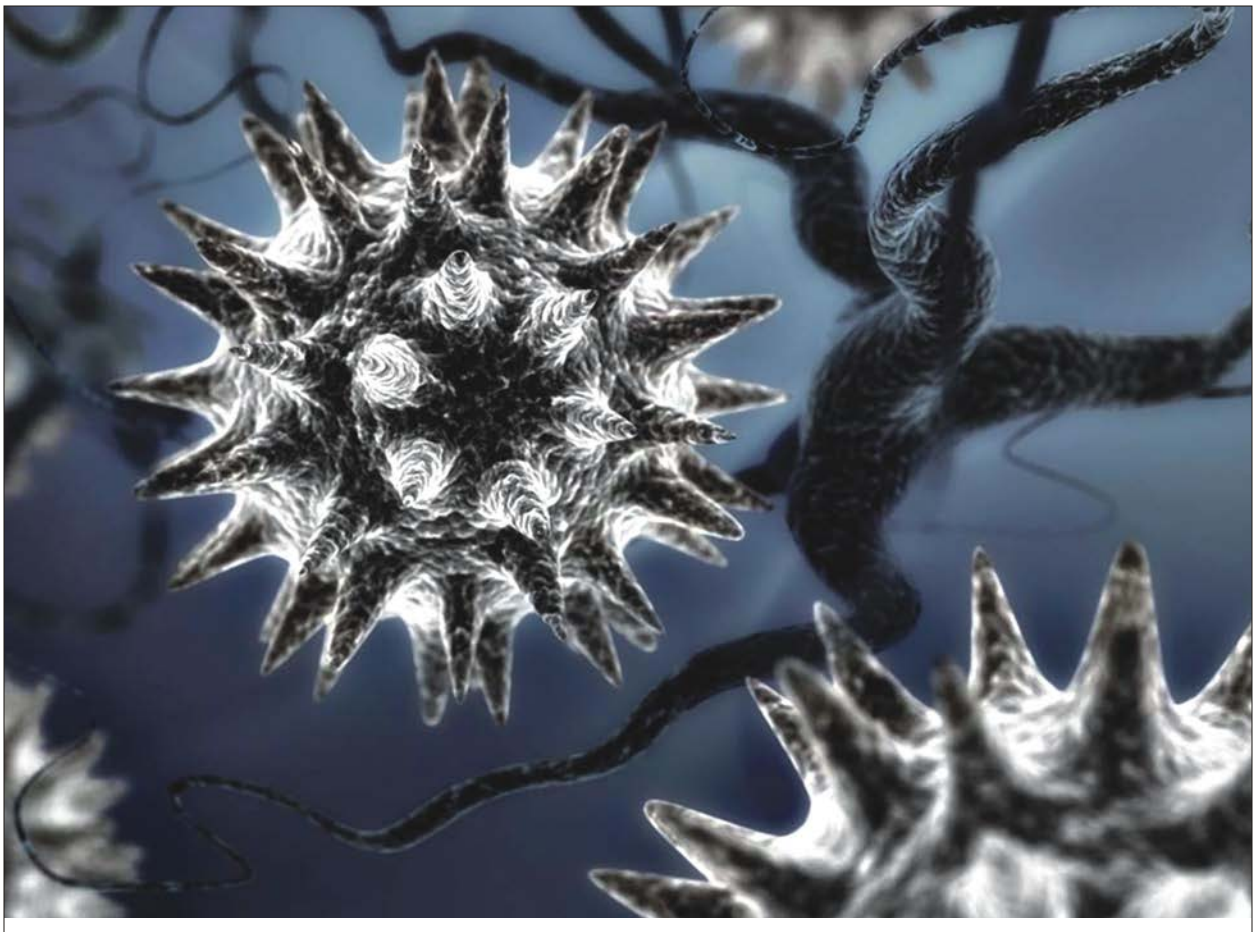


# International Journal of Virology

ISSN 1816-4900

Academic Journals Inc.

www.academicjournals.com

International Journal of Virology 3 (1): 19-27, 2007
ISSN 1816-4900


# QSAR Analysis of HIV-1 Reverse Transcriptase Inhibitory 5-Alkyl-2- [(Aryl and Alkyloxylcarbonylmethyl) Thio]-6- (1-Napthylmethyl) Pyrimidin-4 (3H)-Ones

[1]C. Karthikeyan, [1]Dengale Santosh,
[2]N.S. Hari Narayana Moorthy and [1]Piyush Trivedi
[1]Department of Pharmacy, Drug Design Laboratory,
Shri. G.S. Institute of Technology and Science, 23 Park Road, Indore-452003, India
[2]School of Pharmaceutical Sciences, Rajiv Gandhi Technical University,
Airport Bypass Road, Gandhi Nagar, Bhopal-462036, India

**Abstract:** QSAR analysis of a novel set of HIV-1 reverse transcriptase inhibitors of S-DABO series was investigated by using QuaSAR descriptors of MOE. The MMFF94 force field with root mean square gradient of 0.01 kcal $mol^{-1}$ Å was used to energy minimize the compounds. Correlation between reported biological activity values and QuaSAR descriptors was established by multiple linear regression analysis method. The generated correlations were found to be statistically significant and exhibited good predictive power. The results obtained from the QSAR study reveal that substituents with a permanent dipole and electron-releasing capacity will increase the HIV-1 RT binding affinity of S-DABO derivatives. The findings of the study suggest that HIV-1 inhibitory activity of S-DABO derivatives is dependent on the electronic properties and shape of the molecules.

**Key words:** QSAR, S-DABO, HIV-1 reverse transcriptase inhibitors, NNRTIs, MOE

## Introduction

Human Immunodeficiency Virus (HIV) infection, with its clinical progression to AIDS, is one of the leading causes of morbidity and mortality in the world (Piot *et al*., 2001). The delineation of the lifecycle of Human Immunodeficiency Virus has shown that the virus requires the catalytic activity of three unique enzymes namely protease, integrase and reverse transcriptase for its replication (Darke and Huff, 1994). Among them, Reverse Transcriptase (RT) is the key enzyme which plays an essential and multifunctional role in the replication of the Human Immunodeficiency Virus (HIV) (Jonckheere *et al*., 2000) and thus represents an attractive target for the development of new drugs useful in AIDS therapy (De Clercq, 2000). RT is responsible for the synthesis of double stranded viral DNA from proviral RNA for subsequent incorporation into the host cell chromosomes.

The currently available RT inhibitors can be classified into two groups; Nucleoside reverse transcriptase inhibitors (NRTIs), which act as a chain terminators to block the elongation of HIV-1 viral DNA (De Clercq, 1995) strand, non-nucleoside reverse transcriptase inhibitors (NNRTI), which directly inhibit reverse transcriptase enzyme by binding to the allosteric site, near the polymerase active site (De Clercq, 1998). Non-nucleoside inhibitors of this enzyme (NNRTIs) are especially attractive drug candidates because they do not function as chain terminators and do not bind at the dNTP site making them less likely to interfere with the normal function of other DNA polymerases and therefore less toxic than nucleoside inhibitors (NRTIs) such as AZT. Many structurally

**Corresponding Author:** Piyush Trivedi, Department of Pharmacy, Drug Design Laboratory,
Shri. G.S. Institute of Technology and Science, 23 Park Road, Indore-452003, India
Tel: +91 0731-2368582 Fax: +91 0731-2368582

distinct families of NNRTIs have been identified including TSAO (Balzarini *et al*., 1992), TIBO (Pauwels *et al*., 1990), HEPT (Tanaka *et al*., 1992), α-APAs (Pauwels *et al*., 1993), Pyridones (Goldman *et al*., 1991) and ITUs (Ludovici *et al*., 2001). The currently approved NNRTIs are nevirapine delaviridine and efavirenz (De Clercq, 2001) emirivine (MKC-442) (De Clercq, 2001), GW-420867X (Prince *et al*., 1999) and AG-1549 (Fujiwara *et al*., 1998) (S-1153) are currently being evaluated in clinical studies. Recently, it has been shown that combination of NRTIs, NNRTIs and PIs have been found to decrease HIV viral load, increase CD4 count, decrease mortality and delay disease progression, particularly in AIDS patient with advanced immune suppression (Palella *et al*., 1998). However, the efficacy of NNRTIs is seriously compromised by the emergence of mutant viral strains (De Clercq, 2002a, b) Some mutations, most notably K103N, are selected both *in vitro* and *in vivo* by most currently available NNRTIs (Bacheler, 1999). K103N is also the most frequently observed mutation among patients failing Highly Active Antiretroviral Therapy (HAART) because it confers resistance to all of the clinically approved NNRTIs.

The look out for potent NNRTIs devoid of any resistance-associated problems continues. QSAR is a powerful tool for the design of bioactive compounds and the prediction of corresponding activity with physical and chemical properties. QSAR studies have been successfully applied in many instances to guide the design of potent HIV RT inhibitors (Leonard and Roy, 2004; Gayen *et al*., 2004; Prabhakar *et al*., 2004). Related to the forgoing and in continuation of our efforts (Balaji *et al*., 2004) to develop potent HIV RT inhibitors, the present study strives to apply novel set of QuaSAR descriptors (Lin, 1997) programmed into molecular modeling software MOE (MOE, 2002) for modeling the HIV-1 RT inhibitory by a novel set of S-DABO derivatives reported by He *et al*. (2004). The rationale for selection of the series for QSAR analysis is based on the following

- The S-DABO derivatives exhibit potent activity against mutated drug resistant HIV-1 strains comparable to existing standard drugs such as nevirapine and efavirenz.
- Unlike the standard drugs, the S-DABO analogs also exhibited the capability to inhibit HIV-2 multiplication.
- Exact mechanism of action of these compounds is yet to be known.

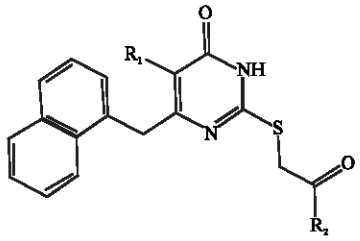
The aforementioned clearly augur well for the application of QSAR analysis on these analogs and considerable efforts were also spared to discern the factors influencing the inhibitory potency of these molecules.

## Materials and Methods

The 21 structurally diverse 5- alkyl- 2[(aryl and alkyloxy carbonyl methyl) thio]-6-(1-napthyl methyl) pyrimidine 4 (3H)-ones used in the present study were taken from the literature (He *et al*., 2004). The activity data have been reported as $IC_{50}$ values where $IC_{50}$ is the experimentally determined inhibitory concentration required to protect the cell against viral (HIV-1 $III_B$ strain) cytopathogenicity by 50% in MT-4 cells (Table 1).

The computing tools used for the present study were molecular operating environment (MOE, 2002), statistical software SYSTAT (Version 10.2) (SYSTAT 10.2, 2003) and our in-house validation program VALSTAT (VALSTAT, 2004). All the computations were carried out on Compaq PIV workstation at Drug Design laboratory, S.G.S.I.T.S., Indore, India on March 2006. Structures of compounds in the series were sketched by using builder module of MOE software and sketched the structures were subsequently energy minimized upto root mean square gradient of 0.01 kcal $mol^{-1}$ Å using MMFF94 force field. The energy-minimized structures were stored in MOE database for descriptor calculation.

Table 1: Structural variations in the S-DABO and their HIV-1 RT inhibitory activity values




| Compound No. | $R_1$ | $R_2$ | $Ic_{50}$ (μM) | -Log $IC_{50}$ (M) |
|---|---|---|---|---|
| 1 | Me | (4'-$CH_3$)Ph | 0.67 | 6.1739 |
| 2 | Me | (4'-$OCH_3$) Ph | 0.37 | 6.4317 |
| 3 | Me | $CH_3$ | 1.12 | 5.9507 |
| 4 | Me | (4'-F) Ph | 0.180 | 6.7447 |
| 5 | Me | (4'-Cl) Ph | 0.55 | 6.2596 |
| 6 | Me | $CH_3O$ | 6.99 | 5.1555 |
| 7 | Me | $CH_3CH_2O$ | 7.00 | 5.1549 |
| 8 | Et | (4'-$CH_3$)Ph | 0.32 | 6.4948 |
| 9 | Et | (4'-$OCH_3$)Ph | 0.045 | 7.3467 |
| 10 | Et | (4'-F) Ph | 0.078 | 7.1079 |
| 11 | Et | (4'-Cl) Ph | 0.26 | 6.585 |
| 12 | Et | $CH_3O$ | 4.72 | 5.326 |
| 13 | i-Pr | Ph | 0.046 | 7.3372 |
| 14 | i-Pr | (4'-$CH_3$) Ph | 0.24 | 6.6197 |
| 15 | i-Pr | (4'-$OCH_3$) Ph | 0.030 | 7.5228 |
| 16 | i-Pr | $CH_3$ | 0.41 | 6.3872 |
| 17 | i-Pr | (4'-F) Ph | 0.078 | 7.1079 |
| 18 | i-Pr | (4'-Cl) Ph | 0.32 | 6.4948 |
| 19 | i-Pr | $CH_3O$ | 4.49 | 5.3477 |
| 20 | i-Pr | $CH_3CH_2O$ | 1.41 | 5.8507 |
| 21 | i-Pr | (2',4'-$CH_3$) Ph | 0.24 | 6.6197 |

Molecular descriptors were calculated for the lowest energy conformers of the compounds in the series using the QuaSAR module of the molecular modeling software MOE. The QuaSAR module of the MOE program provides a widely applicable set of classical molecular descriptors, which can be broadly, classified into two sets, 2D and internal 3D descriptors. The two dimensional descriptors include traditional physicochemical properties, (atom counts and bond counts, mr, logP and vdw_area etc), connectivity-based topological descriptors (Kier and Hall connectivity and Kappa Shape indices; adjacency and distance matrix descriptors), pharmacophore feature descriptors (e.g., donor, acceptor, polar, positive, negative, hydrophobic), partial charge descriptors based on partial equalization of orbital electronegativities method. Quantum-chemical descriptors were also additionally chosen to account for the electronic properties of the molecules, calculated with the semiempirical PM3 Hamiltonian method as implemented by the Molecular Operating Environment program. Over 130 descriptors programmed in to MOE were calculated for each molecule in the series. However, many of the calculated descriptors such as descriptors in the PEOE_VSA-6 to PEOE_VSA+6 series, SlogP_VSA (1-9) series and SMR_VSA (1-7) series were not interpretable hence they were not used for QSAR modeling. Only, those descriptors, which could be easily interpreted, were considered for formulation of QSAR models (Table 2).

Statistical processing of the generated data was performed using statistical software SYSTAT (SYSTAT 10.2, 2003). QSAR models were constructed by multiple linear regression method following a stepwise procedure that is, only one parameter at a time was added to a model and always in the order of most significant to least significant. Statistical parameters were calculated for each step in the process so the significance of the added parameter could be verified. The quality of the regression

Table 2: Descriptors for quantitative models of HIV-RT inhibitory activity of S-DABO series

| LOGIC50 | PetitjeanSC | Kier2 | Chi1_C | PEOE_VSA_FPPOS | PM3_HOMO | PM3_dipole | Weiner path |
|---|---|---|---|---|---|---|---|
| 6.173925 | 1 | 10.7448 | 10.8121 | 0.032609 | -8.81381 | 5.414961 | 2699 |
| 6.431798 | 0.888889 | 11.42146 | 10.23475 | 0.031422 | -8.837 | 4.905065 | 3000 |
| 5.950782 | 1 | 8.589506 | 7.845772 | 0.039648 | -8.94622 | 4.584404 | 1374 |
| 6.744727 | 1 | 10.7448 | 10.23475 | 0.033698 | -8.90695 | 3.665781 | 2699 |
| 6.259637 | 1 | 10.7448 | 10.23475 | 0.032581 | -8.85145 | 4.310317 | 2699 |
| 5.155523 | 0.857143 | 9.273923 | 7.268422 | 0.080963 | -8.87888 | 3.812481 | 1562 |
| 5.154902 | 1 | 9.9723 | 7.975528 | 0.077075 | -8.85365 | 3.759203 | 1775 |
| 6.49485 | 1 | 11.42146 | 11.3501 | 0.031253 | -8.83877 | 5.383857 | 2906 |
| 7.346787 | 0.888889 | 12.10938 | 10.77275 | 0.030161 | -8.85111 | 4.952682 | 3221 |
| 7.107905 | 1 | 11.42146 | 10.77275 | 0.032251 | -8.90421 | 3.68549 | 2906 |
| 6.585027 | 1 | 11.42146 | 10.77275 | 0.031227 | -8.86869 | 4.281878 | 2906 |
| 5.326058 | 0.857143 | 9.9723 | 7.806427 | 0.077075 | -8.88118 | 3.823226 | 1712 |
| 7.337242 | 0.875 | 11.42146 | 11.32894 | 0.031152 | -8.82192 | 4.844153 | 2818 |
| 6.619789 | 1 | 11.62016 | 11.72278 | 0.029912 | -8.82414 | 5.278085 | 3115 |
| 7.522879 | 0.888889 | 12.3008 | 11.14543 | 0.02891 | -8.82533 | 5.553202 | 3444 |
| 6.387216 | 1 | 9.467456 | 8.756455 | 0.035731 | -8.92816 | 4.480147 | 1656 |
| 7.107905 | 1 | 11.62016 | 11.14543 | 0.030825 | -8.89086 | 3.518474 | 3115 |
| 6.49485 | 1 | 11.62016 | 11.14543 | 0.029888 | -8.86769 | 4.188405 | 3115 |
| 5.347754 | 0.857143 | 10.15625 | 8.179105 | 0.073281 | -8.87481 | 3.66648 | 1864 |
| 5.850781 | 1 | 10.85782 | 8.886212 | 0.070081 | -8.86162 | 3.902731 | 2099 |
| 6.619789 | 1 | 11.82315 | 12.13347 | 0.028767 | -8.81125 | 4.722422 | 3366 |

equations was adjudged by the statistical parameters such as correlation coefficient R, squared correlation coefficient $R^2$, Fischer ratio values and standard error of the estimate SEE. Guidelines for the acceptance of regressions were: The squared correlation coefficient $R^2$, above 0.7 or higher (R>0.80), minimum intercorrelation between the descriptors found in the same equation (<0.7), Fischer ratio values indicating 99% level of significance.

## Results and Discussion

Linear regression analysis of the using the biological activity parameter as dependent variable and the reduced pool of descriptors as predictor variable resulted in several correlations. The generated correlations were evaluated for statistical significance and the most significant correlations were chosen on the basis of standard test of significance and correlation coefficient. The best correlations selected are summarized below.

$$-\text{LogIC}_{50} = [4.40391(\pm 2.34397)] + \text{PEOE\_VSA\_FPPOS}\ [-23.2418\ (\pm 9.70254)] + \text{Kier2}\ [0.271892\ (\pm 0.189855)] \quad (1)$$

$$N = 21,\ R = 0.91,\ R^2 = 0.82,\ SEE = 0.316,\ F_{(2,18\ (F=5.85))} = 42.00,\ P>0.000$$

$$-\text{LogIC}_{50} = [-86.4896\ (\pm 46.2168)] + \text{Chi1\_C}\ [0.586258(\pm 0.135277)] + \text{PetitjeanSC}\ [-3.96893(\pm 2.74307)] + \text{PM3\_HOMO}\ [-10.2435(\pm 5.23658)] \quad (2)$$

$$N = 21,\ R = 0.92,\ R^2 = 0.84,\ SEE = 0.305,\ F_{(3,17\ (F=4.84))} = 30.83,\ ,\ P>0.000.$$

$$-\text{LogIC}_{50} = [-114.351(\pm 33.7159)] + \text{Weiner Path}\ [0.00112734\ (\pm 0.000185163)] + \text{PM3\_HOMO}\ [-13.1177(\pm 3.73701)] + \text{PM3\_DIPOLE}\ [0.340935\ (\pm 0.184433)] \quad (3)$$

$$N = 20,\ R = 0.963789,\ R^2 = 0.928889,\ SEE = 0.203038,\ F_{(3,16\ (F=4.94))} = 69.66,\ P>0.000$$

In the equations, n is the number of molecules. R is the correlation coefficient, $R^2$ is the squared correlation coefficient, SEE is the standard error of estimate, F is the Fischer ratio values at 99% confidence levels and p-value is the significance level. The figures within the parentheses are 95% confidence limits.

Models 1-3 manifests good statistical quality and explains more than 80% of variance in the biological activity as established by high squared correlation coefficient ($R^2$>0.8). Further, low values of standard error of estimate indicate accuracy of the statistical fit. The F-test values are significant at of the correlations exceeds the tabulated F-value (given in parentheses of the calculated F values) by a large margin as desired in linear regression. The p-values less than 0.000 also indicate that there is indeed a significant relationship between the predictor variables and dependant variables in the selected correlations.

Absence of collinear descriptors in the selected correlations was established by calculation of correlation matrix (Table 3) and Variance Inflation Factor (VIF) values (Table 4). VIF value (Cho *et al.*, 2001) was calculated from $1/1-R^2$, where $R^2$ is the multiple correlation coefficient of one descriptor's effect regressed on the remaining molecular descriptors. VIF values larger than 5 indicates that the information of the descriptors may be hidden by the correlation of the descriptors. A perusal of the correlation matrix and VIF values of descriptors recorded in Table 3 and 4, which shows that the descriptors used in the regressions are reasonably orthogonal to each other.

The biparametric model (Eq. 1) includes the partial charge descriptor PEOE_VSA_FPPOS (Lin, 1997) and topological descriptor Kier2 (Hall and Kier, 1991). The partial charge descriptor PEOE_VSA_FPPOS represents the fractional positive polar vander Waals surface area of the molecule. Mathematically, it can be defined as the sum of vander Waals surface area (vi) such that partial charge (qi) is greater than 0.2 divided by the total surface area. The descriptor PEOE_VSA_FPPOS takes a negative weight in the correlation, which suggests that increase in the molecular surface area bearing a polar positive charge will decrease the HIV-1 RT inhibitory potency of S-DABO derivatives. Development of a fractional positive polar partial charge on an atom is always associated with electron withdrawal by electronegative atom in its immediate vicinity. The aforementioned fact point towards the alkoxy substituents in the $R_2$ position of pyrimidine ring as they form ester linkage, which leaves partial positive polar charge on the alkyl groups because of electron withdrawing ability of the carboxyl moiety. The observation also leads to hypothesis that presence of polar positive partial charges on the alkyl groups somehow impedes the interaction of alkyl substituents with its complementary group in the enzyme. The topological descriptor Kier2 denotes Kier's kappa shape index, which encodes information related to the degree of star graph-likeness and linear graph likeness of the molecule. The descriptor values are higher for a linear molecule and decreases with branching in the molecule. Thus, the positive coefficient of the descriptor in model 1 implies that non-branched molecule will exhibit better HIV-1 RT inhibitory activity than branched counterparts.

The triparametric model (Eq. 2) comprises of two topological descriptors Chi1_C (Hall and Kier, 1991; Kier and Hall, 1977), petitjeanSC (Petitjean, 1992) and a quantum chemical descriptor PM3_HOMO (Mati and Lobanov, 1996). The topological descriptor Chi1_C refers to Kier and Halls carbon connectivity index of order 1. In general, the descriptor encodes information regarding degree of branching, cyclization in the molecule. Mathematically, it can be defined as

$$\text{Chi1\_C} = \Sigma(\delta_i\delta_j)^{-1/2} \quad (4)$$

Where $\delta_i$ and $\delta_j$ are the vertex connectivity degree of carbon atoms i and j, respectively and the summation extends to all bonded pairs of non hydrogen carbon atoms in the group or molecule.

The value of the $\delta_i$ increase with branching in the molecule. Thus, the positive coefficient of the descriptor Chi1_C in Eq. 2 suggests that non-branched S_DABO derivatives will have increased HIV-1 RT inhibitory potency. The topological parameter Petitjean Shape coefficient bears a negative weight in model 2 which suggest that molecular shape is an important determinant in the binding of S-DABO derivatives to HIV-1 Reverse transcriptase. The quantum chemical descriptor PM3_HOMO in the

Table 3: Correlation matrix showing the inter-correlation of molecular descriptors used in models

| | LOG IC50 | Petitjean SC | Kier2 | Chi1_C | PEOE_ VSAS FPPO | PM3_ HOMO | PM3_ dipole | Weiner path |
|---|---|---|---|---|---|---|---|---|
| LOGIC50 | 1.00 | 0.13 | 0.76 | 0.82 | -0.86 | 0.21 | 0.43 | 0.81 |
| PetitjeanSC | 0.13 | 1.00 | 0.04 | 0.31 | -0.40 | -0.17 | -0.02 | 0.15 |
| Kier2 | 0.76 | 0.04 | 1.00 | 0.88 | -0.61 | 0.64 | 0.39 | 0.96 |
| Chi1_C | 0.82 | 0.31 | 0.88 | 1.00 | -0.84 | 0.56 | 0.52 | 0.95 |
| PEOE_VSA_FPPOS | -0.86 | -0.40 | -0.61 | -0.84 | 1.00 | -0.19 | -0.56 | -0.76 |
| PM3_HOMO | 0.21 | -0.17 | 0.64 | 0.56 | -0.19 | 1.00 | 0.60 | 0.59 |
| PM3_dipole | 0.43 | -0.02 | 0.39 | 0.52 | -0.56 | 0.60 | 1.00 | 0.45 |
| Weiner path | 0.81 | 0.15 | 0.96 | 0.95 | -0.76 | 0.59 | 0.45 | 1.00 |

Table 4: VIF values descriptors in generated correlations

| Model No. | Intercept/descriptors | VIF |
|---|---|---|
| 1 | PEOE_VSA_FPPOS | 1.579 |
| | Kier2 | 1.579 |
| 2 | Chi1_C | 1.949 |
| | Petitjean SC | 1.369 |
| | PM3_HOMO | 1.808 |
| 3 | Weinerpath | 1.584 |
| | PM3_DIPOLE | 1.956 |
| | PM3_HOMO | 1.592 |

model 2 denotes energy associated highest occupied molecular orbital and can be related to the ionization potential of the molecule. The coefficient of descriptor bears a negative sign in model 2, the negative values of HOMO, which corresponds to more electron releasing group favors HIV-1 reverse transcriptase inhibitory activity of S- DABO derivatives.

Another triparametric model (Eq. 3) of good statistical quality was obtained with the following descriptors; Weiner Path (Wiener, 1947), PM3_HOMO (Mati and Lobanov, 1996) and PM3_DIPOLE (Mati and Lobanov, 1996). Weiner Path descriptor is contributing towards the activity in the first model. Wiener Path index is defined as the half the sum of all entries in a distance matrix.

$$W = \tfrac{1}{2} \sum_{i} \sum_{j} d_{ij}$$

Wiener path index is a global descriptor and has contributions from all the atoms of the molecule. It is inversely related to the degree of compactness of the molecule and decreases with increase in the branching and cyclicity of the molecules. Thus, the positive coefficient of the descriptor Wiener Path against colon carcinoma cells in model 1 suggest decreased branching in the side chain and resultant increase in its flexibility is conducive HIV-1 RT inhibitory activity of S-DABO derivatives. The negative coefficient of the quantum chemical descriptor PM3_HOMO reinforces the conclusion drawn from Eq. 2. Interestingly, another quantum chemical descriptor PM3_DIPOLE bears a positive weight in the Eq. 3. The descriptor PM3_DIPOLE accounts for dipole-dipole interaction between functional groups at R2 and the receptor. The positive coefficient associated with this descriptor suggests that charge distribution in the molecule is related with binding affinity of the molecules to the enzyme.

Compound number 13 was found to be an outlier in Eq. 3 on account of large deviation of calculated activity from the experimentally determined value (studentized residual = 3.08). The outlying behavior of the compound is not immediately apparent and merits further studies.

The predictive power of the generated correlations was evaluated by cross validation method following a 'leave-one-out' scheme using inhouse program VALSAT. The reliability of the correlations was tested in a cross validation with the determination of $r^2cv$ (cross validated $r^2$) or $q^2$. In this method, one data point is removed systematically from the dataset and a QSAR model is constructed on the basis of reduced dataset and subsequently used to predict the activity of the removed data point. This

Table 5: Experimental (-Log $IC_{50}$) and predicted activity values (Model 1-3) for HIV-1 RT inhibition

| Compound No. | -Log $IC_{50}$ | Model 1 | Model 2 | Model 3 |
|---|---|---|---|---|
| 1 | 6.173925 | 6.59995 | 6.16218 | 6.14915 |
| 2 | 6.431798 | 6.80371 | 6.51535 | 6.64407 |
| 3 | 5.950782 | 5.68086 | 5.7062 | 6.30447 |
| 4 | 6.744727 | 6.52659 | 6.78717 | 6.7896 |
| 5 | 6.259637 | 6.59364 | 6.20688 | 6.27312 |
| 6 | 5.155523 | 5.00416 | 5.38603 | 5.18697 |
| 7 | 5.154902 | 5.36912 | 4.74386 | 5.03947 |
| 8 | 6.49485 | 6.80355 | 6.76421 | 6.74715 |
| 9 | 7.346787 | 6.9454 | 6.88864 | 7.03741 |
| 10 | 7.107905 | 6.73555 | 7.05764 | 6.94514 |
| 11 | 6.585027 | 6.79776 | 6.71408 | 6.73378 |
| 12 | 5.326058 | 5.32339 | 5.76107 | 5.3943 |
| 13 | 7.337242 | 6.74567 | 6.96025 | - |
| 14 | 6.619789 | 6.88906 | 6.83575 | 6.7295 |
| 15 | 7.522879 | 6.99686 | 6.78883 | 7.09613 |
| 16 | 6.387216 | 6.06019 | 6.06444 | 6.05308 |
| 17 | 7.107905 | 6.82531 | 7.15812 | 6.93816 |
| 19 | 6.49485 | 6.90021 | 6.95761 | 6.97106 |
| 19 | 5.347754 | 5.48718 | 5.93629 | 5.4322 |
| 20 | 5.850781 | 5.69529 | 5.45071 | 5.54956 |
| 21 | 6.619789 | 6.98389 | 6.98438 | 6.64135 |

Table 6: Comparison of cross validation parameters for generated QSAR models

| Model No. | $q^{2a}$ | $S_{PRESS}{}^{b}$ | SDEP[c] |
|---|---|---|---|
| 1 | 0.776483 | 0.356481 | 0.330037 |
| 2 | 0.758418 | 0.381351 | 0.343115 |
| 3 | 0.885757 | 0.257348 | 0.230179 |

[a] = Squared correlation coefficient of prediction. [b] = Standard deviation of prediction. [c] = Standard error of prediction

procedure is repeated until a complete set of predicted activities. The correlation coefficient between the experimentally determined activity and activity predicted by LOO method is calculated (cross validated $r^2$) (Table 5). High values of $q^2$ considered as proof of high predictive ability of the models. It is worth mentioning that all the generated correlations exhibits good predictive ability as established by high $q^2$ values (>0.6) and the best being recorded for biparametric Eq. 3. The $q^2$ values for the obtained correlations are given in the Table 5.

Further confirmation on predictive ability of the correlations was obtained by determining the uncertainty in the prediction ($S_{PRESS}$) and standard error due to prediction (SDEP). The $S_{PRESS}$ and SDEP values should be low for a regression equation to have good predicivity. The $S_{PRESS}$ and SDEP values for the correlations obtained are presented in the Table 6.

## Conclusions

Finally to conclude, the QSAR analysis of a series of 21 HIV-1 RT inhibitory S-DABO derivatives using a novel set of QuaSAR descriptors resulted in quantitative models of good statistical significance. The generated QSAR models also showed good predictive potential as established by their high $q^2$ values (>0.7) and hence can be used in the prediction of biological activity of novel molecules prior to their synthesis. Further, the QSAR study suggest that HIV-1 inhibitory activity of S-DABO derivatives is related to the electronic properties and topology of the molecule. From the results of the QSAR study, it appears that substituents with a permanent dipole and electron-releasing capacity will increase the HIV-1 RT binding affinity of S-DABO derivatives. Additionally, the study also indicates that molecular branching and increase in molecular surface area bearing a polar positive partial charge decreases the HIV-1 inhibitory potency of S-DABO derivatives

**Acknowledgements**

The authors gratefully acknowledge Prof. Lemont B. Kier, Emeritus Faculty of Medicinal Chemistry, Virginia Commonwealth University and Prof. Michael Petitjean, p for providing reprints of their research publications. The authors wish to thank TATA ELXSI for providing molecular modeling software MOE for the present study.

**References**

Bacheler, L.T., 1999. Resistance to non-nucleoside inhibitors of HIV-1 reverse transcriptase. Drug Resist. Updates, 2: 56-67.

Balaji, S., C. Karthikeyan, N.S. Hari Narayan Moorthy and P. Trivedi, 2004. QSAR modelling of HIV-1 reverse transcriptase inhibition by benzoxazinones using a combination of P_VSA and pharmacophore feature descriptors. Bioorg. Med. Chem. Lett., 14: 6089-6094.

Balzarini, J., M.J. Pe´rez-Pe´rez, A. San-Fe´lix, S. Vela´zquez, M.J. Camarasa and E. De Clercq, 1992. [2',5'-Bis-*O*-(*tert*-butyldimethylsilyl)]- 3'-spiro-5"-(4"-amino-1",2"-oxathiole-2",2"-dioxide) (TSAO) derivatives of purine and pyrimidine nucleosides as potent and selective inhibitors of human immunodeficiency virus type 1. Antimicrob. Agents Chemother., 36: 1073-1080.

Cho, D.H., S.K. Lee, B.T. Kim and K.T. No, 2001. Quantitative Structure-Activity Relationship (QSAR) study of new fluorovinyloxyacetamides. Bull. Korean Chem. Soc., 22: 388-394.

Darke, P.L. and J.R. Huff, 1994. Advances in Pharamacology. Augest, J., M.W. Ander, F. Murad (Eds.), San Diego, Academic, pp: 399-415.

De Clercq, E., 1995. Toward improved anti-HIV chemotherapy: Therapeutic strategies for intervention with HIV infections. J. Med. Chem., 38: 2491-2517.

De Clercq, E., 1998. The role of non-nucleoside reverses transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection. Antiviral Res., 38: 153-179.

De Clercq, E., 2000. Novel compounds in preclinical/early clinical development for the treatment of HIV infections. Rev. Med. Virol., 10: 255-277.

De Clercq, E., 2001. New developments in anti-HIV chemotherapy. Il Farmaco, 56: 3-12.

De Clercq, E., 2002a. New anti-HIV agents and targets. Med. Res. Rev., 22: 531-565.

De Clercq, E., 2002b. New developments in anti-HIV chemotherapy. Biochim. Biophys. Acta, 1587: 258-275.

Fujiwara, T., A. Sato, M. El-Farrash, S. Miki and K. Abe *et al.*, 1998. S-1153 inhibits replication of known drug-resistant strains of human immunodeficiency virus type 1. Antimicrob. Agents Chemother., 42: 1340-1345.

Gayen, S., B. Debnath, S. Samantha and T. Jha, 2004. QSAR study on some anti-HIV HEPT analogues using physicochemical and topological parameters. Biorg. Med. Chem., 12: 1493-1503.

Goldman, M.E., J.H. Nunberg, J.A. O'Brien, J.C. Quintero and W.A. Schleif *et al.*, 1991. Pyridone derivatives: Specific human immunodeficiency virus type 1 reverse transcriptase inhibitors with antiviral activity. Proc. Natl. Acad. Sci. USA., 88: 6863-6867.

Hall, L.H. and L.B. Kier, 1991. The molecular connectivity chi indices and kappa shape indices in structure-property modeling. Rev. Comput. Chem., Vol. 2.

He, Y., C. Fener, S. Guangfu, W. Yueping, E. De Clercq, J. Balzarini and C. Pannecouque, 2004. 5-Alkyl-2-[(aryl and alkyloxylcarbonylmethyl)thio]-6- (1-naphthylmethyl) pyrimidin-4(3H)-ones as an unique HIV reverse transcriptase inhibitors of S-DABO series. Bioorg. Med. Chem. Lett., 14: 3173-3176.

Jonckheere, H., J. Anne and E. De Clercq, 2000. The HIV-1 reverse transcription (RT) process as target for RT inhibitors. Med. Res. Rev., 20: 129-154.

Kier, L.B. and L.H. Hall, 1977. Structure-activity studies on hallucinogenic amphetamines using molecular connectivity. J. Med. Chem., 20: 1631-1636.

Leonard, J. and K. Roy, 2004. Classical QSAR Modeling of HIV-1 Reverse Transcriptase Inhibitor 2-Amino-6-arylsulfonylbenzonitriles and Congeners. QSAR Combi. Sci., 23: 23-35.

Lin, A., 1997. QuaSAR-descriptors, Journal of Chemical Computing Group. http://www. chemcomp. com/Journal_of_CCG/Features/descr.htm.

Ludovici, D.W., M.J. Kukla, P.G. Grous, S. Krishnan and K. Andries *et al*., 2001. Evolution of anti-HIV drug candidates. Part 1: From R-Anilinophenylacetamide (RAPA) to imidazolyl thiourea (ITU). Bioorg. Med. Chem. Lett., 11: 2225-2228.

Mati, K. and V.S. Lobanov, 1996. Quantum-chemical descriptors in QSAR/QSPR studies. Chem. Rev., 96: 1027-1043.

MOE., 2002. Molecular modeling package developed by chemical Computing Group Inc., Canada.

Palella, F.J., K.M. Delaney, A.C. Moorman, M.O. Loveless, J. Fuhrer, G.A. Satten, D.J. Aschman and S.D. Holmberg, 1998. For the HIV outpatient study investigators. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. N. Eng. J. Med., 338: 853-860.

Pauwels, R., K. Andries, J. Desmijter, D. Schols and M.J. Kukla *et al*., 1990. Potent and selective inhibition of HIV-1 replication *in vitro* by a novel series of TIBO derivatives. Nature, 343: 470-474.

Pauwels, R., K. Andries, Z. Debryser, P. Van Daele and D. Schols *et al*., 1993. Potent and highly selective human immunodeficiency virus type 1 (HIV-1) inhibition by a series of R-anilinophenylacetamide derivatives targeted at HIV-1 reverse transcriptase. Proc. Natl. Acad. Sci. USA., 90: 1711-1715.

Petitjean, M., 1992. Applications of the radius-diameter diagram to the classification of topological and geometrical shapes of chemical compounds. J. Chem. Inf. Comput. Sci., 32: 331-337.

Piot, P., M. Bartos, P.D. Ghys, N. Walker and B. Schwartlander, 2001. The global impact of HIV/AIDS. Nature, 410: 968-973.

Prabhakar, Y.S., V.R. Soloman, R.K. Rawal, M.K. Gupta and S.B. Katti, 2004. CP-MLR/PLS Directed structure-activity modeling of the HIV-1 RT inhibitory activity of 2,3-Diaryl-1,3-thiazolidin-4-ones. QSAR Combi. Sci., 23: 234-244.

Prince, W., K. Moore, L. Cass, N. Dallow, A. Jones, J.P. Kleim, P. Mutch and M. St. Clair, 1999. GW420867X, a new non-nucleoside reverse transcriptase inhibitor (NNRTI)-initial phase I evaluation. Abstracts of the 12th International Conference on Antiviral Research, Jerusalem, Israel, March 21-25, 41: A48-A50.

SYSTAT 10.2, 2003. A statistical package developed by SYSTAT software Inc, USA, 2003.

Tanaka, H., H. Takashima, M. Ubasawa, K. Sekiya, I. Nitta, M. Baba, S. Shigeta, R.T. Walker, E. De Clercq and T. Miyasaka, 1992. Synthesis and antiviral activity of 1-[(2-hydroxyethoxy)-methyl]-6-(phenylthio)thymine (HEPT) as potent and selective anti-HIV-1 agents. J. Med. Chem., 35: 4713-4719.

VALSTAT., 2004. A Statistical program developed by Dr. Arun Kumar Gupta, SGSITS, Indore.

Wiener, H., 1947. Structural determination of paraffin boiling points. J. Am. Chem. Soc., 69: 17-20.