

QSAR MODELING OF HIV-1 REVERSE TRANSCRIPTASE INHIBITORY ACTIVITY WITH PETT DERIVATIVES

V. Ravichandran^a, V. K. Mourya^b, R. K. Agrawal^{a*}

^aPharmaceutical Chemistry Research Laboratory, Dept. of Pharmaceutical Sciences,
Dr. Hari Singh Gour University, Sagar (M.P) - 470 003, India.

^bGovt. College of Pharmacy, Osmanpura, Aurangabad, Maharashtra, India.

In pursuit of better anti-HIV agents, QSAR studies were performed on a series of phenyl ethyl thiourea (PETT) analogues using WIN CAChe 6.1. Stepwise multiple linear regression analysis was performed to derive QSAR models which were further evaluated for statistical significance and predictive power by internal and external validation. The best QSAR model was selected, having correlation coefficient (r) = 0.835 and cross validated squared correlation coefficient (q^2) = 0.642. The QSAR model indicates that the molecular weight, valence connectivity index and critical pressure play an important role in the HIV-1 RT inhibitory activities. The results of the present study may be useful on the designing of more potent PETT analogues as HIV-1 RT inhibitory agents.

Keywords: QSAR; HIV-1 RT inhibitory activity; multiple linear regressions, phenyl ethyl thiourea.

(Received January 22, 2008; accepted January 26, 2008)

1. Introduction

HIV-1 (Human Immunodeficiency Virus Type-1) is the pathogenic retrovirus and causative agent of AIDS or AIDS-related complex (ARC) [1]. When viral RNA is translated into a polypeptide sequence, it is assembled in a long polypeptide chain, which includes several individual proteins namely, reverse transcriptase, protease, integrase, etc. Before these enzymes become functional, they must be cut from the longer polypeptide chain.

Acquired immune deficiency syndrome (AIDS) is a formidable pandemic that is still wreaking havoc world wide. The catastrophic potential of this virally caused disease may not have been fully realized. The causative moiety of the disease is human immunodeficiency virus (HIV), which is a retrovirus of the lentivirus family [2]. The three viral enzymes; reverse transcriptase, protease and integrase encoded by the gag and gag-pol genes of HIV play an important role in the virus replication cycle. Among them, viral reverse transcriptase (RT) catalyzes the formation of proviral DNA from viral RNA, the key stage in viral replication. Its central role in viral replication makes RT a prime target for anti-HIV-therapy [3].

Two main categories of HIV RT inhibitors have been discovered to date. The first category of inhibitors is nucleoside analogues (e.g., AZT, 3TC, ddI, ddC) and the second category of inhibitors is nonnucleoside analogues. Nevirapine, delaviridine and efavirenz are the only nonnucleoside reverse transcriptase inhibitors (NNRTI) that have received regulatory approval with several NNRTIs (MKC442, Troviridine, S-1153/ AG1549. PNU142721, ACT and HBY1293/GW420867X) are currently undergoing clinical trials. Efavirenz was the first potent anti-HIV drug to be approved by FDA and studies have shown that efavirenz penetrates into the cerebrospinal fluid, a common viral sanctuary. The therapeutic efficacy of the drug is mainly restricted due to the development of viral resistance associated with mutations that includes K103N, L1001 and Y188L [4].

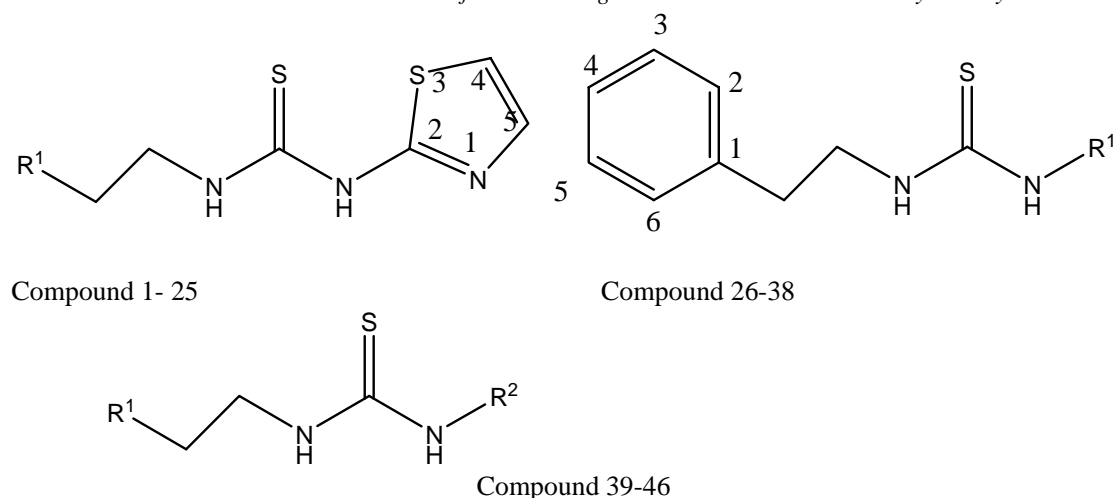
QSAR analyses of HIV-1 reverse transcriptase inhibitors [5], HIV-1 protease inhibitors [6,7] and HIV-1 integrase inhibitors [8] and gp 120 envelope glycoprotein [9] were reported. The present group of authors has developed a few quantitative structure-activity relationship models to predict anti-HIV activity of different group of compounds [10-16]. In continuation of such efforts, in this article, we have performed QSAR analysis

* Corresponding author: dragrawal2001@yahoo.co.in

to explore the correlation between physicochemical and biological activity of PET derivatives using modeling software WIN CACHe 6.1 (molecular modeling software, a product of Fujitsu private limited, Japan) and statistical software STATISTICA version 6 (StatSoft, Inc., Tulsa, USA).

In the present work we have taken 71 PETT compounds and their HIV-1 RT inhibitory activity from the reported work [17,18]. Many of these compounds inhibited wild type HIV-1 RT with IC_{50} values between 0.001 μ M and 3.9 μ M. There is high structural diversity and a sufficient range of the biological activity in the selected series of PET derivatives. It insists as to select these series of compounds for our QSAR studies. All the HIV-1 RT inhibitory activities used in the present study were expressed as $pIC_{50} = -\log_{10} IC_{50}$. Where IC_{50} is the micro molar concentration of the compounds producing 50% reduction in the HIV-1 RT activity is stated as the means of at least two experiments. The compounds which did not show confirmed HIV-1 RT inhibitory activity in the above cited literature have not been taken for our study. We carried out QSAR analysis and established a QSAR model to guide further structural optimization and predict the potency and physicochemical properties of clinical drug candidates.

Table 1. Structures of PETT analogs and their HIV-1 RT inhibitory activity



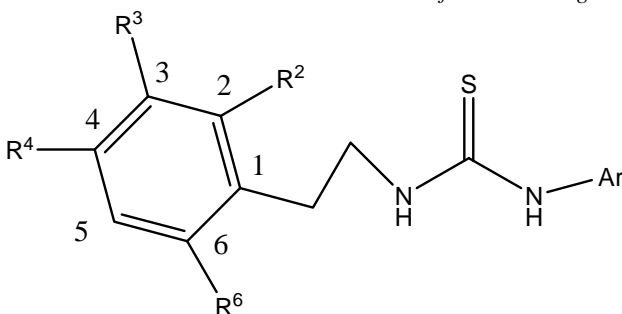
| Comp No | R ¹ | R ² | pIC ₅₀ (μ M) |
|-----------------|--------------------------|----------------|------------------------------|
| | | | Experimental ^b |
| 1 | Phenyl | - | 0.0457 |
| 2 | 2-fluorophenyl | - | 1.222 |
| 3 ^a | 3-fluorophenyl | - | 0.824 |
| 4 | 4-fluorophenyl | - | 0 |
| 5 ^a | 2-methoxyphenyl | - | 1.398 |
| 6 | 3-methoxyphenyl | - | 0.824 |
| 7 | 4-methoxyphenyl | - | 0.455 |
| 8 | 2-methylphenyl | - | 1.096 |
| 9 | 2-nitrophenyl | - | 0.824 |
| 10 | 2-hydroxyphenyl | - | -0.041 |
| 11 | 2-chlorophenyl | - | 0.222 |
| 12 | 3-ethoxyphenyl | - | 1.221 |
| 13 | 3-propoxyphenyl | - | 0.698 |
| 14 ^a | 3-isopropoxyphenyl | - | 0.398 |
| 15 | 3-phenoxyphenyl | - | -0.041 |
| 16 | 2,6-dimethoxyphenyl | - | 1.046 |
| 17 | 2,5-dimethoxyphenyl | - | 0.699 |
| 18 | 3-bromo-6-methoxyphenyl | - | 1.522 |
| 19 | 2-fluoro-6-methoxyphenyl | - | 2 |
| 20 ^a | 2-ethoxy-6-fluorophenyl | - | 2 |
| 21 | 2,6-difluorophenyl | - | 2.221 |

| | | | |
|-----------------|----------------------------------|-----------------------|--------|
| 22 | 2-chloro-6-fluorophenyl | - | 0.698 |
| 23 | 2-pyridyl | - | -0.279 |
| 24 | 3-pyridyl | - | 0.187 |
| 25 | 2-furyl | - | 1 |
| 26 | 4-methylthiazol-2-yl | - | 0.221 |
| 27 | 4-ethylthiazol-2-yl | - | 0.455 |
| 28 | 4-propylthiazol-2-yl | - | 0.698 |
| 29 | 4-isopropylthiazol-2-yl | - | -0.398 |
| 30 ^a | 4-butylthiazol-2-yl | - | 0.698 |
| 31 | 4-cyanothiazol-2-yl | - | 0.259 |
| 32 | 4-(trifluoro methyl)thiazol-2-yl | - | 0.698 |
| 33 | 4-(ethoxy carbonyl)thiazol-2-yl | - | -0.380 |
| 34 | 5-chlorothiazol-2-yl | - | -0.278 |
| 35 | 1,3,4-thiazol-2-yl | - | -0.591 |
| 36 | 2-pyridyl | - | 1.698 |
| 37 | 5-bromo-2-pyridyl | - | 1.823 |
| 38 | 5-methyl-2-pyridyl | - | 1.522 |
| 39 ^a | 2,6-difluorophenyl | 4-cyano thiazoly-2-yl | 0.698 |
| 40 | 2,6-difluorophenyl | 5-bromo-2-pyridyl | 2.259 |
| 41 | 2,6-difluorophenyl | 5-methyl-2-pyridyl | 2.222 |
| 42 ^a | 2-ethoxy-6-fluorophenyl | 5-methyl-2-pyridyl | 3 |
| 43 | 2-ethoxy-6-fluorophenyl | 5-bromo-2-pyridyl | 2.522 |
| 44 | 2-pyridyl | 5-methyl-2-pyridyl | 2.346 |
| 45 | 2-pyridyl | 5-bromo-2-pyridyl | 2.221 |
| 46 | 2,6-difluorophenyl | 4-ethylthiazol-2-yl | 1.301 |

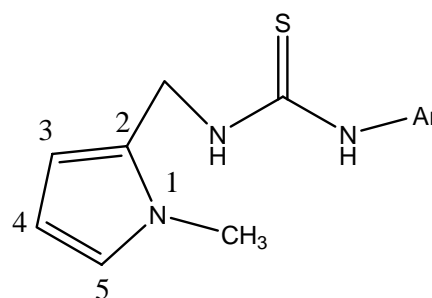
a - test set compounds

b - the experimental IC₅₀ values (in micro molar) were converted into -logIC₅₀ (pIC₅₀, in micro molar).

Table 2. Structures of PETT analogs and their HIV-1 RT inhibitory activity.



Compound 47-69



Compound 70 & 71

| Comp. No. | R ² | R ³ | R ⁴ | R ⁶ | Ar | pIC ₅₀ (μM) |
|-----------------|----------------|----------------------------------|----------------|----------------|--------------------|---------------------------|
| | | | | | | Experimental ^b |
| 47 | F | (CO)N(Me) ₂ | H | F | 5-bromo-2-pyridyl | 1.823 |
| 48 | F | CH ₂ NAc | H | F | 5-bromo-2-pyridyl | 3 |
| 49 | F | CN | H | F | 5-chloro-2-pyridyl | 1.638 |
| 50 | F | N(Me) ₂ | H | F | 5-chloro-2-pyridyl | 1.346 |
| 51 ^a | F | N(Me) ₂ | H | F | 5-bromo-2-pyridyl | 2.045 |
| 52 | F | OCH ₃ | H | F | 5-bromo-2-pyridyl | 2.096 |
| 53 ^a | F | OC ₂ H ₅ | H | F | 5-bromo-2-pyridyl | 2.154 |
| 54 | F | CH ₂ OCH ₃ | H | F | 5-bromo-2-pyridyl | 2.221 |

| | | | | | | |
|-----------------|------------------|--------------------------------|----|--------------------------------|--------------------|-------|
| 55 | Cl | OC ₂ H ₅ | H | F | 5-bromo-2-pyridyl | 2.397 |
| 56 | Cl | OC ₂ H ₅ | H | F | 5-chloro-2-pyridyl | 2.397 |
| 57 | Cl | OC ₂ H ₅ | H | F | 5-iodo-2-pyridyl | 1.921 |
| 58 | Cl | OC ₂ H ₅ | H | F | 5-cyano-2-pyridyl | 2.221 |
| 59 | H | OCH ₃ | H | OCH ₃ | 5-chloro-2-pyridyl | 2.096 |
| 60 | H | OC ₂ H ₅ | H | OC ₂ H ₅ | 5-bromo-2-pyridyl | 2.301 |
| 61 | F | H | H | OC ₂ H ₅ | 5-bromo-2-pyridyl | 2.301 |
| 62 | F | F | H | OC ₂ H ₅ | 5-bromo-2-pyridyl | 1.745 |
| 63 ^a | F | F | H | OCH ₃ | 5-bromo-2-pyridyl | 2.221 |
| 64 | F | OCH ₃ | H | OCH ₃ | 5-chloro-2-pyridyl | 2.301 |
| 65 | F | OC ₂ H ₅ | H | OCH ₃ | 5-chloro-2-pyridyl | 2.522 |
| 66 | OCH ₃ | OCH ₃ | H | F | 5-bromo-2-pyridyl | 3 |
| 67 | F | N(Me) ₂ | H | F | 5-bromo-2-pyridyl | 2.301 |
| 68 | F | CN | H | F | 5-bromo-2-pyridyl | 1.698 |
| 69 | Cl | OC ₂ H ₅ | Cl | F | 5-bromo-2-pyridyl | 1.677 |
| 70 | - | - | - | - | 5-cyano-2-pyridyl | 1.154 |
| 71 ^a | - | - | - | - | 5-chloro-2-pyridyl | 1.046 |

a - test set compounds

b - the experimental IC₅₀ values (in micro molar) were converted into -logIC₅₀ (pIC₅₀, in micro molar).

2. Modeling

All 71 PETTs structures (Table 1 and 2) were built on workspace of Win CAChe 6.1 (molecular modeling software, a product of Fujitsu private limited, Japan) and energy minimization of the molecules was done using Allinger's MM2 force field followed by AM1 (Austin model) method available in MOPAC module until the root mean square gradient value becomes smaller than 0.001 kcal/mol Å. Most stable structure for each compound was generated and used for calculating various physico-chemical descriptors like thermodynamic, steric and electronic values of descriptors.

All the calculated descriptors (25 descriptors calculated by Win CAChe 6.1, the complete descriptors data set of all compounds will be provided on request) were considered as independent variable and biological activity as dependent variable. STATISTICA version 6 (StatSoft, Inc., Tulsa, USA) software was used to generate QSAR models by stepwise multiple linear regression analysis. Statistical measures used were n-number of compounds in regression, r-correlation coefficient, r²-squared correlation coefficient, F-test (Fischer's value) for statistical significance, SEE-standard error of estimation, q²-cross validated correlation coefficient and correlation matrix to show correlation among the parameters.

3. Results and discussion

A data set of 71 compounds (Table 1 and 2) for HIV-1 RT inhibitory activity was used for the present QSAR study. The QSAR studies of the PETTs series resulted in several QSAR equations. Intercorrelation between the descriptors is given in table 3. The selected descriptors are molecular weight (MW), critical pressure (CP), valence connectivity index second order (VCI2) and valence connectivity index fourth order (VCI4) (Table 4). The best equation when we considered only one parameter is Eq.1.

$$\log (1/IC_{50}) = -2.172 (\pm 0.537) + 0.010 (\pm 0.002) MW$$

n = 60, r = 0.650, r² = 0.422, r²_{adj} = 0.412, SEE = 0.733, F = 43.08, P < 0.001, q² = 0.384, S_{PRESS} = 0.768

(1)

Table 3. Selected physicochemical parameters of PETT derivatives.

| Compd. No | MW | CP | VCI2 | VCI4 |
|-----------------|----------|----------|----------|----------|
| 1 | 263.3869 | 37.13495 | 4.751703 | 2.184906 |
| 2 | 281.3774 | 34.56141 | 4.857513 | 2.231703 |
| 3 ^a | 281.3774 | 34.56141 | 4.896025 | 2.232702 |
| 4 | 281.3774 | 34.56141 | 4.89257 | 2.191303 |
| 5 ^a | 293.4131 | 31.95541 | 5.077951 | 2.440062 |
| 6 | 293.4132 | 31.95541 | 5.117633 | 2.368198 |
| 7 | 293.4131 | 31.95541 | 5.114179 | 2.332433 |
| 8 | 277.4138 | 32.57855 | 5.192589 | 2.421149 |
| 9 | 309.3924 | 35.60016 | 5.151926 | 2.430906 |
| 10 | 281.4023 | 42.83051 | 5.281243 | 2.57705 |
| 11 | 297.8318 | 34.88881 | 5.263843 | 2.461435 |
| 12 | 307.4399 | 28.7503 | 5.346521 | 2.497676 |
| 13 | 321.4669 | 26.00428 | 5.761971 | 2.550923 |
| 14 ^a | 321.4669 | 26.21776 | 6.080115 | 2.5867 |
| 15 | 355.484 | 27.18334 | 6.328624 | 2.929333 |
| 16 | 323.4395 | 27.78851 | 5.40719 | 2.695289 |
| 17 | 323.4396 | 27.78851 | 5.443882 | 2.607854 |
| 18 | 372.3095 | 33.99944 | 6.136333 | 2.880862 |
| 19 | 311.4041 | 29.89321 | 5.186754 | 2.493617 |
| 20 ^a | 325.4305 | 26.986 | 5.415642 | 2.593106 |
| 21 | 299.3682 | 32.2464 | 4.966317 | 2.29206 |
| 22 | 315.8225 | 32.54139 | 5.372646 | 2.549009 |
| 23 | 264.3749 | 41.94743 | 4.58718 | 2.06997 |
| 24 | 264.3749 | 41.94743 | 4.627378 | 2.086237 |
| 25 | 253.3487 | 43.33955 | 4.345468 | 1.973923 |
| 26 | 277.4138 | 32.57855 | 5.529445 | 2.447315 |
| 27 | 291.4406 | 29.28168 | 5.619534 | 2.650928 |
| 28 | 305.4673 | 26.46107 | 6.015981 | 3.010697 |
| 29 | 305.4677 | 26.68021 | 6.303963 | 2.799503 |
| 30 ^a | 319.4944 | 24.02922 | 6.369534 | 3.055741 |
| 31 | 288.3967 | 31.66833 | 5.191488 | 2.367551 |
| 32 | 345.412 | 25.9249 | 6.057937 | 2.857912 |
| 33 | 335.4503 | 26.59772 | 5.800155 | 2.708712 |
| 34 | 297.8317 | 34.8888 | 5.239922 | 2.518567 |
| 35 | 264.3747 | 41.94743 | 4.597211 | 2.043658 |
| 36 | 257.3587 | 32.06407 | 4.2382 | 1.768119 |
| 37 | 336.2549 | 34.1187 | 5.298939 | 2.015307 |
| 38 | 271.3857 | 28.38386 | 4.744012 | 1.873167 |
| 39 ^a | 324.3777 | 27.78852 | 5.242434 | 2.496379 |
| 40 | 372.2358 | 29.79539 | 5.513552 | 2.12246 |
| 41 | 307.3666 | 25.07517 | 4.958625 | 1.980321 |
| 42 ^a | 319.4023 | 23.45092 | 5.179062 | 2.181877 |
| 43 | 384.2715 | 27.70083 | 5.733989 | 2.324017 |
| 44 | 272.3734 | 31.56167 | 4.579488 | 1.75823 |
| 45 | 337.2426 | 38.34025 | 5.134415 | 1.90037 |
| 46 | 327.4218 | 25.81962 | 5.590281 | 2.79734 |
| 47 | 443.3145 | 23.24783 | 6.900401 | 2.58418 |
| 48 | 443.3145 | 24.33843 | 6.722125 | 2.58284 |
| 49 | 352.7943 | 23.24784 | 5.334068 | 2.141875 |
| 50 | 370.8528 | 21.69381 | 6.088257 | 2.388554 |
| 51 ^a | 415.3042 | 24.00568 | 6.566814 | 2.511132 |

| | | | | |
|-----------------|----------|----------|----------|----------|
| 52 | 402.2621 | 26.03082 | 5.855521 | 2.314264 |
| 53 ^a | 416.2888 | 23.65668 | 6.08441 | 2.423592 |
| 54 | 416.2889 | 23.65668 | 6.245398 | 2.493591 |
| 55 | 432.7433 | 23.84186 | 6.461566 | 2.725254 |
| 56 | 388.2921 | 21.55304 | 5.98301 | 2.602676 |
| 57 | 479.744 | 21.85639 | 6.772664 | 2.804939 |
| 58 | 378.8571 | 19.96548 | 5.729768 | 2.575308 |
| 59 | 351.856 | 23.24783 | 5.51256 | 2.315676 |
| 60 | 424.3611 | 21.43347 | 6.448894 | 2.664909 |
| 61 | 398.2984 | 25.1003 | 5.998397 | 2.400664 |
| 62 | 416.2888 | 23.65668 | 6.084409 | 2.449885 |
| 63 ^a | 384.2715 | 27.70083 | 5.769508 | 2.303487 |
| 64 | 414.2979 | 24.31443 | 6.075958 | 2.515138 |
| 65 | 428.3246 | 22.16619 | 6.304847 | 2.624466 |
| 66 | 414.2976 | 24.31443 | 6.074787 | 2.551072 |
| 67 | 415.304 | 24.00568 | 6.595736 | 2.437967 |
| 68 | 422.3235 | 22.31301 | 6.752365 | 2.642094 |
| 69 | 467.188 | 22.67574 | 6.991711 | 3.100554 |
| 70 | 285.3724 | 25.81962 | 4.691972 | 1.960616 |
| 71 ^a | 294.8073 | 28.17333 | 4.945214 | 1.987984 |

Table 4. Correlation matrix of the selected physicochemical parameters and biological activity.

| Variables | log (1/IC ₅₀) | MW | CP | VCI2 | VCI4 |
|---------------------------|---------------------------|--------|--------|-------|------|
| log (1/IC ₅₀) | 1 | | | | |
| MW | 0.631 | 1 | | | |
| CP | -0.557 | -0.721 | 1 | | |
| VCI2 | 0.334 | 0.770 | -0.715 | 1 | |
| VCI4 | -0.091 | 0.467 | -0.445 | 0.753 | 1 |

The above equation is not statistically significant one because r^2 value is too low and SEE value is too high. So we have considered the best equation containing two parameters is Eq.2.

$$\log (1/IC_{50}) = 0.443 (\pm 0.628) + 0.014 (\pm 0.001) MW - 1.663 (\pm 0.290) VCI4$$

$n = 60, r = 0.795, r^2 = 0.631, r^2_{adj} = 0.619, SEE = 0.590, F = 49.69, P < 0.001, q^2 = 0.592, S_{PRESS} = 0.629$

(2)

The above equation is statistically significant one but the internal predictivity is less. So we have considered the best equation containing three parameters is Eq.3.

$$\log (1/IC_{50}) = 3.127 (\pm 1.206) + 0.011 (\pm 0.002) MW - 1.769 (\pm 0.280) VCI4 - 0.047 (\pm 0.018) CP$$

$n = 60, r = 0.818, r^2 = 0.670, r^2_{adj} = 0.652, SEE = 0.564, F = 38.50, P < 0.001, q^2 = 0.616, S_{PRESS} = 0.593$

(3)

The above equation is statistically significant one. The r^2 and internal predictivity of the model is good. When we have considered the best equation containing four parameters is Eq.4.

$$\log (1/IC_{50}) = 4.100 (\pm 1.245) + 0.017 (\pm 0.003) MW - 1.001 (\pm 0.438) VCI4 - 0.053 (\pm 0.018) CP - 0.831 (\pm 0.373) VCI2$$

$n = 60, r = 0.835, r^2 = 0.697, r^2_{adj} = 0.675, SEE = 0.545, F = 32.13, P < 0.001, q^2 = 0.642, S_{PRESS} = 0.577$

(4)

Out of above four models, model 4 was selected as the best model on the basis of high q^2 values and r^2 value. The values given in the parentheses are 95% confidence intervals of the regression coefficients. Eq. 4 could explain 69.7% and predict 64.2% of the variance of the HIV-1 RT inhibitory activity data. The calculated HIV-1 RT inhibitory activity values by Eq.4 are given in Table 5. This model showed good correlation coefficient (r)

of 0.835 between descriptors [molecular weight, valence connectivity index and critical pressure] and HIV-1 RT inhibitory activity. This model also indicates statistical significance > 99.9% with

Table 5. Observed, predicted and calculated activities for training and test set of compounds. using Eq.4.

| Compd. No. | Obs. Act | Cal. Act | Pred. Act (LOO) | Pred. Act |
|---------------------|-----------------|-----------------|------------------------|------------------|
| Training Set | | | | |
| 1 | 0.0458 | 0.4993 | 0.5228 | - |
| 2 | 1.2218 | 0.816 | 0.8001 | - |
| 4 | 0 | 0.8203 | 0.8501 | - |
| 6 | 0.8239 | 0.8072 | 0.8067 | - |
| 7 | 0.4559 | 0.8427 | 0.8536 | - |
| 8 | 1.0969 | 0.3761 | 0.343 | - |
| 9 | 0.8239 | 0.7976 | 0.796 | - |
| 10 | -0.0414 | -0.3411 | -0.3998 | - |
| 11 | 0.2218 | 0.5039 | 0.5153 | - |
| 12 | 1.2218 | 0.9013 | 0.8893 | - |
| 13 | 0.699 | 0.8673 | 0.8775 | - |
| 15 | -0.0414 | 0.5334 | 0.5849 | - |
| 16 | 1.0458 | 0.9995 | 0.9947 | - |
| 17 | 0.699 | 1.0448 | 1.0645 | - |
| 18 | 1.5229 | 0.673 | 0.5528 | - |
| 19 | 2 | 1.0584 | 1.002 | - |
| 21 | 2.2218 | 1.1035 | 1.0568 | - |
| 22 | 0.699 | 0.7685 | 0.7714 | - |
| 23 | -0.2788 | 0.5069 | 0.5934 | - |
| 24 | 0.1871 | 0.4551 | 0.4843 | - |
| 25 | 1 | 0.5467 | 0.4825 | - |
| 26 | 0.2218 | 0.0423 | 0.0183 | - |
| 27 | 0.4559 | 0.2022 | 0.1808 | - |
| 28 | 0.699 | -0.0873 | -0.2275 | - |
| 29 | -0.3979 | -0.173 | -0.0931 | - |
| 31 | 0.2596 | 0.6679 | 0.6815 | - |
| 32 | 0.699 | 0.7403 | 0.7431 | - |
| 33 | -0.3802 | 0.9013 | 0.9591 | - |
| 34 | -0.2788 | 0.4743 | 0.5121 | - |
| 35 | -0.5911 | 0.5214 | 0.6453 | - |
| 36 | 1.699 | 1.5229 | 1.4984 | - |
| 37 | 1.8239 | 1.5895 | 1.5677 | - |
| 38 | 1.5229 | 1.4104 | 1.3936 | - |
| 40 | 2.2596 | 2.1625 | 2.1561 | - |
| 41 | 2.2218 | 1.9275 | 1.8917 | - |
| 43 | 2.5229 | 2.1032 | 2.0856 | - |
| 44 | 2.3468 | 1.508 | 1.3972 | - |
| 45 | 2.2218 | 1.6296 | 1.5205 | - |
| 46 | 1.301 | 0.9168 | 0.872 | - |
| 47 | 1.8239 | 2.0721 | 2.1167 | - |
| 48 | 3 | 2.1773 | 2.0828 | - |
| 49 | 1.6383 | 2.3309 | 2.3921 | - |
| 50 | 1.3468 | 1.8149 | 1.8616 | - |
| 52 | 2.0969 | 2.4068 | 2.4247 | - |
| 54 | 2.1549 | 2.4733 | 2.4914 | - |
| 55 | 2.3979 | 2.1309 | 2.1132 | - |
| 56 | 2.3979 | 2.031 | 2.0094 | - |
| 57 | 1.9208 | 2.7038 | 2.821 | - |

| | | | | |
|-----------------|--------|--------|--------|--------|
| 58 | 2.2218 | 2.2113 | 2.2099 | - |
| 59 | 2.0969 | 1.993 | 1.9869 | - |
| 60 | 2.301 | 2.1832 | 2.1769 | - |
| 61 | 2.301 | 2.1791 | 2.1741 | - |
| 62 | 1.7447 | 2.4495 | 2.4892 | - |
| 64 | 2.301 | 2.3273 | 2.3287 | - |
| 65 | 2.5229 | 2.3813 | 2.3724 | - |
| 66 | 3 | 2.2958 | 2.2533 | - |
| 67 | 2.301 | 1.9531 | 1.8937 | - |
| 68 | 1.699 | 1.8404 | 1.8573 | - |
| 69 | 1.6778 | 1.9701 | 2.0191 | - |
| 70 | 1.1549 | 1.7651 | 1.8551 | - |
| Test Set | | | | |
| 3 | 0.824 | - | - | 0.753 |
| 5 | 1.397 | - | - | 0.737 |
| 14 | 0.397 | - | - | 0.539 |
| 20 | 2 | - | - | 1.111 |
| 30 | 0.699 | - | - | -0.088 |
| 39 | 0.699 | - | - | 1.292 |
| 42 | 3 | - | - | 1.804 |
| 51 | 2.045 | - | - | 1.924 |
| 54 | 2.222 | - | - | 2.243 |
| 63 | 2.222 | - | - | 2.069 |
| 71 | 1.045 | - | - | 1.524 |

Where Cal.Act - calculated activity, Pred.Act (LOO) – predicted activity by leave one out method, Pred.Act - predicted activity of test set compounds by Eq.4.

F value $F(4,55) = 32.13$. The predictive ability of the selected model was also confirmed by external $r^2_{CV_{ext}}$ method. According to Tropsha et al., the proposed QSAR model is predictive as it satisfies the conditions $r^2_{CV_{ext}} > 0.5$ and $r^2_{pred} > 0.6$ ($r^2_{CV_{ext}} = 0.529$, $r^2_{pred} = 0.627$). The robustness of this model was checked by Y-randomization test. The low r^2 and r^2_{CV} values indicate that the good results in our original model are not due to a chance correlation or structural dependency of the training set.

In this QSAR equation, the positive contribution of MW on the biological activity showed that the increase in molecular weight of molecule leads to better HIV-1 RT inhibitory activity. The negative coefficient of VCI2 and VCI4 showed that increase of branching is detrimental for the activity. The negative coefficient of CP showed that increase of critical pressure of the molecules is not favor for activity. Small groups with high molecular weight like Cl, Br, F and CN are favorable for activity. The substitution of higher branched alkane's and aromatic rings to the phenyl ethyl ring or N-phenyl ring of PET is not favorable for HIV-1 RT inhibitory activity.

4. Conclusions

Therefore, obtained data by adequate designed QSAR studies allow observing aspects and essential structural characteristics to have an increased biological activity, suggesting certain structural requirements for an increased for HIV-1 RT inhibitory potential. Our results open very interesting perspectives regarding PETT derivatives with for HIV-1 RT inhibitory activity.

Acknowledgment

One of the authors V. Ravichandran is thankful to AICTE, New Delhi for providing (QIP) Senior Research Fellowship.

References

- [1] R. C. Gallo, S. Z. Salahuddin, M. Popovic, G. M. Shearer, M. Kaplan, B. F. Haynes, T. J. Palker, R. Redfield, J. Oleske, B. Safai, *Science* **224**, 500 (1984).
- [2] E. D. Clercq, *J. Med. Chem.* **38**, 2491 (1995).
- [3] J. Milton, M. J. Slater, A. J. Bird, D. Spinks, G. Scott, C. E. Price, S. Downing, D. V. S. Green, Madar, R. Bethell, D. K. Stammers, *Bioorg. Med. Chem. Lett.* **8**, 2623 (1998).
- [4] S. D. Young, S. F. Britcher, L. O. Tran, L. S. Payne, W. C. Lumma, T. A. Lyle, J. R. Huff, P. S. Anderson, D. B. Olsen, S. S. Carroll, *Antimicrob. Agents. Chemother.* **39**, 2602 (1995).
- [5] P. Pungpo, S. Hannongbua, *J. Mol. Graphics & Modell.* **18**, 581 (2000).
- [6] A. C. Nair, P. Jayatilleke, X. Wang, S. Miertus, W. J. Welsh, *J. Med. Chem.* **45**, 973 (2002).
- [7] P. R. N. Jayatilleke, A. C. Nair, R. Zauhar, W. J. Welsh, *J. Med. Chem.* **43**, 4446 (2000).
- [8] K. Raghavan, J. K. Buolamwini, M. R. Fesen, Y. Pommier, K. W. Kohn, *J. Med. Chem.* **38**, 890 (1995).
- [9] A. K. Debnath, S. Jiang, N. Strick, K. Lin, P. Haberfield, *J. Med. Chem.* **37**, 1099 (1994).
- [10] V. Ravichandran, R. K. Agrawal, *Bioorg. Med. Chem. Lett.* **17**, 2197 (2007).
- [11] V. Ravichandran, V. K. Mourya, R. K. Agrawal, *Arkivoc* **XIV**, 204 (2007).
- [12] V. Ravichandran, V. K. Mourya, R. K. Agrawal, *Internet Electron. J. Mol. Des.* **6**, 363 (2007).
- [13] V. Ravichandran, B. R. Prashanthakumar, S. Sankar, R. K. Agrawal, *Med. Chem. Res. (Inpress)* (2007).
- [14] V. Ravichandran, P. K. Jain, V. K. Mourya, R. K. Agrawal, *Med. Chem. Res. (Inpress)* (2007).
- [15] K. K. Sahu, V. Ravichandran, V. K. Mourya, R. K. Agrawal, *Acta. Chim. Slov. (Inpress)* (2007).
- [16] K. K. Sahu, V. Ravichandran, V. K. Mourya, R. K. Agrawal, *Med. Chem. Res.* **15**, 418 (2007).
- [17] F. W. Bell, A. S. Cantrell, M. Hogberg, S. R. Jaskunas, N. G. Johansson, C. L. Jordan, M. D. Kinnic, P. Lind, J. M. Morin, R. Noreen, B. Oberg, J. A. Palkowitz, C. A. Parrish, P. Pranc, C. Sahlberg, R. J. Ternansky, R. T. Vasileff, L. Vrang, S. J. West, H. Zhang, X. X. Zhou, *J. Med. Chem.* **38**, 4929 (1995).
- [18] A. S. Cantrell, P. Engelhardt, M. Hogberg, S. R. Jaskunas, N. G. Johansson, C. L. Jordan, M. D. Kinnic, P. Lind, J. M. Morin, R. Noreen, B. Oberg, J. A. Palkowitz, C. A. Parrish, P. Pranc, C. Sahlberg, R. J. Ternansky, R. T. Vasileff, L. Vrang, S. J. West, H. Zhang, *J. Med. Chem.* **39**, 4261 (1996).