INTERNATIONAL JOURNAL OF CURRENT RESEARCH IN CHEMISTRY AND PHARMACEUTICAL SCIENCES
(p-ISSN: 2348-5213: e-ISSN: 2348-5221)
www.ijcrcps.com
Coden: IJCROO(USA) Volume 3, Issue 3 - 2016
Research Article
SOI: http://s-o-i.org/1.15/ijcrcps-2016-3-3-6

Structural Characterization for HIV-1 Non-Nucleoside Reverse Transcriptase inhibition on PETT Derivatives

Megha Chourey[a], Tripti Kamalpuria[a], Arpan Bharadwaj[a], Nameeta Bende[b*] and Abhilash Thakur[c]

[a] Department of Chemistry, Govt. Madhav Science PG College, Ujjain, M.P., India
[b] Department of Chemistry, Govt. Holkar Science College, Indore, M.P., India
[c] Department of Applied Science, National Institute of Technical Teachers Training and Research, Bhopal, India
*Corresponding Author: namitabende@gmail.com

Abstract

This work describes the structure activity relationship studies of PETT (Phenyl Ethyl Thaio-Azail Thaio-Urea) derivatives as HIV-1 reverse transcriptase inhibitors. To analysis the structure activity relationship topological and physicochemical properties along with indicator parameters are tested. The application of various structural parameters indicated that the studied set of 25 compounds is structure specific in nature for mamodeling the HIV-1 reverse transcriptase inhibition activity logIC_{50} (50% of inhibitory concentration of PETT derivatives for RTs). The emphasis has also been made on the structural features leading the modeling activity of more potent PETT derivative for the inhibition of Reverse Transcriptase.

Keywords: Topological indices, physicochemical parameters, NNRTI's, molecular modeling.

Introduction

HIV, the human immunodeficiency virus, is a retrovirus of the Lentiviridac family that causes AIDS, a debilitating and deadly disease of the human immune system.

Since the discovery of HIV as the causative agent of AIDS (acquired immune deficiency syndrome) enormous efforts have been invested worldwide in the search for effective drugs against this devastating disease. This major epidemic is caused by two variants of the human immuno-deficiency virus, HIV-1 and HIV-2. Worldwide, The predominant virus is HIV-1.

NNRTIs are considered as primer target to treat HIV-1. In the present study molecular modeling and CADD approach is used to understand and develop the new biologically active chemical entity belongs to well known family of PETT (Phenyl Ethyl Thaio-Azail Thaio-Urea) derivatives as NNRTIs.[1] Several heterocyclic thioureas have been already reported as a new class of potent NNRTIs such phenethylthiazolyl-thiourea (PETT) derivatives.[2-5] Uckun and co-worker[6] described the synthesis of a series of thiazole thioureas with alkyl, aryl, heteroaryl substituents as newly identified NNRTI of HIV, including mutant strains of HIV, and effective in the treatment of multi-drug resistant HIV infection. Generation of this class of derivatives attracts the people working on Anti-HIV drug molecules to understand the unique features responsible for the special biological activity.[3]

HIV is a virus that gradually attacks immune system cells. As HIV progressively damages these cells, the body becomes more vulnerable to infections, which it will have difficulty in fighting off.[7-9] It is at the point of very advanced HIV infection that a person is said to have

AIDS. It can be years before HIV has damaged the immune system enough for AIDS to develop.[10]

To be effective, a designed drug must discriminate successfully the macromolecular target from alternative structures present in the organism. Not only the affinity for the desired target, but also the selectivity over potential competitors, is of crucial importance.
As the efficacy of existing compounds increases, it becomes more difficult to discover new chemical entities with substantial advantages.

Molecular modeling provides the opportunity to utilize a quicker and more cost-effective method of enabling drug discovery and development from the research bench to the patients, bedside. Molecular modeling is accessible, “hands on”, and does not need to be intimidating. It provides some options to the chemistry researchers that are difficult or impossible using traditional experimental techniques. It is relatively young discipline which has not yet come full circle. It creates models based on known data and then uses these models to predict unknown elements of expectation.[11]

The molecular modeling techniques are derived and extended with the various classical mechanical, quantum-chemical and other mathematical theorems generated for real chemical information.

Molecular models are allowing us to “think like a molecule.” We can imagine what is molecule and how a molecule may act and these models provide a window on the molecular world.

Molecular modeling is easy to perform with currently available software, but the difficulty lies in getting the right model and proper interpretation.

In the present study efforts have been made to develop the correct structural information for modeling the new and more potent PETT derivatives to treat the HIV-1.

Materials and Methodology

Quantitative structure-activity relationships (QSAR) have been established for set of 25 analogues of PETT (Phenyl Ethyl Thaio-Azail Thaio-Urea) a potent inhibitor of the HIV-1 reverse transcriptase (RT). The activity of these compounds was Adopted from the literature[1]

Three separate descriptors were used namely 2D-topological descriptors, physicochemical properties and hydrophobic parameter logP(Octanol/Water partition coefficient).

2D-topological descriptors such as Wiener index(W)[12], Randic connectivity index(χ)[13], Balaban J index(J)[14], Szeged index(Sz)[15], Shultz molecular topological index(MTI)[16] and Electrotopological index(S)[17] were tested in –mono, -di, -tri and –tetra variate combinations. Similarly in case of physicochemical properties Molar refractivity (MR), Molar volume (MV), parachor (Pc), index of refraction (η), surface tension (ST), density (D) and polarizability (Pol) were tested in various combinations. Since logP is an important property effecting biological activities, therefore, it is tested separately from other physicochemical properties.All the physicochemical properties are calculated using ACD chemsketch software[18].

All the regressions are carried out using maximum r^2 method[19]. Step-wise regression has been performed for obtaining the best model.

Results and Discussion

As mentioned in the introduction PETT derivatives are one of the very important class of NNRTIs and having significant role in the treatment of HIV, thus analysis of the relationship between biological activity and the structure of the molecule and various molecular properties of PETT derivatives become very important and essential to develop the new compounds with optimized activity and better potential. In the present study efforts have been made to do the same. In the present study Structural analysis has been done for the set of 25 compounds mentioned in Table1

.

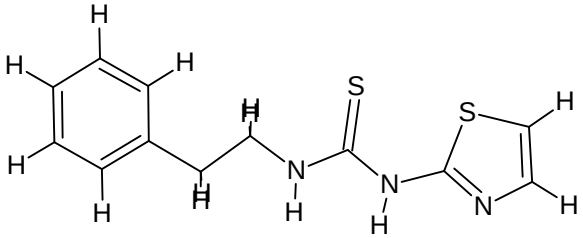

Figure 1. Parent Structure of PETT Derivatives

Table 1: Various Substituents with their biological activity

Comp No.	R'	pIC_{50}
1	Phenyl	0.0457
2	2-fluorophenyl	1.222
3	3-fluorophenyl	0.824
4	4-fluorophenyl	0
5	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	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	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

To understand the relationship between biological responses and the specific chemical system or compounds it is become essential to analyse the relationship on the basis of structure of the compound or on the basis of structural properties particularly physicochemical properties those may play the role either directly or indirectly in availability of the compounds on the biological response site through circulatory system or may play the role directly in interaction between compounds and biological systems. In the context of significance of physicochemical parameters it is worthy to analyse their role in terms of biological response for this particular series of compounds. Thus the parameters tested to analyse the biological response and its dependence over the various structural features directly or indirectly are representing the various structural features viz size, shape, branching, connectivity and electronic or refractive nature of the molecule along with the specific features such as inter and intra molecular forces, volume, steric properties and substitution on the specific position of the molecule and its specific type.

All these parameters tested in the present study can also be classify, as topological indices or graph theoretical descriptors, classical physicochemical properties, nonconventional physicochemical parameters and indicator parameters.

Topological parameters used in present study represent the size, connectivity, branching, electro-topological stat etc. Classical physicochemical properties are the properties having history of test in QSAR studies but non conventional physicochemical properties are the properties those are used rarely in the history of QSAR. Indicator parameters are used in present study to explore the type of substitution, effect of substitution and the positions of the substitution. All these parameters are recorded in Table 2, Table 3, Table 4 and Table 5 respectively.

All these parameters tested in the present study are correlated with the biological activity $\log IC_{50}$ and mutually. The univariate correlation between the biological activity $\log IC_{50}$ and parameters are presented in form of correlation matrix, also correlation matrix display the mutual correlation amongst the parameters and are presented in Table 6.

Table2: Physicochemical parameters used in present investigation

Comp No.	MR	MV	Pc	RI	ST	D	Pol
1	76.97	199.1	569.6	1.7	67.0	1.322	30.51
2	76.96	203.3	576.8	1.681	64.7	1.383	30.51
3	76.96	203.3	576.8	1.681	64.7	1.383	30.51
4	76.96	203.3	576.8	1.681	64.7	1.383	30.51
5	83.65	223.1	626.3	1.673	62.1	1.315	33.16
6	83.65	223.1	626.3	1.673	62.1	1.315	33.16
7	83.65	223.1	626.3	1.673	62.1	1.315	33.16
8	81.8	215.3	607.3	1.684	63.2	1.287	32.42
9	83.52	210.9	625.1	1.722	77.1	1.461	33.11
10	78.85	197.5	584.7	1.73	76.7	1.414	31.26
11	81.87	211.0	605.5	1.703	67.7	1.411	32.45
12	88.28	239.6	666.1	1.658	59.7	1.282	34.99
13	92.91	256.1	705.9	1.645	57.7	1.255	36.83
14	92.87	256.5	703.3	1.644	56.6	1.253	36.81
15	103.41	269.6	760.8	1.692	63.3	1.318	40.99
16	90.33	247.1	683.0	1.651	58.3	1.308	35.81
17	90.33	247.1	683.0	1.651	58.3	1.308	35.81
18	91.34	239.2	676.8	1.688	64.0	1.555	36.21
19	83.64	227.3	633.4	1.657	60.3	1.369	33.16
20	88.28	243.8	673.2	1.643	58.1	1.334	34.99
21	76.96	207.5	683.9	1.663	62.6	1.442	30.51
22	81.86	215.2	612.6	1.685	65.6	1.467	32.45
23	75.06	192.3	563.8	1.708	73.8	1.374	29.75
24	75.06	192.3	563.8	1.708	73.8	1.374	29.75
25	69.27	181.9	523.6	1.686	68.6	1.392	27.46

Table:3 Topological parameters used in present investigation

COMP NO.	MTI	W	J	S
1	2739	635	1.509	68.234
2	3110	728	1.558	80.54
3	3156	739	1.534	76.376
4	3202	750	1.513	73.758
5	3553	839	1.587	88.336
6	3645	861	1.546	87.119
7	3737	883	1.509	86.678
8	3110	728	1.588	76.925
9	4004	952	1.629	117.744
10	3110	728	1.558	97.848
11	3110	728	1.558	83.726
12	4212	1002	1.548	97.923
13	4861	1163	1.543	108.956
14	4787	1145	1.564	123.313
15	7224	1660	1.283	109.863
16	4439	1063	1.69	113.523
17	4547	1089	1.649	111.217
18	4010	954	1.626	100.277
19	3956	941	1.65	111.863
20	4499	1077	1.668	124.541
21	3495	825	1.614	142.069
22	3495	825	1.614	109.42
23	2739	635	1.509	70.225
24	2739	635	1.509	69.228
25	2320	534	1.506	64.634

Table:4 Connectivity indices used in present investigation

Comp No	X0	X1	X2	X3	X4	X5
1	11.924	8.343	6.926	5.244	4.164	2.616
2	12.795	8.754	7.444	5.771	4.483	2.812
3	12.795	8.737	7.56	5.565	4.51	2.782
4	12.795	8.737	7.548	5.655	4.306	2.9
5	13.502	9.292	7.635	6.073	4.863	3.081
6	13.502	9.275	7.729	5.989	4.723	3.059
7	13.502	9.275	7.717	6.063	4.616	2.998
8	12.795	8.754	7.444	5.771	4.483	2.812
9	14.372	9.665	8.375	6.306	5.135	3.272
10	12.795	8.754	7.444	5.771	4.483	2.812
11	12.795	8.754	7.444	5.771	4.483	2.812
12	14.209	9.775	8.11	6.108	5.023	3.209
13	14.916	10.275	8.463	6.377	5.107	3.421
14	15.097	10.131	8.95	6.216	5.233	3.326
15	16.615	11.793	10.061	7.803	6.461	4.517
16	15.097	10.241	8.354	6.854	5.575	3.647
17	15.097	10.224	8.438	6.832	5.347	3.538
18	14.372	9.686	8.269	6.409	5.114	3.36
19	14.372	9.703	8.163	6.54	5.257	3.33
20	15.079	10.203	8.543	6.675	5.479	3.543
21	13.665	9.165	7.972	6.226	4.955	2.933
22	13.665	9.165	7.972	6.226	4.955	2.933
23	11.924	8.343	6.926	5.244	4.164	2.616
24	11.924	8.343	6.926	5.244	4.164	2.616
25	11.217	7.843	6.573	4.994	3.987	2.003

Table: 5 Indicator parameters used in present investigation

Comp No	IOMe	IX
1	0	0
2	0	1
3	0	1
4	0	1
5	1	0
6	1	0
7	1	0
8	0	0
9	0	0
10	0	0
11	0	1
12	0	0
13	0	0
14	0	0
15	0	0
16	1	0
17	1	0
18	1	1
19	1	1
20	0	1
21	0	1
22	0	1
23	0	0
24	0	0
25	0	0

Table 6. Correlation of various parameters with biological activity in from of correlation matrix.

	BA	MR	MV	PC	IR	ST	D	POL
BA	1.00000							
MR	.05740	1.00000						
MV	.18369	.97071	1.00000					
PC	.05120	.94113	.92553	1.00000				
IR	-.53836	-.44195	-.64316	-.45061	1.00000			
ST	-.45995	-.56346	-.72914	-.51790	.94892	1.00000		
D	.14556	-.33345	-.43483	-.36295	.53320	.54451	1.00000	
POL	.05767	1.00000	.97072	.94104	-.44202	-.56351	-.33307	1.00000
LOGP	-.26525	.04326	.06623	-.02027	-.11837	-.28124	-.32620	.04279
SMTI	.04763	.94101	.90958	.89068	-.41077	-.49991	-.32375	.94108
W	.07411	.94949	.92623	.90082	-.44356	-.52496	-.33017	.94957
J	.61252	-.08092	.03048	-.06651	-.34856	-.21919	.19525	-.08063
TIE	.49331	.58739	.63078	.55869	-.47847	-.41653	.02655	.58766
X0	.25351	.93510	.93991	.88320	-.53688	-.58872	-.24575	.93528
X1	.16700	.95487	.94388	.90730	-.49092	-.55982	-.30549	.95500
X2	.15828	.92233	.89966	.85940	-.43299	-.50781	-.19963	.92247
X3	.26485	.88709	.86767	.81813	-.43639	-.51680	-.14658	.88735
X4	.27565	.89838	.88300	.84337	-.45225	-.51447	-.17660	.89857
X5	.12164	.93571	.90058	.88692	-.39527	-.48971	-.24356	.93586
TA1	.39881	.24461	.26330	.23326	-.21751	-.23650	.26076	.24507
TA2	.25175	.09738	.09843	.25900	-.05996	-.03497	.24655	.09761
IOME	.30400	.27701	.31964	.25042	-.33215	-.38161	-.05131	.27752
IX	.42747	-.17580	-.13255	-.22333	-.10825	-.11977	.56718	-.17553

	LOGP	SMTI	W	J	TIE	X0	X1	X2
LOGP	1.00000							
SMTI	-.03823	1.00000						
W	-.05322	.99843	1.00000					
J	-.39619	-.23869	-.19095	1.00000				
TIE	-.32299	.62378	.64928	.40795	1.00000			
X0	-.16730	.94085	.95606	.09559	.78829	1.00000		
X1	-.11587	.98002	.98802	-.04903	.70880	.98572	1.00000	
X2	-.10515	.96643	.97220	-.06959	.75273	.97381	.97283	1.00000
X3	-.19252	.92177	.92828	.05049	.74775	.96141	.95972	.93913
X4	-.19939	.94231	.94869	.03035	.76601	.97124	.97317	.96080
X5	-.10041	.95723	.96127	-.06267	.65179	.95956	.97860	.94981
TA1	-.22081	.13577	.15288	.43526	.22207	.27707	.23018	.18432
TA2	-.24794	.00901	.01921	.31303	.05286	.10064	.07887	.02288
IOME	-.25436	.14753	.16800	.36225	.12430	.28308	.25254	.14041
IX	.10944	-.14747	-.14246	.28907	.16510	-.02392	-.10458	.00000

	X3	X4	X5	TA1	TA2	IOME	IX
X3	1.00000						
X4	.98033	1.00000					
X5	.96182	.96043	1.00000				
TA1	.31489	.30889	.27878	1.00000			
TA2	.13418	.14071	.14490	.84644	1.00000		
IOME	.33581	.25912	.28011	.59201	.46485	1.00000	
IX	.02984	-.00656	-.05290	.23196	.12480	-.09651	1.00000

From the perusal of Table 6, except Balaban branching Index and Refractive Index, none of the topological or physicochemical parameters are near by the statistically significant correlation with the biological activity. These two parameters are having the significant statistical value but they are marginal to the scale of statistical significance not self explanatory to the activity of the studied compounds.

Table 6 explores the dominance of Balaban branching index in characterizing the activity. The comparison amongst the refractive index and branching, Branching dominates the biological characterization with higher magnitude. It is also explore from the correlation that the increase in branching is not favoring the biological activity. But the increase in refractive index or properties may favor the biological activity.

As these two parameters are not self explanatory we fallow the bi-parametric combinations. Out of the bi-parametric combinations best model obtained from the combination of IR and density D. the model obtained from the bi parametric combination is as below

$\log IC_{50}$ = -24.2203(±4.7505)RI + 5.7212 (±1.5978) D – 33.6949 (Eq. 1)
n = 25, Se = 0.4764, R = 0.7425, F = 13.516, Q = 1.56

Equation 1 exhibits the domination of parameters for refractive properties and the density. While branching dose not appears to be very significant within the combination. It shows that branching properties or its effect will be dominated by other physicochemical parameters.

As we analysis the statistical parameters it is observed that with the good correlation coefficient eq. 1 having poor quality factor Q. thus we tried tri-parametric combination. Out of these combinations fallowing combinations are found significant and presented in eq.2 and 3 respectively.

$\log IC_{50}$ = -23.8028(±4.6944)RI + 5.0563 (±1.6585) D – 0.1458(±0.1139)logP +34.0687 (Eq. 2)
n = 25, Se = 0.4697, R = 0.7641, F = 9.819, Q = 1.63

$\log IC_{50}$ = -19.0856(±5.4277)RI + 4.2701 (±1.7446) D + 2.4439(±1.4122)J+23.2455 (Eq. 3)
n = 25, Se = 0.4562, R = 0.7793, F = 10.826, Q = 1.71

As we have pass through the equations it is observed that increase in the parameters or addition of partition coefficient logP in bi-parametric combination reduces the magnitude of refractive properties in modeling the activity $\log IC_{50.}$ Information generated by the statistics revels that the refractive properties may also having the non linear relationship with the biological activity in combinations of other parameters.

Addition of the branching index in bi parametric combination increases the value of regression but the magnitude of refractive properties and density both decreases. Branching also having the poor magnitude. Model demonstrates that increase in the branching in molecule may not favor the biological activity of the compounds. Features though may increase the refractive properties favoring the activity of the compounds. These two properties leading the activity with negative effect of each other revels that the branching in the molecule may dominate the characterization but higher branching may reduce the biological activity. It may also stat that the branching responsible for increase in the refractive index also not favoring the biological activity. There is continues decrease in the magnitude of refractive parameters from eq1 to 3 it inform its linear relationship with the biological activity.

Comparison of both the eq 2 and 3 explore that the logP is dominated by the branching in modeling the anti HIV activity for PETT derivatives. It is also informed by the comparison that the logP may not have the direct role in biological activity. It may play the partial role while drug molecule passes through the various barriers. Higher magnitude of branching shows the dominance over the logP.

Analysis of magnitude of residue from eq. 3 indicate to fallow the out liar concept. Analysis of residue shows the exceptional behavior of compound no 4.

After the deletion of compound no 4 model obtained is as fallows.

$\log IC_{50}$ = -20.2326(±5.1988)RI + 4.7189 (±1.6773) D + 1.9865(±1.3665)J+25.2927 (Eq. 4)
n = 24, Se = 0.4337, R = 0.7799, F = 11.846, Q = 1.80

After the deletion of compound no. 4 increase in the value of r is very less but significant lowering in the value of Se justify the deletion. Eq. also explore that the magnitude of refractive parameters increase with the deletion of compound. It reflects the non linear behavior of the refractive properties in terms of biological activity. There is also change in the magnitude of branching with deletion of compounds revels the case specific relationship between the branching and biological activity under study.

Deletion of compound did not change the direction of behavior for density D. This informs about the linear relationship or role between biological activity and the physicochemical parameter D.

Equation 4 also having the compound with exceptional behavior from the parent series. After the deletion of compound no. 14 model obtained is as below.

$\log IC_{50}$ = -21.6175(±5.0291)RI + 4.3158 (±1.6186) D + 1.9105(±1.3055)J+28.3161 (Eq. 5)
n = 23, Se = 0.4141, R = 0.8257, F = 13.572, Q = 1.99

Statistics generated after the deletion of the compound justify the deletion of compound no. 14. Remarkable improvement in the value of r, F and Q confirm the exceptional behavior of the compound from the parent series.

Equation 4 also having one out liar. After the deletion of compound no 17 model obtained is as below

$logIC_{50}$ = -22.3945(±4.6740)RI + 3.8996 (±1.5132) D + 2.3742(±1.2305)J+29.5012 (Eq. 6)
n = 22, Se = 0.3836, R = 0.8606, F = 17.136, Q = 2.24

Deletion of compound no 17 improves the statistics significantly. The direction of change in magnitude for parameters confirm the findings of eq. 4.

This equation also having the out liars. Model obtained after out liar the compound no 22 is as below

$logIC_{50}$ = -23.3869(±4.4079)RI + 4.7310 (±1.4843) D + 2.3611(±1.1521)J+30.0875 (Eq. 7)
n = 21, Se = 0.3591, R = 0.8858, F = 20.645, Q = 2.47

Equation 7 also extending the study in same direction with increase in value of correlation coefficient and F ratio. But model also having the compound no 16 as out liar.
After the deletion of compound no 16 model obtained is as below.

$logIC_{50}$ = -23.5964(±4.0094)RI + 4.0830 (±1.3833) D – 3.2121 (±1.1208) J+30.0346 (Eq. 8)
n = 20, Se = 0.3266, R = 0.9121, F = 26.400, Q = 2.79

As we delete the compound no 16 from the calculation it is obvious that there will be increase in the value of r with lower number of compounds but at the very same time increase in the value of F ratio justify the deletion of compounds.
Equation 8 also having the one out liar and after the deletion of compound no 13 model obtained is as below.

$logIC_{50}$ = -25.6360 (±3.6258)RI + 3.5041 (±1.2397) D + 3.1695 (±0.9853) J+34.3692 (Eq. 9)
n = 19, Se = 0.2870, R = 0.9370, F = 35.948, Q = 3.26

Continues increase in the regression value and quality factor Q justify the all deletions. eq also explore the increase in value of refractive property may favor the activity and the same with branching may reduce the activity. Biological activity calculated by the eq. 9 presented in table 7 and graphically presented in figure 3.

As we pass from eq. 3 to eq. 9 it is observed that, with the deletion of compounds from the calculations there are increase in the magnitude of refractive properties. It is also observed that the branching may dominates the modeling with increase in the magnitude but physicochemical property density shows the variable nature with change in the direction of magnitude. As density not leading in characterization of activity its change can be considered as induced force to support the activity.

The present study also explore that the higher branching in the molecule not favoring the activity. The substitution which may increase the branching in the molecule may reduce the anti HIV activity of the PETT derivatives. It may be due to non permeability of higher branching molecule through the barriers or induced repulsion from the receptor site. Refractive properties are not playing the direct role in directing the biological activity but it reflects the nature of substitution on the molecule and may play the role in permeability of the moiety through the barriers. Non linier behavior of the branching and refractive properties demonstrates that the branching due to substitution having higher refractive properties may help the derivatives to show the potential anti HIV activity.

All these models are validated and critically discussed using variety of cross validation parameters recorded in Table 8

.

Table 8. Cross validation parameters calculated for the eq. 1 to 9.

S. No.	Parameters	N	PRESS	SSY	R^2_{cv}	PSE	SPRESS	Press/SSY
1.	2	25	4.99390	6.13640	0.18618	0.44694	0.47644	0.81382
2.	3	25	4.63240	6.49790	0.28709	0.43046	0.46967	0.71291
3.	3	25	4.37060	6.75970	0.35343	0.41812	0.45621	0.64657
4.	3	24	3.76220	6.68530	0.43724	0.39593	0.43372	0.56276
5.	3	23	3.25820	6.98240	0.53337	0.37638	0.41411	0.46663
6.	3	22	2.64850	7.56410	0.64986	0.34697	0.38359	0.35014
7.	3	21	2.19280	7.98880	0.72552	0.32314	0.35915	0.27448
8.	3	20	1.70620	8.44580	0.79798	0.29208	0.32655	0.20202
9.	3	19	1.23580	8.88510	0.86091	0.25503	0.28703	0.13909

*****N = number of data set; PRESS (predicted residual sum squares), SSY (sum of the squares of response value), r^2_{cv}(overall predictive ability), S_{PRESS} (uncertainty of prediction), PSE (predictive square of error).

PRESS (predicted residual sum of squares) is an important cross-validation parameter as it is a good estimate of the real predictive error of the models. Its value less than SSY (sum of the squares of response value) indicate that the model predicts better than chance and can be considered statistically significant. In the present case all the models represented by eq. 1 to 9 have PRESS < SSY indicating them to be better than chance and statistically significant. In case of model 8 and 9 the value of PRESS is very-very less than the SSY thus may be considered as good predictive models.

Furthermore, the ratio PRESS/SSY is used to estimate the confidence interval of the psychotomimetic activity. To have a reliable QSAR model, PRESS/SSY should be smaller than 0.4. In our case the ratio PRESS/SSY ranges between .85 – 0.139 indicating that some of the proposed models are unreasonable QSAR models and required other parameters to elucidate the structural requirements. Models represented in form of eq.6 to 9 are good, reliable QSAR models and can be used to predict the biological activity of new compound. As shown in table 9 model 9 having the highest reliability and having the good predictability of the activity.

The indication of the performance of the model is obtained from r^2_{cv} (the overall predictive ability). In present case highest r^2_{cv} is found for the model expressed by equation (9), indicating that it has good predictive power. Models presented from eq.5 to 8 also showing the positive coefficients with the significant statistical values of r^2_{cv}. As we pass from model 1 to 9 the value of r^2_{cv} is increasing this increase in the value of r^2_{cv} verify the deletion of compounds from the calculation and also justify the addition of third parameter in bi parametric combination.

Another useful cross-validated parameter is $S_{PRESS,}$ which is used in deciding uncertainty of prediction. However, this parameter in the present case is of no value as it is almost equivalent to the standard error of estimation (Se). Under such situation the parameter PSE is used. The lowest value of PSE, the better is the predictive power which indicates that the model has excellent correlation ability. Based on PSE values (Table 9) once again we observed that the model 9 [equation (9)] has the reasonable correlation ability and predictability.

For the cross-validation analysis we have used the computer program “ANALYSIS” developed by Thakur and Bhadoria. The result obtained from the cross validation are given in table 9.

Table 9. Modeling parameters calculated for the compounds of minimum residue.

Comp. No.	TE	DpM	RMSg
1.	18.14866	1.641	0.09602
10.	17.85118	1.466	0.09814
12.	23.36665	1.677	0.09674
19.	21.91771	3.645	0.0897
24.	23.92687	1.55	0.09464

* TE = Total Energy; DpM = Dipole Moment; RMSg = Root Mean Square Velocity Gradient

As interpretation made from the cross-validation parameters proposed model is very good and just border line case to be an excellent predictive model.

At this stage it is worthy to mention that we can out liar only 2 or 3 compounds but for behavioral analysis we made around 6 out liars from the parent series of the compounds and we made their analysis separately.

Analysis of out liar compounds, explore the role of connectivity in their exceptional behavior. The models obtained from six compounds is as below

$$\log IC_{50} = 1.3193\ (\pm 0.2055)\ \chi 5 - 3.8748 \qquad \text{(Eq.10)}$$
$$n = 6,\ Se = 0.1181,\ R = 0.9655,\ F = 41.206,\ Q = 8.17$$

Regression value of eq. 10 explores the domination of higher order of connectivity in modeling of biological activity. It also shows that the higher order of connectivity reduces the anti HIV activity of out liar compounds. Overall analysis express that the connectivity which may reduce the branching favor the biological activity. These out liar compounds having the molecular structure with higher connectivity that may increase the branching in the molecule thus showing the exceptional behavior. We analysis the molecular modeling properties for the compounds having the minimum residue to understand the real structural requirements to develop the new PETT derivative with optimized anti HIV activity. For the purpose we fallow the energy minimization through molecular mechanics using MM+ force field and obtained the energy parameters along with dipole moment and flexibility features. Optimized structures of the molecules are presented in figure 4 to 8 and obtained parameters are presented in table 9 and 10

Table 10. Modeling parameters calculated for the compounds of minimum residue.

Comp. No.	ASA	SAG	HE
1.	370.42	457.67	-9.67
10.	372.65	467.14	-14.12
12.	462.55	534.98	-10.22
19.	394.12	479.15	-8.8
24.	360.67	452.67	-10.91

* ASA = Approximate Surface Area; SAG = Surface Area Grid; HE = Hydration Energy

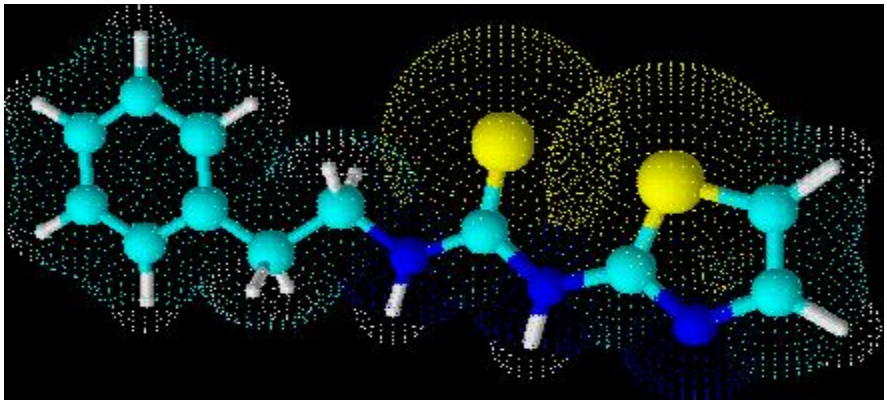

Figure 4. Optimized structure for compound no 1.

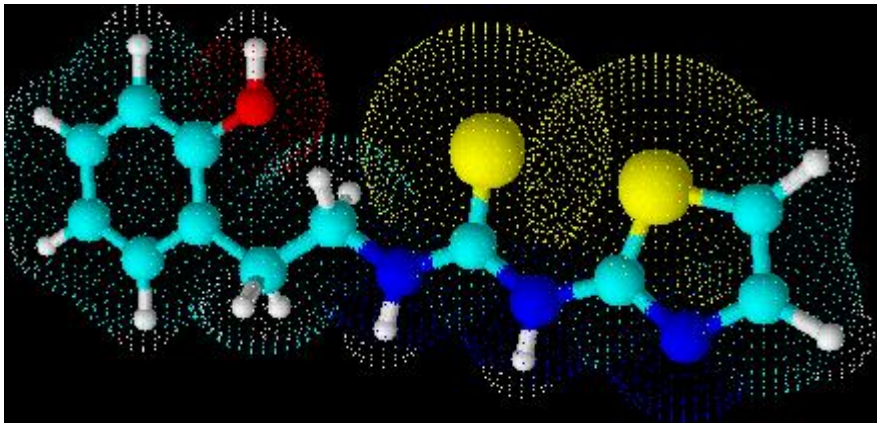

Figure 5. Optimized structure for compound no 10.

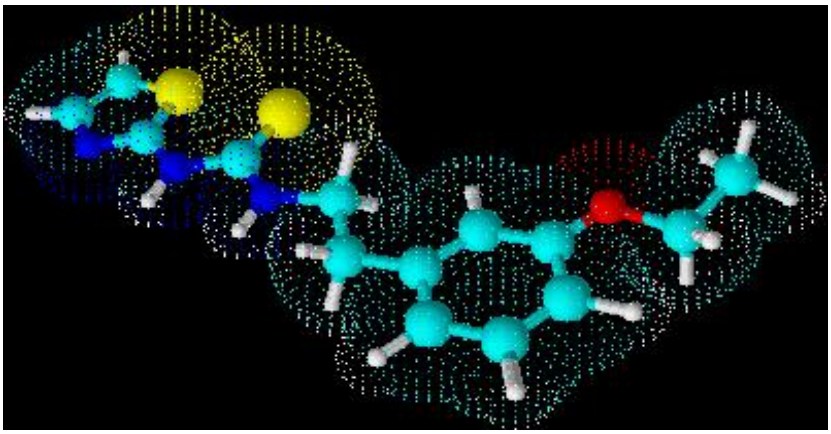

Figure 6. Optimized structure for compound no 12.

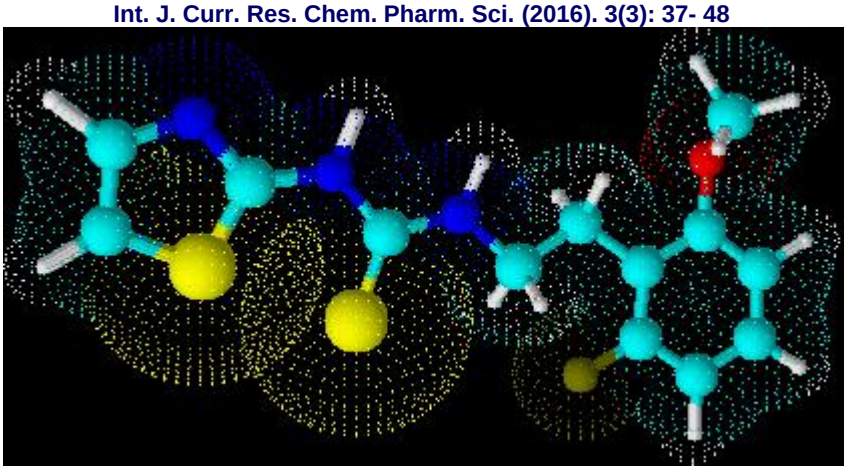

Figure 7. Optimized structure for compound no 19.

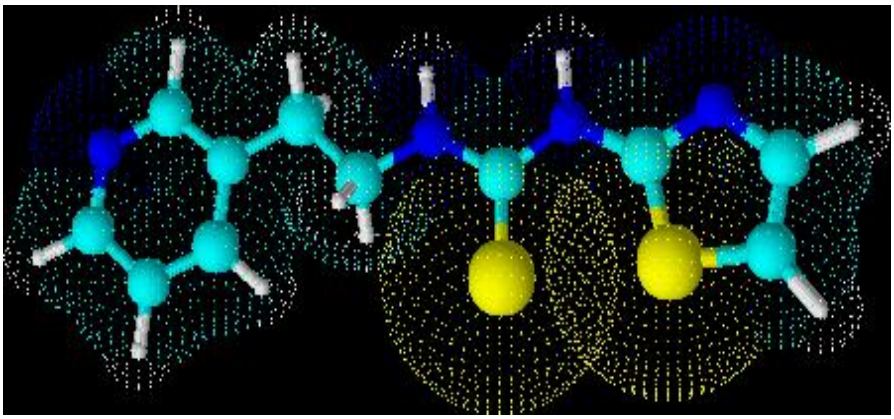

Figure 8. Optimized structure for compound no 24.

Relation between the biological activity and modeling parameters are presented in fallowing equations.

$\log IC_{50}$ = 0.8262 (±0.2894) DpM – 0.9663

(Eq.11)

n = 5, Se = 0.5357, R = 0.8550, F = 8.152, Q = 8.17

$\log IC_{50}$ = -208.574 (±0.2894) RMSg + 20.5072

(Eq.12)

n = 5, Se = 0.6750, R = -0.7569, F = 4.023, Q = 8.17

Regression analysis explores the role of dipole moment in modeling of anti HIV activity of PETT derivatives. Eq. 11 exhibits that the substitution or connectivity or branching by that the value of dipole moment decreases will favor the biological activity i.e., increase in the dipole moment may not allow the molecule to cross the membrane or it may reduce the interaction between the receptor and drug molecule. Eq.12 demonstrates that the flexibility in the molecule may favor the biological activity of the PETT derivatives.

Conclusion

On the basis of result and discussion made above conclusion can be drawn.

The NNRT's inhibition activity is highly structure dependent for the both series of PETT derivatives under study.

Physicochemical property density lead the anti HIV activity with positive effect in studied set of PETT derivatives.

Permeability parameters or phenomenon may govern the biological behavior of compounds while transportation of molecule in biological system.

In studied set of compounds, sulpher is present as the part of five member ring.

Thus order of connectivity changed his behavior and becoming the secondary feature in modeling the biological activity for the compounds.

In case of six out lair compounds connectivity lead the activity because of its partial involvement in reducing the branching.

The new derivative that may design must have the substitution that can increase the Dipole moment in the derivative.

Derivatives must have the less branching but higher order of connectivity for potential anti HIV activity.

Flexibility in molecule may favor the receptor ligand interaction thus this parameter should be in consideration while designing the new derivatives.

References

1. V. Ravichandrana, V. K. Mouryab, R. K. Agrawala; Qsar Modeling Of Hiv-1 Reverse Transcriptase Inhibitory Activity With Pett Derivatives, Digest Journal of Nanomaterials and Biostructures, 2008, Vol. 3, 1, 9 – 17.
2. Walid Fathalla, Syntheses and reactions of methyl [3-(4-phenyl-thiazol-2-yl)-thioureido] alkanoates and related compounds ARKIVOC 2008 (xii) 245-255
3. Heinisch, G.; Matuszczak, B.; Pachler, S.; Rakowitz, D. *Antivir. Chem. Chemother.* **1997**, *8*, 443.
4. Ren, J.; Diprose, J.; Warren, J.; Esnouf, R. M.; Bird, L. E.; Ikemizu, S.; Slater, M.; Milton, J.; Balzarini, J.; Stuart, D. I; Stammers, D. K. *J. Biol. Chem.* **2000,** *275*, 5633.
5. Ahgren, C.; Backro, K.; Bell, F. W.; Cantrell, S.; Clemens, M.; Colacino, J. M.; Deeter, M. J. B.; Engelhardt, J. A.; Hogberg, M.; Jaskunas, S. R. *Antimicrob. Agents Chemother.* **1995**, *39,* 1329.
6. Uckun, F. M.; Venkatachalam, T. K. *United States Patent* **2005**, 6960606.
7. Rajni Garg, Satya P. Gupta, Hua Gao,§ Mekapati Suresh Babu, Asim Kumar Debnath, and Corwin Hansch, Comparative Quantitative Structure-Activity Relationship Studies on Anti-HIV, Drugs; *Chem. Rev.* 1999, *99,* 3525-3601
8. M. Marastoni, A. Baldisserotto, C. Trapella, J. McDonald, F. Bortolotti, R. Tomatis, HIV protease inhibitors: synthesis and activity of *N*-aryl-*N'*-hydroxyalkyl hydrazide pseudopeptides; European Journal of Medicinal Chemistry (2005) ,40, 445–451
9. Yeni, P. G.; Hammer, S. M.; Hirsch, M. S.; Saag, M. S.; Schechter, Fauci, A. S. Host factors and the pathogenesis of HIV-induced disease. *Nature* 1996, *384*, 529-534.
10. Berger, E. A.; Murphy, P. M.; Farber, J. M. Chemokine Receptors as HIV-1 Coreceptors: Roles in Viral Entry, Tropism, and Disease. *Annu. ReV. Immunol.* 1999, *17*, 657-700.
11. Abhilash Thakur, Mamta Thakur, Arpan Bharadwaj, Suprajnya Thakur, SAR and QSAR Study: Modelling of New DAPY Derivatives, Eur. J. Med. Chem., 2008, 43, 471-477.
12. Wiener H, J. Am. Chem. Soc., 69, 2636, 1947.
13. Randic M, J. Am. Chem. Soc., 97, 6609, 1975.
14. Balaban, A. T. From Chemical Graphs to 3D Molecular Modeling. In From Chemical Topology to Three-Dimensional Geometry; Balaban, A. T., Ed.; Plenum: New York, 1997; pp 1-24.
15. Khadikar P.V, Deshpsnde N.V, Kale P.P, Dobrynin A, Gutman I and Domotor G ; J, Chem. Inf. Comp. Sci., 35, 547, 1995.
16. Shultz HP, J.Am.Chem.Inf.CompSci.,29,227,1989.
17. Kier L.B.and Hall L.H.,J. Am.Chem. Inf.CompSci.,37,548,1997.
18. ACD-Lab software for calculating the referred physicochemical parameters; Chemsketh S.O, *WWW. Acdlabs.com.*
19. Chaterjee,S.; Hadi,A.S.; Price,B., *Regression Analysis by Examples*, 3rd ed. Wiley VCH: New york, 2000.

Access this Article in Online	
Quick Response Code	Website: www.icrcps.com
	Subject: Chemistry of Viruses

How to cite this article:

Megha Chourey, Tripti Kamalpuria, Arpan Bharadwaj, Nameeta Bende and Abhilash Thakur. (2016). Structural Characterization for HIV-1 Non-Nucleoside Reverse Transcriptase inhibition on PETT Derivatives. Int. J. Curr. Res. Chem. Pharm. Sci. 3(3): 37-48.