International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

QSAR Modeling for Predicting Anti HIV Activity of PETT Derivatives

Megha Choure[1], Nameeta Bende[2], Arpan Bharadwaj[3], Suresh Kumar Verma[4]

[1,3]Department of Chemistry, Madhav Science PG College, Ujjain

[2]Department of Chemistry, Government Holkar Science College, Indore

[4]Department of Chemistry, Government S. K. P.G. College, Sikar

Abstract: ***The research includes the structural requirements of PETT (Phenyl Ethyl Thaio-Azail Thaio-Urea) derivatives as anti HIV molecules. For analysis of the structural features in respect of biological activity (NNRTI's) topological and physicochemical properties along with indicator parameters are investigated. The application of a multiple linear regression analysis indicated that a combination of parameters representing the branching, connectivity and specific substitution in the studied compounds yielded a statistically significant model for the prediction of activity, logIC_{50} (50% of inhibitory concentration of PETT derivatives for RTs). The emphasis has been made on the structural features required to model new and more potent PETT derivative for the inhibition of reverse transcriptase.***

Keywords: Molecular modeling, Topological indices, NNRTI's, Physicochemical parameters, PETT derivative.

1. Introduction

Molecular modeling is a valuable and essential part of the drug design process that describes the generation, manipulation of the drug molecule. It also represents two dimensional and three dimensional structural features of the molecules and associated physicochemical and quantum molecular properties.[1-4]

This field has been called many things in different disciplines, but in general, it is a design of new molecules based on desired properties. In organo- matellic, pharmaceutical and chemical development, this effort is focused on modeling the chemical and physical features that characterize the various functions of the system, so that better binds, and therefore, more potent or precise feature can be developed. Evolutionary technique can help achieve the design of totally new molecules, some of that were never even thought of before. [5]

In the past few decades there has been a hiatus in the momentum of research and discovery of a "novel medicinal compound". The particular drug development perhaps is augmented due to two vital factors, first strict empirical and rational approach to drug design that incorporate massive screening of large corporate libraries of synthesized or naturally-occurring compounds. It may be considered as an integrated whole approach which essentially involves various steps namely chemical synthesis, evolution for activity spectrum, toxicological studies, metabolism of drug i.e. biotransformation and the study of metabolites formed, assay procedures and lastly galencial formulation and bio-pharmaceutics and secondly high standard of safely and therapeutic efficacy together with tremendous increased costs of research and development and finally the clinical trials.[8] In the light of these processes drug discovery was mainly the result of chance discovery.

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.

In the present study molecular modeling and Computer Aided Drug Design (CADD) approach is used to understand the role of structural features in the development of the new biologically active chemical entity belongs to well known family of PETT (Phenyl Ethyl Thaio-Azail Thaio-Urea) derivatives as NNRTIs (Non-Nucleoside Reverse Transcriptase Inhibitors).[6] . Several heterocyclic thioureas have been already reported as a new class of potent NNRTIs such as phenethylthiazolyl-thiourea (PETT) derivatives. [7-10] Uckun and co-worker[11] described the synthesis of a series of thiazole thioureas with alkyl, aryl, heteroaryl substituent's 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.[12]

2. Experimental and Methodology

Quantitative structure-activity relationships (QSAR) have been established for set of 23 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 [6].

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)[13], Randic′ connectivity index(χ)[14], Balaban index(J)[15], Szeged index(Sz)[16], Shultz molecular topological index(MTI)[17] and Electrotopological index(S)[23] 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.

3. Result and Discussion

As mentioned in the introduction, to analyse the relationship between biological activity and the structure of the molecule and various molecular properties of PETT derivatives mentioned in Table1, we tested the parameters representing the various structural features along with the specific substitution on the molecule, in Table 2. Indicator parameters are used in present study to explore the effect of substitution of specific type and positions. The univariate correlation in form of correlation matrix is presented in Table 3.

From the perusal of Table 3, except fifth order connectivity 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 also not having the significant statistical value.

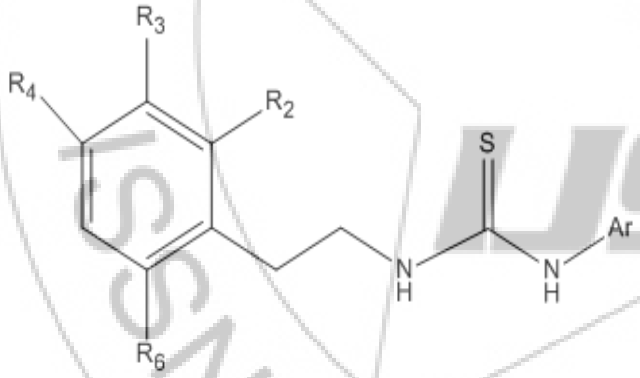

Figure 1: Parent Structure of PETT Derivatives

To understand the real role of structure in modeling the NNRTI activity of these PETT derivatives we have tested the different bi, tri and tetra parametric combination of various parameters representing the specific structural features. Amongst the various bi parametric combinations the combination of fifth order connectivity index and indicator parameter representing the methyl substitution having the significant statistical value and explore the role of these two structural features in NNRT inhibition activity.

The model obtained from the combination of these descriptors is as below

$\log IC_{50}$ = 0.8900(±0.3569)χ5 + 0.3114 (±0.1628) I_{Me} − 2.0432 (Eq. 1)
n = 23, Se = 0.3114, R = 0.5711, F = 4.841, Q = 1.834

From the perusal of eq. 1 it is observed that the connectivity in the molecule playing leading role over the presence of methyl group as substitution in directing the biological activity. Correlation value of both the parameters seems to be linear with biological activity numerically but the value of regression is not adequate to explain the behavior of these structural features in guiding the anti HIV activity as NNRTI's. Preliminary information exposed by the eq. is, the increase in the order of connectivity in the PETT derivatives may decrees the inhibition of NNRT and may not having the good activity against HIV. Same with the methyl substitution on the molecule, i.e., presence of methyl may not favor the anti HIV activity of the molecule due to its bulky nature or may increase the order of connectivity in the molecule.

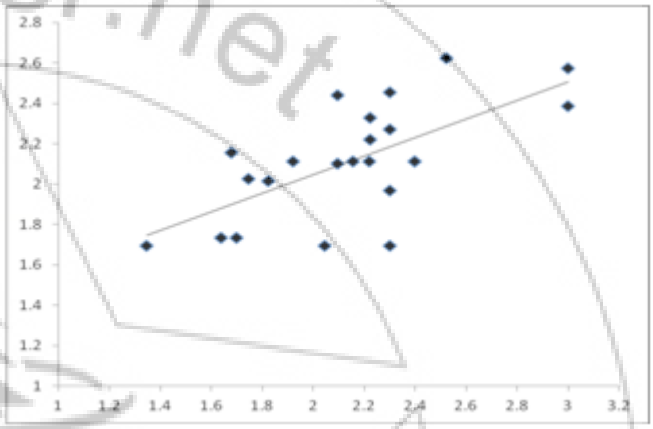

Figure 2: Graph obtained between obs. and calculated $\log IC_{50}$

Further elucidation of structure requires the addition of other descriptor in the model. Thus various tri parametric combinations are tested and with addition to the eq.1.

The best model obtained from the combination of descriptors tested in eq.1 and Balaban branching index. The model obtained is as fallows

$\log IC_{50}$ = -6.9322(±3.6695) J + 1.3410(±0.4122) χ5 + 0.3269(±0.1535) I_{ME} + 7.5527 (Eq. 2)
n = 23; Se = 0.3227; R = 0.6578; F = 4.831; Q = 2.038

Eq. 2 exhibits the similar findings with slightly improved statistics along with the elucidation of branching features in directing the anti HIV activity specifically inhibition of NNRT's. Model shows the favorable presence of branching for the studied biological activity. The substitution or any structural change in molecule that increases the branching in molecule may increase the inhibition activity of compounds against the NNRT's. With the addition of descriptor in eq.1 also affect the magnitude of other parameters in model. With the branching index the role of connectivity increases but at the very same time role of methyl substitution decreases in modeling the activity. Model having the two compounds comp. no. 2 and 21 behaving exceptional (on the basis of residue) to their parent series of compounds. Thus these two compounds are excluded from the calculation and model obtained for the set of 21 compounds is as below

$\log IC_{50}$ = -7.5712(±3.2464) J + 1.3120(±0.3865) χ5 + 0.4046(±0.1304) I_{ME} + 8.6865 (Eq. 3)

n = 21; Se = 0.2693; R = 0.7363; F = 6.711; Q = 2.734

Improvement in the statistics justify the deletion of compound no 2 and 21 from the calculation. It also exhibits the increase in magnitude of descriptor IME along with the decrease in the magnitude of connectivity parameter that may indicate the behavioral deviation of descriptors from linear to non-linear.

This model also having the compound with exceptional behavior (on the basis of residue) and after the deletion of that compound model obtained is as below

$\log IC_{50}$ = 6.5940(±3.1298)J + 1.3879(±0.3690)χ5 + 0.3703(±0.1252) I_{ME} + 6.7227 (Eq. 4)
n = 20; Se = 0.2552; R = 0.7640; F = 7.478; Q = 2.99

Model obtained after the deletion of compound no 20 explore the exceptional behavior of the compound. Also it exhibits the change in the magnitude of the descriptors that guide the study towards the nonlinear relationship between descriptor and the inhibition activity of the compounds. Comparison between the magnitudes of descriptors in the model, explore the dominance of connectivity in directing or characterizing the inhibition of NNRT's. Direction of relationship between the connectivity and biological activity remain same in all the equations with the change in magnitude shows the behavioral change of descriptor with change in structure and other associate parameters. Same thing happen with the parameter indicating the role of presence of methyl substitution on the compounds. This change with decries and increase in magnitude explore the compound dominating behavior of the descriptors. Surprisingly the orientation of relationship is changed in case of branching parameter from eq. 3 to eq. 4 it shows the non linear behavior of the descriptor in respect of biological activity studied.

It is also shown by the model that the methyl substitution or branching that may increase the order of connectivity is unfavorable for biological or anti HIV activity up to certain extend. The behavior of indicator parameter in the models demonstrate the hindrance between the interaction of active site and the chemical entity due to presence of bulky substitution and additional branching contain the alkyl substitution or aromatic substitution. As shown by the eq. 3 and 4 branching having the nonlinear relation with inhibition activity may favors the inhibition of NNRT's if contain the halogen substitution and hetro-atoms in branching or in substitution. The result obtained from the eq. 4 are recorded in table 4 along with residue values and graphically presented in figure 2.

4. Conclusion

On the basis of result and discussion made above conclusion can be drawn that the NNRT's inhibition activity is highly structure dependent for the series of PETT derivatives under study. Branching and order of the connectivity in the molecule plays the leading role in characterizing the inhibition activity or anti HIV activity. Increase in the order of connectivity the biological activity for the studied series. For the studied set of compounds Me substitution minimize the inhibition activity with bulky nature.

Reference

[1] RAMA RAO NADENDLA, MOLECULAR MODELING: A POWERFUL TOOL FOR DRUG DESIGN AND MOLECULAR DOCKING , *RESONANCE*, 2004, 9(5), 51-60.
[2] Lipkowitz, K. B. and McCracken R. O., Molecular modeling: a tool for predicting anthelmintic activity in vivo Parasitology Research,1993, 79(6), 475-479, ORIGINAL INVESTIGATIONS.
[3] Thakur,M., Ojha, L., Thakur, A. and Solanki, B.K., Study of Active Site and SAR on Some Benzodiazepines, *J. Comp. Biology & Bioinformatics*, 2010, 2(2), 10-19.
[4] Thakur,A. and Thakur,M., Chem-Bioinformatics: Computational modeling of TIBO Derivatives, *Anti Inf. Agents in Medicinal Chem*, 2010, 9 (2), 52-58.
[5] GERLAND,U. AND HWA,T. *EVOLUTIONARY SELECTION BETWEEN ALTERNATIVE MODES OF GENE REGULATION*, *PNAS*, 2009, 106, 8841-8846 .
[6] Ravichandrana,V., Mouryab ,V. K., Agrawala, R. K., QSAR Modeling of HIV-1 Reverse Transcriptase Inhibitory Activity with PETT Derivatives, *Digest Journal of Nanomaterials and Biostructures*, 2008, 3(1), 9 – 17.
[7] Walid, F., Syntheses and Reactions of Methyl [3-(4-phenyl-thiazol-2-yl)-Thioureido] Alkanoates and Related Compounds, *ARKIVOC* , 2008 (xii) 245-255.
[8] Hasegawa, K., Kimura, T. and Funatsu, K. ,GA Strategy for Variable Selection in QSAR Studies: Application of GA-Based Region Selection to a 3D-QSAR Study of Acetylcholines- terase Inhibitors. *J. Chem. Inf ComputSci.*, 1999,39,112-1209.
[9] Ren, B., A New Topological Index for QSPR of Alkanes. *J.Chem.Inf.Comput.Sci*., 1999, 39, 139-143.
[10] Balaban, A.T. (1988b). Modeling the Chemical Structures Through Graphs & of These Graphs Through Topological Indices. *Rev. Chimie (Bucarest*), 1988,39,1026-1031.
[11] Hong-Ping Xie ,Jian-Hui Jiang ,Hui Cui ,Guo-Li Shen and Ru-Qin Yu, A new redundant variable pruning approach—minor latent variable perturbation–PLS used for QSAR studies on anti-HIV drugs,*Computers & Chemistry*,2002,26(6), 591–600.
[12] Wiener, H., *Correlation of Heat of Isomerization, and Differences in Heats of Vaporiza-tion of Isomers, among the Paraffin Hydrocarbons. J.Am. Chem.Soc.*, 1947, *69,2636-2638*.
[13] Randic ,M., *On Characterization of Molecular Branching. J.Am. Chem.Soc.*, 1975, *97, 6609- 6615*.
[14] Balaban, A. T. From Chemical Graphs to 3D Molecular Modeling. In From Chemical Topology to Three-Dimensional Geometry, *Ed.; Plenum: New York*, 1997, 1-24.
[15] Khadikar,P.V.,The Szeged Index and an Analogy with the Wiener Index. *J.Chem.Inf.Comput.Sci*., 1995,35, 547-550.
[16] Shultz, H.P., *Topological Organic Chemistry. 1. Graph Theory and Topological Indices of Alkanes, J. Chem.Inj.Comput.Sci.*, 1989, 29,227-228.

[17] Kier, L.B.and Hall, L.H., *The E-State as an Extended Free Valence. J.Chem.Znj.Com.Sci,* 1997, 37,548-552.
[18] ACD-Lab software for calculating the referred physicochemical parameters; Chemsketh S.O, *WWW. Acdlabs.com.*
[19] Chaterjee,S., Hadi,A.S. and Price,B., Regression Analysis by Examples, *3rd ed. Wiley VCH:New york,* 2000.

Dr. Arpan Bharadwaj: He is Professor of Chemistry in Madhav Science PG College, Ujjain, India. He has published 5 research papers in International and National Journals.

Dr. Suresh Kumar Verma: He is Professor of Chemistry in Govt.S. K. P. G. College, Sikar. He has cleared NET (CSIR) 27 times and GATE, 13 times.

Author Profile

Dr. Megha Choure: She is a working as guest lecturer in the Madhav Science P.G. College, Ujjain, India, She has completed her Ph.D research work in the same Institute.

Dr. Nameeta Bende: She is an Asst. Professor of Chemistry in Govt. Holkar Science College, Indore, India. She has presented 15 research papers in International and National Conferences all over India.

Table 1: Various substituent's for PETT derivatives with their biological activity.

Comp	R^2	R^3	R^4	R^6	Ar	$logIC_{50}$
1	F	$(CO)N(Me)_2$	H	F	5-bromo-2-pyridyl	1.823
2	F	CH_2NAc	H	F	5-bromo-2-pyridyl	3
3	F	CN	H	F	5-chloro-2-pyridyl	1.638
4	F	$N(Me)_2$	H	F	5-chloro-2-pyridyl	1.346
5	F	$N(Me)_2$	H	F	5-bromo-2-pyridyl	2.045
6	F	OCH_3	H	F	5-bromo-2-pyridyl	2.096
7	F	OC_2H_5	H	F	5-bromo-2-pyridyl	2.154
8	F	CH_2OCH_3	H	F	5-bromo-2-pyridyl	2.221
9	Cl	OC_2H_5	H	F	5-bromo-2-pyridyl	2.397
10	Cl	OC_2H_5	H	F	5-chloro-2-pyridyl	2.397
11	Cl	OC_2H_5	H	F	5-iodo-2-pyridyl	1.921
12	Cl	OC_2H_5	H	F	5-cyano-2-pyridyl	2.221
13	H	OCH_3	H	OCH_3	5-chloro-2-pyridyl	2.096
14	H	OC_2H_5	H	OC_2H_5	5-bromo-2-pyridyl	2.301
15	F	H	H	OC_2H_5	5-bromo-2-pyridyl	2.301
16	F	F	H	OC_2H_5	5-bromo-2-pyridyl	1.745
17	F	F	H	OCH_3	5-bromo-2-pyridyl	2.221
18	F	OCH_3	H	OCH_3	5-chloro-2-pyridyl	2.301
19	F	OC_2H_5	H	OCH_3	5-chloro-2-pyridyl	2.522
20	OCH_3	OCH_3	H	F	5-bromo-2-pyridyl	3
21	F	$N(Me)_2$	H	F	5-bromo-2-pyridyl	2.301
22	F	CN	H	F	5-bromo-2-pyridyl	1.698
23	Cl	OC_2H_5	Cl	F	5-bromo-2-pyridyl	1.677

Table 2: Different descriptors tested in the present study for the PETT derivatives.

Comp. No.	J	$\chi 5$	I_{ME}
1.	1.740	4.876	0
2.	1.688	4.897	0
3.	1.666	4.320	0
4.	1.696	4.434	0
5.	1.696	4.434	0
6.	1.666	4.320	1
7.	1.676	4.641	0
8.	1.676	4.641	0
9.	1.676	4.641	0
10.	1.676	4.641	0
11.	1.676	4.641	0

12.	1.664	4.739	0
13.	1.640	4.441	1
14.	1.676	4.889	0
15.	1.657	4.447	0
16.	1.701	4.697	0
17.	1.680	4.471	1
18.	1.705	4.627	1
19.	1.719	4.948	1
20.	1.709	4.863	1
21.	1.696	4.434	0
22.	1.666	4.320	0
23.	1.717	4.868	0

* J = Balaban Branching Index; $\chi 5$ = Fifth order connectivity index; I_{ME} = Indicator parameter 1 if Me substitution is present.

Table 3: Observed and calculated values of $logIC_{50}$ with residue

Comp. No.	$logIC_{50}$(Obs.)	z_{50}(Calc.)	Residue
	1.823	2.016	-0.193
	3.000	2.388	0.611
	1.638	1.733	-0.095
	1.346	1.693	-0.347
	2.045	1.693	0.352
	2.096	2.103	-0.007
	2.154	2.112	0.042
	2.220	2.112	0.108
	2.397	2.112	0.285
	2.397	2.112	0.285
	1.921	2.112	-0.191
	2.221	2.327	-0.106
	2.096	2.442	-0.346
	2.301	2.456	-0.155
	2.301	1.968	0.332
	1.745	2.025	-0.280
	2.221	2.220	0.001
	2.301	2.272	0.029
	2.522	2.625	-0.103
	3.000	2.573	0.427
	2.301	1.693	0.608
	1.698	1.733	-0.035
	1.677	2.157	-0.480