Asian Journal of Biochemistry

ISSN 1815-9923

Academic
Journals Inc.

www.academicjournals.com

Asian Journal of Biochemistry 5 (4): 238-254, 2010
ISSN 1815-9923 / DOI: 10.3923/ajb.2008.62.78

2D and 3D QSAR: Modeling of TIBO Derivatives as Reverse Transcriptase 1 Inhibitors*

[1]A. Thakur, [2]B.K. Tiwari, [3]M. Thakur, [2]N.D. Pandey,
[2]S.S. Narvi, [4]S. Thakur and [5]A. Bharadwaj
[1]Department of Chemistry, Rishiraj Institute of Technology, Indore, India
[2]Department of Chemistry, MNNIT, Allahabad, India
[3]Department of Chemistry, Ujjain Engineering College, Ujjain, India
[4]Department of Physics, Shri Vaishnav Institute of Technology and Science, Indore, India
[5]Department of Chemistry, Govt. Madhav Science College, Ujjain, India

Abstract: This report describes QSAR and SAR studies on the Inhibition of Reveres Transcriptase (RT) by 79 TIBO (Tetrahydoimidazobenzodizepin-2-one) derivatives using both classical and unconventional physicochemical properties and quantum molecular descriptors along with indicator parameters. The application of a multiple linear regression analysis indicated that a combination of classical physicochemical descriptors and the indicator parameters yielded a s tatistically significant model for the activity, log 1/C (50% of inhibition concentration of TIBO derivatives for RTs). The final selection of a potential TIBO compound for the inhibition of Reveres Transcriptase is made by quantum molecular modeling. We have found that, among the a number of Quantum and modeling parameters, the electron density on the 9th atom correlated best with the activity.

Key words: Reveres transcriptase, QSAR, correlation

INTRODUCTION

The Acquired Immune-Deficiency Syndrome (AIDS) is one of the most hazardous diseases, which is caused by infection with the Human Immunodeficiency Virus (HIV). Reverse Transcriptase (RT) is the key for HIV replication and is not required for normal host cell replication. During the inspection of effective therapies facing HIV, Reverse Transcriptase (RT) has been identified as one of the most promising targets (Garg *et al*., 1999).

Substituted TIBO (tetrahydoimidazobenzodizepin-2-one) derivatives find extensive applications in inhibition of reverse transcriptase (Garg *et al*., 1999). It is, thus, important to understand the potential inhibition activity, electronic and steric features. Thus, there is a compelling need to understand the mechanisms and correlation modes of potential inhibition activity with reverse transcriptase. Over the past several decades, the Hansch group has studied the effect and significance of the molecular, steric and hydrophobic parameters in the modeling of a number of biological activities of numerous compounds. In a review based on the QSAR study on NNRTI (non-nucleoside reverse transcriptase inhibitors), the Hansch group (Garg *et al*., 1999) has reported that the use of steric and molecular parameters yielded excellent statistics on the large set of TIBO derivatives. Inspired by the pioneering work of Hansch, (Garg *et al*., 1999) and in the continuation of our earlier works, (Wlodawer, 2002; Barre *et al*., 1983; Gallo, 1984; De Clercq, 1992; Mitsuya and Broder, 1986; Reardon and Miller, 1990; Richman *et al*., 1987; Fischl *et al*., 1989; Debyser *et al*., 1992) we have revisited the work of Hansch (Garg *et al*., 1999) to see if we can further develop a significant QSAR

Corresponding Author: A. Thakur, 205, Sudama Nagar Anp. Sec. Indore 452009, India
Tel: 91-734-2526885 Fax: 91-731-2763618

*Originally Published in Asian Journal of Biochemistry, 2008

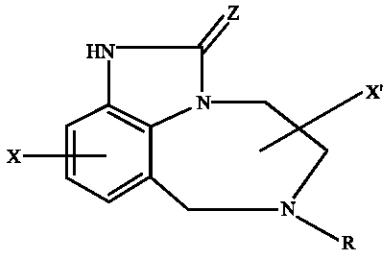

Fig. 1: Parent structure of TIBO derivative used in present study

model with an entirely different set of parameters. To achieve this objective, we have used a large set of Molecular descriptors such as Molar Refractivity (MR), Molar Volume (MV), Parachore (Pc), the Refractive index (ç), Surface Tension (ST), Density (D), Hydration Energy (HE), Approximate Surface Area (ASA), Surface Area Grid (SAG) and quantum chemical parameters, along with the indicator parameters used for the structural and positional specifications.

To obtain a statistically significant model, we have used a number of methods including the maximum R^2 method, which was followed by stepwise regression analyses (Loya *et al*., 1992, 1994, 1995a, b; Chaterjee and Hadi, 2000).

We note that the maximum R^2 method actually includes a combination of standard error, adjusted R^2 value (R_A^2), R, standard error of estimation and the F-ratio value. The predictive ability of the model is discussed on the basis of the predictive correlation coefficients. To validate our model further, we have used quantum molecular modeling parameters and on the basis of these parameters, we have analyzed the structural behavior of these molecules. For the molecular modeling, we have optimized the geometries of molecules using the molecular mechanics method by applying the MM+ force field method and Huckel molecular orbital theory using CNDO (Complete neglect of differential overlap) methods. Our motivation for the Huckel molecular orbital theory study is that the electronic and quantum parameters strongly depend on the degree of sophistication and such electronic parameters vary with the degree of the level of theory. The accuracy of a molecular mechanics or quantum mechanical method depends on the database used to parameterize the method.

MATERIALS AND METHODS

Activity

The log 1/C (IC_{50}) of the TIBO derivatives (Fig. 1) is adopted from the literature (Garg *et al*., 1999).

Physicochemical Parameters

The Molar volume, Parachor, Molar Refractivity, Refractive Index, Surface Tension, Density and PolarI Zability, for the set of TIBO derivatives were calculated from ACD Lab software (www.acdlabs.com) and the un-conventional physicochemical parameters Approximate Surface area, Surface area grid and Hydration Energy were calculated using hyperchem 7 demo-version. Molecular modeling parameters were calculated by applying MM^+ force field (Molecular mechanics) using Hyperchem 7 demo-version (www.hyper.com).

Indicator Parameters

Indicator parameters are the dummy parameters sometimes used for accounting those structural feature not covered in any molecular descriptor used. They assumed only two values

1 or 0. If the assumed structural feature is present; then the indicator parameters are 1 otherwise it is 0. The details of such parameters, used in the present study are already given in the Result and Discussion section (Table 2).

Statistical Analysis

Maximum R^2 method together with stepwise regression (Chaterjee and Hadi, 2000) was carried for arriving at statistically significant models. In present study linear mathematical models are developed to study Quantitative Structure/Property-Activity Relationship. Multiple linear regression analysis is used to develop these models.

RESULTS AND DISCUSSION

The set of 79 TIBO derivatives and their adopted activity i.e., the inhibition activity of Reverse transcriptase-1 had been expressed as log 1/C is shown in Table 1.

A very low-level degeneracy is present in the activity log1/C (mol L^{-1}). As a result of the occurrence of degeneracy in activity log1/C, it becomes essential to examine the degeneracy in the molecular descriptors also (Table 1). A perusal of Table 2 and 3, which contains the unconventional physicochemical parameters and classical physicochemical parameters calculated for TIBO derivatives, shows that the low level degeneracy is observed in the unconventional and classical physicochemical descriptors (Balaban, 1992; Balaban and Balaban, 1991) has shown that the indices/descriptors, in spite of their degeneracy, can be used successfully in developing statistically significant QSAR models.

The correlation among the descriptors like unconventional physicochemical parameters (Table 2), classical physicochemical properties (Table 3), indicator parameters (Table 2), logP (Table 2) and activities shows (Table 4) that, except for the MR, logP, SAG and indicator parameters I_Z and I_R, all other unconventional and classical physicochemical parameters do not correlate well with the biological activity log1/C.

Out of the set of unconventional physicochemical descriptors used, Initial bi-parametric regression analyses indicate that the combination of surface area grid and the indicator parameter I_Z plays the significant role in modeling the activity log1/C. But the statistics obtained from this combination is not adequate to explain the structure activity relationship. In the case of trivariate correlations, the combination of SAG, I_Z, I_R resulted little better but not as required. In case of tetra and penta variate correlations the results are encouraging and the best results obtained from the tetra variate combination of SAG, I_Z, I_R, I_X with the biological activity log1/C. The model obtained from the above variables is:

Table 1: Substituents and biological activity of TIBO derivatives used in present study

S. No.	X	Z	R	X'	Obs.log1/C (mol L^{-1})
1	H	S	DMA	5-Me	7.36
2	9-Cl	S	DMA	5-Me	7.47
3	8-Cl	S	DMA	5-Me	8.37
4	8-F	S	DMA	5-Me	8.24
5	8-SMe	S	DMA	5-Me	8.30
6	8-OMe	S	DMA	5-Me	7.47
7	8-OC_2H_5	S	DMA	5-Me	7.02
8	8-CN	O	DMA	5-Me	5.94
9	8-CN	S	DMA	5-Me	7.25
10	8-CHO	S	DMA	5-Me	6.73
11	8-$CONH_2$	O	DMA	5-Me	5.20
12	8-Br	O	DMA	5-Me	7.33
13	8-Br	S	DMA	5-Me	8.52
14	8-I	O	DMA	5-Me	7.06
15	8-I	S	DMA	5-Me	7.32
16	8-C≡CH	O	DMA	5-Me	6.36
17	8-C≡CH	S	DMA	5-Me	7.53
18	8-Me	O	DMA	5-Me	6.00

Table 1: Continued

S.No.	X	Z	R	X'	Obs.log1/C (mol L^{-1})
19	8-Me	S	DMA	5-Me	7.87
20	9-NO_2	O	CPMb	5-Me	4.48
21	8-NH_2	O	CPM	5-Me	3.07
22	8-NMe_2	O	CPM	5-Me	5.18
23	9-NH_2	O	CPM	5-Me	4.22
24	9-NMe_2	O	CPM	5-Me	5.18
25	9-NHCOMe	O	CPM	5-Me	3.80
26	9-NO_2	S	CPM	5-Me	5.61
27	9-F	S	DMA	5-Me	7.60
28	9-CF_3	O	DMA	5-Me	5.23
29	9-CF_3	S	DMA	5-Me	6.31
30	9-Me	O	DEA c	5-Me	6.50
31	10-OMe	O	DMA	5-Me	5.18
32	10-OMe	S	DMA	5-Me	5.33
33	9,10-di-Cl	S	DMA	5-Me	7.60
34	10-Br	S	DMA	5-Me	5.97
35	H	O	$CH_2CH = CH_2$	5-Me	4.15
36	H	O	2-MA	5-Me	4.33
37	H	O	CH_2CO_2Me	5-Me	3.07
38	H	O	$CH_2C{\equiv}CH$	5-Me	3.24
39	H	O	CH_2-2-furanyl	5-Me	3.97
40	H	O	$CH_2CH_2CH = CH_2$	5-Me	4.30
41	H	O	$CH_2CH_2CH_3$	5-Me	4.05
42	H	O	CPM	5-Me	4.36
43	H	O	$CH_2CH = CHMe$	5-Me	4.24
44	H	O	$CH_2CH_2CH_2Me$	5-Me	4.00
45	H	O	DMA	5-Me	4.90
46	H	O	$CH_2C(Br) = CH_2$	5-Me	4.21
47	H	O	$CH_2C(Me) = CHMe$	5-Me	4.54
48	H	O	$CH_2C(C_2H_5) = CH_2$	5-Me	4.43
49	H	O	$CH_2CH = CHC_6H_5$	5-Me	3.91
50	H	O	$CH_2C(CH = CH_2) = CH_2$	5-Me	4.15
51	8-Cl	S	DMA	H	7.34
52	9-Cl	S	DMA	H	6.80
53	H	O	2-MA	5,5-di-Me	4.64
54	H	O	2-MA	4-Me	4.50
55	9-Cl	S	2-MA	4-Me	6.17
56	9-Cl	S	CPM	4-Me	5.66
57	H	O	C3H7	4-$CHMe_2$	4.13
58	H	O	2-MA	4-$CHMe_2$	4.90
59	H	O	2-MA	4-C_3H_7	4.32
60	H	O	DMA	7-Me	4.92
61	8-Cl	O	DMA	7-Me	6.84
62	9-Cl	O	DMA	7-Me	6.80
63	H	S	C_3H_7	7-Me	5.61
64	H	S	DMA	7-Me	7.11
65	8-Cl	S	DMA	7-Me	7.92
66	9-Cl	S	DMA	7-Me	7.64
67	H	O	DMA	4,5-di-Me	4.25
68	H	S	DMA	4,5-di-Me	5.65
69	H	S	CPM	4,5-di-Me	4.87
70	H	S	DMA	5,7-di-Me	5.94
71	9-Cl	O	DMA	5,7-di-Me	6.64
72	9-Cl	S	DMA	5,7-di-Me	6.32
73	H	S	DMA	4,7-di-Me	4.59
74	9-Cl	O	DMA	5-Me	6.74
75	9-Cl	S	CPM	5-Me	7.47
76	H	S	CPM	5-Me	7.22
77	H	O	C_3H_7	5-Me	4.22
78	H	S	C_3H_7	5-Me	5.78
79	H	S	2-MA	5-Me	7.59

Table 2: Unconventional physicochemical parameters, logP and indicator parameters used in present study

Comp. No.	ASA	SAG	HE	logP	I_Z	I_R	I_X
1	438.76	510.84	-3.17	1.738	1	1	0
2	480.00	534.89	-2.87	2.430	1	1	1
3	469.77	523.81	-2.89	2.430	1	1	1
4	448.16	503.91	-2.99	2.031	1	1	1
5	414.26	510.37	-2.30	2.323	1	1	0
6	490.10	540.80	-3.69	1.591	1	1	0
7	524.58	571.08	-3.01	2.076	1	1	0
8	469.95	505.20	-5.48	1.111	0	1	0
9	481.32	527.20	-6.74	1.141	1	1	0
10	461.87	527.72	5.14	1.237	1	1	0
11	393.00	504.26	-6.66	0.191	0	1	0
12	387.23	469.37	-1.49	2.662	0	1	1
13	479.12	532.41	-2.88	2.692	1	1	1
14	479.58	525.02	-1.60	3.047	0	1	1
15	491.90	540.99	-2.87	3.077	1	1	1
16	419.85	496.18	-1.63	2.003	0	1	0
17	509.10	537.95	-2.84	2.003	1	1	0
18	371.28	468.52	-1.02	2.172	0	1	0
19	370.01	477.08	-1.82	2.202	1	1	0
20	411.98	488.62	-6.74	0.335	0	0	0
21	306.74	447.64	-4.91	-0.685	0	0	0
22	385.65	489.37	-1.31	0.777	0	0	0
23	366.60	474.32	-6.87	-0.658	0	0	0
24	438.56	513.99	-1.03	0.777	0	0	0
25	430.28	517.59	-4.13	-0.477	0	0	0
26	412.13	498.06	-7.87	0.365	1	0	0
27	455.10	515.89	-2.91	2.031	1	1	1
28	414.89	489.32	-1.44	2.761	0	1	1
29	412.85	500.28	-2.36	2.791	1	1	1
30	513.12	561.28	-0.16	2.968	0	0	0
31	415.83	492.11	-2.49	1.555	0	1	0
32	397.44	497.21	-3.15	1.585	1	1	0
33	510.00	551.10	-2.36	3.655	1	1	1
34	391.52	489.44	-2.11	3.075	1	1	1
35	355.11	439.45	-3.75	0.456	0	0	0
36	255.11	421.32	-3.07	1.033	0	0	0
37	382.83	456.90	-3.99	-2.285	0	0	0
38	338.89	418.83	-2.86	-0.009	0	0	0
39	387.71	494.47	-3.51	0.305	0	0	0
40	368.66	464.20	-3.62	0.863	0	0	0
41	378.80	447.14	-1.90	0.846	0	0	0
42	344.06	457.18	-2.12	0.609	0	0	0
43	394.49	476.00	-2.59	1.176	0	0	0
44	413.90	480.63	-1.52	1.382	0	0	0
45	425.09	485.69	-1.85	1.753	0	1	0
46	345.25	454.58	-4.14	1.107	0	0	0
47	412.07	480.63	-1.99	1.670	0	0	0
48	359.47	471.84	-3.17	1.440	0	0	0
49	288.45	477.24	-3.26	2.419	0	0	0
50	252.48	429.32	-3.24	1.179	0	0	0
51	455.80	507.10	-3.25	2.342	1	1	1
52	464.89	523.93	-3.24	2.342	1	1	1
53	382.28	457.13	-2.58	1.263	0	0	0
54	377.94	459.86	-2.82	1.239	0	0	0
55	426.51	501.45	-3.49	1.916	1	0	1
56	417.66	505.57	-2.74	1.198	1	0	1
57	382.47	476.26	-1.29	1.807	0	0	0
58	285.38	447.06	-2.46	2.236	0	0	0
59	289.58	458.80	-2.47	2.311	0	0	0
60	424.23	489.13	-1.68	1.993	0	1	0
61	449.59	500.82	-1.49	2.640	0	1	1

Table 2: Continued

Comp. No.	ASA	SAG	HE	logP	I_Z	I_R	I_X
62	358.69	465.60	-1.51	2.640	0	1	1
63	387.00	469.63	-3.24	1.116	1	0	0
64	439.40	512.45	-3.16	2.023	1	1	0
65	400.83	429.54	-2.88	2.670	1	1	1
66	477.59	529.67	-2.74	2.670	1	1	1
67	431.07	501.04	-1.53	2.047	0	1	0
68	454.76	523.94	-2.51	2.077	1	1	0
69	368.23	483.99	-2.53	0.933	1	0	0
70	446.09	517.48	-2.74	2.111	1	1	0
71	400.91	485.98	-1.21	2.728	0	1	1
72	401.09	497.75	-1.98	2.758	1	1	1
73	437.23	514.57	-2.51	2.317	1	1	0
74	465.94	511.87	-1.51	2.400	0	1	1
75	411.05	502.99	-3.13	1.268	1	0	1
76	369.00	480.22	-3.44	0.639	1	0	0
77	353.44	444.40	-2.05	0.846	0	0	0
78	393.56	471.57	-3.20	0.876	1	0	0
79	268.10	436.50	-4.12	1.063	1	0	0

ASA = Approximate Surface Area ($A^{\circ 2}$), SAG = Surface Area Grid ($A^{\circ 2}$), HE = Hydration Energy (kcal mol^{-1}), logP = Hydrophobic parameter (Octanol/water Partition coefficient) $I_Z = 1$ if S atom at Z position, $I_R = 1$ if Acyclic structure at R position, $I_X = 1$ if halogens present at X position

Table 3: Classical physicochemical properties used in present study for TIBO derivatives

Comp. No.	MR	MV	Pc	η	ST	D	α
1	87.14	235.2	651.1	1.662	58.7	1.22	34.54
2	91.97	246.1	688.2	1.670	61.1	1.30	36.46
3	91.97	246.1	688.2	1.670	61.1	1.30	36.46
4	87.25	239.8	658.4	1.648	56.8	1.27	34.59
5	99.78	264.8	741.8	1.677	61.5	1.25	39.55
6	93.50	257.1	709.7	1.647	58.0	1.23	37.07
7	98.14	273.3	749.8	1.637	56.6	1.21	38.90
8	84.71	236.5	657.3	1.635	59.6	1.25	33.58
9	91.71	244.6	698.7	1.672	66.5	1.27	36.35
10	92.11	248.0	698.3	1.664	62.7	1.27	36.51
11	88.51	243.3	682.8	1.647	61.9	1.29	35.08
12	87.86	239.9	660.7	1.653	57.5	1.45	34.83
13	94.86	248.0	702.1	1.690	64.2	1.47	37.60
14	93.06	246.4	683.6	1.679	59.1	1.61	36.89
15	100.06	254.6	725.0	1.715	65.7	1.62	39.66
16	87.20	244.0	667.4	1.633	55.9	1.21	34.57
17	94.20	252.1	708.9	1.670	62.4	1.23	37.34
18	84.77	242.8	647.9	1.615	50.6	1.17	33.60
19	91.76	251.0	689.4	1.651	56.8	1.20	36.37
20	86.18	238.3	666.7	1.642	61.1	1.32	34.16
21	77.36	203.2	580.1	1.686	66.3	1.34	30.67
22	86.88	238.5	656.8	1.648	57.4	1.25	34.44
23	77.36	203.2	580.1	1.686	66.3	1.34	30.67
24	86.88	238.5	656.8	1.648	57.4	1.25	34.44
25	86.77	236.6	663.9	1.654	61.9	1.32	34.39
26	86.77	219.5	650.6	1.720	77.1	1.45	34.40
27	87.25	239.8	658.4	1.648	56.8	1.27	34.59
28	85.13	258.2	671.6	1.573	45.7	1.31	33.75
29	91.76	251.0	689.4	1.651	56.8	1.20	36.37
30	94.03	275.3	728.0	1.598	48.9	1.13	37.27
31	86.51	248.9	668.2	1.611	51.8	1.21	34.29
32	93.50	257.1	709.7	1.647	58.0	1.23	37.07
33	96.79	257.0	725.4	1.676	63.4	1.38	38.37
34	94.86	248.0	702.1	1.690	64.2	1.47	37.60
35	70.97	196.6	534.7	1.641	54.6	1.23	28.13
36	75.37	212.7	571.1	1.626	51.8	1.20	29.88
37	73.06	204.4	571.1	1.633	60.8	1.34	28.96

Table 3: Continued

Comp. No.	MR	MV	Pc	η	ST	D	α
38	69.01	186.2	525.0	1.662	63.1	1.29	27.35
39	80.14	221.8	614.6	1.642	58.9	1.29	31.77
40	75.60	212.9	574.8	1.628	53.1	1.20	29.97
41	71.20	201.3	545.6	1.625	53.9	1.21	28.22
42	73.74	200.0	552.1	1.658	58.0	1.28	29.23
43	75.74	210.9	573.3	1.637	54.5	1.21	30.02
44	75.83	217.6	585.7	1.613	52.4	1.19	30.06
45	80.14	227.0	609.6	1.623	51.9	1.19	31.77
46	78.47	209.8	583.9	1.671	59.9	1.53	31.10
47	80.14	227.0	609.6	1.623	51.9	1.19	31.77
48	80.00	229.0	611.2	1.615	50.7	1.18	31.71
49	95.82	253.9	707.1	1.678	60.1	1.25	37.98
50	79.77	224.3	600.3	1.629	51.2	1.20	31.62
51	87.36	229.2	650.2	1.687	64.6	1.34	34.63
52	87.36	229.2	650.2	1.687	64.6	1.34	34.63
53	80.03	228.8	609.5	1.616	50.3	1.18	31.72
54	75.37	212.7	571.1	1.626	51.8	1.20	29.88
55	87.20	231.8	649.7	1.675	61.6	1.32	34.56
56	80.96	202.2	592.7	1.732	73.7	1.45	32.09
57	80.44	234.5	623.7	1.601	50.0	1.16	31.89
58	84.61	245.9	649.2	1.604	48.5	1.16	33.54
59	84.63	245.2	651.2	1.606	49.6	1.16	33.55
60	80.14	227.0	609.6	1.623	51.9	1.19	31.77
61	84.97	238.0	646.7	1.632	54.5	1.28	33.68
62	84.97	238.0	646.7	1.632	54.5	1.28	33.68
63	78.20	209.5	587.1	1.669	61.6	1.24	31.00
64	87.14	235.2	651.1	1.662	58.7	1.22	34.54
65	91.97	246.1	688.2	1.670	61.1	1.30	36.46
66	91.97	246.1	688.2	1.670	61.1	1.30	36.46
67	84.75	243.9	647.6	1.611	49.6	1.16	33.59
68	91.75	252.0	689.1	1.648	55.8	1.19	36.37
69	85.35	225.0	631.6	1.682	62.0	1.27	33.83
70	91.75	252.0	689.1	1.648	55.8	1.19	36.37
71	89.58	254.8	684.8	1.620	52.1	1.25	35.51
72	96.57	262.9	726.3	1.655	58.1	1.27	38.28
73	91.75	252.0	689.1	1.648	55.8	1.19	36.37
74	84.97	238.0	646.7	1.632	54.5	1.28	33.68
75	85.57	219.1	630.7	1.709	68.6	1.40	33.92
76	80.74	208.2	593.6	1.703	66.0	1.31	32.01
77	71.20	201.3	545.6	1.625	53.9	1.21	28.22
78	78.20	209.5	587.1	1.669	61.6	1.24	31.00
79	82.37	220.9	612.6	1.668	59.1	1.23	32.65

MR = Molar Refractivity (cm^3), MV = Molar Volume (cm^3), Pc = Parachore (cm^3), η = Index of Refraction, ST = Surface Tension (dynes cm^{-1}), D = Density(g cm^{-3}), α

$$\log 1/C = 0.0052(\pm 0.0036)SAG + 1.2207(\pm 0.2176)I_{ZS} + 0.9095(\pm 0.2390)I_R + 0.8256(\pm 0.2281)I_X + 1.9551 \quad (1)$$

n = 79, Se = 0.8158, R = 0.8407, R^2_A = 0.6907, F = 44.591

Similarly, in case of classical physicochemical parameters bi parametric correlation of MR and indicator parameter I_Z shows good potential to model the activity log1/C but not as much as required describing the structure activity relationship in quantitative manner. In case of tri, tetra and penta variate correlation combination of MR with the indicator parameters shows better results and best result obtained from correlation of MR, I_Z, I_R and I_X.

The model obtained from the above combination is below:

$$\log 1/C = 0.0389(\pm 0.0186)MR + 1.1267(\pm 0.2231)I_Z + 0.7675(\pm 0.2536)I_R + 0.7972\ (\pm 0.2251)\ I_X + 1.2993 \quad (2)$$

Table 4: Correlation matrix between unconventional, classical physicochemical parameters, logP, indicator parameters and log1/C

	log1/C	ASA	SAG	HE	I_Z	I_R	I_X
log1/C	1.00000						
ASA	0.56849	1.00000					
SAG	0.56786	0.89307	1.00000				
HE	0.13361	0.12506	0.09975	1.00000			
I_Z	0.68473	0.42716	0.48384	-0.06121	1.00000		
I_R	0.66077	0.60169	0.55579	0.21253	0.40039	1.00000	
I_X	0.57601	0.36062	0.29789	0.15093	0.33509	0.45454	1.00000

	MR	MV	PC	η	ST	D	Pol	I_{ZS}	I_R	I_X	log1/C
MR	1.000										
MV	0.905	1.000									
Pc	0.983	0.955	1.000								
ç	0.285	-0.145	0.135	1.000							
ST	0.222	-0.180	0.115	0.938	1.000						
D	0.226	-0.067	0.128	0.687	0.658	1.000					
Pol	1.000	0.905	0.983	0.285	0.222	0.226	1.000				
I_{ZS}	0.568	0.307	0.480	0.633	0.558	0.209	0.568	1.000			
I_R	0.667	0.669	0.680	0.032	-0.004	0.100	0.667	0.400	1.000		
I_X	0.392	0.270	0.339	0.310	0.225	0.534	0.392	0.335	0.454	1.000	
log1/C	0.687	0.539	0.636	0.384	0.292	0.256	0.687	0.684	0.660	0.576	1.000

	logP	I_{ZS}	I_R	I_X	log1/C
logP	1.00000				
I_{ZS}	0.31679	1.00000			
I_R	0.63070	0.40039	1.00000		
I_X	0.58102	0.33509	0.45454	1.00000	
log1/C	0.61599	0.68473	0.66077	0.57601	1.00000

n = 79, Se = 0.8037, R = 0.8458, R^2_A = 0.7001, F = 46.516

Same procedure was followed for the estimation of log1/C from the logP and the best model obtained from above variables is:

$$\log 1/C = 0.2713(\pm 0.126)\,\log P + 1.3211(\pm 0.2014)\,I_{ZS} + 0.8008(\pm 0.2433)\,I_R + 0.6022\,(\pm 0.2469)\,I_X + 4.1368 \quad (3)$$

n = 79, Se = 0.8023, R = 0.8464, R^2_A = 0.7011, F = 46.732

To confirm our results we compared the calc. log1/C values with observed ones shown in (Table 3 and Fig. 2, 3).

Substitutional effects are shown by the indicator parameters. Correlation matrix (Table 5-7) shows that all three indicator parameters I_Z, I_R and I_X are having good correlation coefficients (0.684, 0.660 and 0.576, respectively) individually with biological activity log1/C.

Equations suggest that the positive correlation coefficient of unconventional physicochemical parameter surface area grid, classical physicochemical property MR show direct relationship with biological activity log1/C. The hydrophobic parameter logP also shows positive correlation coefficient bears direct relationship with biological activity.

The positive correlation coefficient of indicators I_Z, I_R and I_X also shows positive impact on the biological activity quantitatively.

In view of this above, we have concentrated on the results given by Eq. 3. Further regression analysis indicated that the model expressed by Eq. 3 has nineteen outliers in five different steps (compounds 5, 22, 24, 28, 29, 30, 32, 34, 45, 52, 60, 67, 68, 70, 72, 73, 75, 76 and 79), the deletion of which give the following models with excellent statistics:

$$\log 1/C = 0.2269(\pm 0.1106)\log P + 1.2569(\pm 0.1739)\,I_{ZS} + 1.1072(\pm 0.2095)\,I_R + 0.6662\,(\pm 0.2096)\,I_X + 4.0364 \quad (4)$$

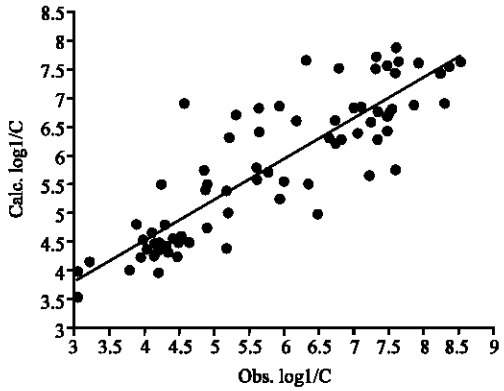

Fig. 2: Relationship obtained between observed and calculated log1/C from Eq. 3

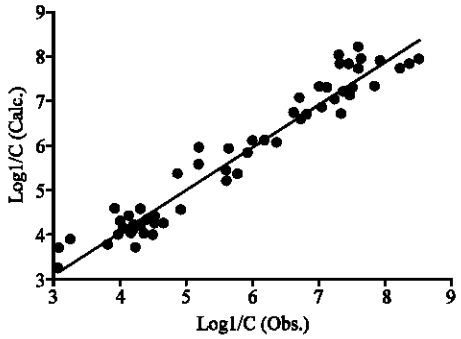

Fig. 3: Relationship obtained between observed and calculated log1/C from Eq. 8.

n = 74, Se = 0.6570, R = 0.9025, R^2_A = 0.8038, F = 75.757

$$\log 1/C = 0.2240(\pm 0.0939)\log P + 1.2700(\pm 0.1515)\, I_Z + 1.2027(\pm 0.1802)\, I_R + 0.7192\,(\pm 0.1852)\, I_X + 4.0316 \quad (5)$$

n = 69, Se = 0.5542, R = 0.9326, R^2_A = 0.8616, F = 106.863

$$\log 1/C = 0.2603(\pm 0.0823)\log P + 1.2883(\pm 0.1381)\, I_Z + 1.3628(\pm 0.1618)\, I_R + 0.5943\,(\pm 0.1717)\, I_X + 3.9875 \quad (6)$$

n = 66, Se = 0.4833, R = 0.9505, R^2_A = 0.8971, F = 142.709

$$\log 1/C = 0.2784(\pm 0.0673)\log P + 1.2539(\pm 0.1154)\, I_{ZS} + 1.5201(\pm 0.1350)\, I_R + 0.5040\,(\pm 0.1456)\, I_X + 3.9011 \quad (7)$$

n = 61, Se = 0.3931, R = 0.9697, R^2_A = 0.9404, F = 220.894

$$\log 1/C = 0.2820(\pm 0.0634)\log P + 1.2077(\pm 0.1033)\, I_Z + 1.5974(\pm 0.1391)\, I_R + 0.4475\,(\pm 0.1386)\, I_X + 3.9106 \quad (8)$$

n = 60, Se = 0.3703, R = 0.9736, R^2_A = 0.9440, F = 249.772

Table 5: Observed and calculated log1/C of TIBO derivatives used in present study

Comp. No.	log1/C(Obs.)	log1/C(Calc.)[a]	Residual	log1/C(Calc.)[b]	Residual
1	7.36	6.73	0.62	7.21	0.15
2	7.47	7.52	-0.05	7.85	-0.38
3	8.37	7.52	0.84	7.85	0.52
4	8.24	7.41	0.82	7.74	0.50
5	8.30*	6.88	1.41	7.37	0.93
6	7.47	6.69	0.77	7.16	0.31
7	7.02	6.82	0.19	7.30	-0.28
8	5.94	5.23	0.70	5.82	0.12
9	7.25	6.56	0.68	7.04	0.21
10	6.73	6.59	0.13	7.06	-0.33
11	5.20	4.98	0.21	5.56	-0.36
12	7.33	6.26	1.06	6.71	0.62
13	8.52	7.59	0.92	7.92	0.60
14	7.06	6.36	0.69	6.81	0.25
15	7.32	7.69	-0.37	8.03	-0.71
16	6.36	5.48	0.87	6.07	0.29
17	7.53	6.80	0.72	7.28	0.25
18	6.00	5.52	0.48	6.12	-0.12
19	7.87	6.85	1.01	7.34	0.53
20	4.48	4.22	0.26	4.01	0.47
21	3.07	3.95	-0.88	3.72	-0.65
22	5.18*	4.34	0.83	4.13	1.05
23	4.22	3.95	0.26	3.73	0.49
24	5.18*	4.34	0.83	4.13	1.05
25	3.80	4.00	-0.20	3.78	0.02
26	5.61	5.55	0.05	5.22	0.39
27	7.60	7.41	0.18	7.74	-0.14
28	5.23*	6.28	-1.05	6.73	-1.50
29	6.31*	7.61	-1.30	7.95	-1.64
30	6.50*	4.94	1.55	4.75	1.75
31	5.18	5.35	-0.17	5.95	-0.77
32	5.33*	6.68	-1.35	7.16	-1.83
33	7.60	7.85	-0.25	8.19	-0.59
34	5.97*	7.69	-1.72	8.03	-2.06
35	4.15	4.26	-0.11	4.04	0.11
36	4.33	4.41	-0.08	4.20	0.13
37	3.07	3.51	-0.44	3.27	-0.20
38	3.24	4.13	-0.89	3.91	-0.67
39	3.97	4.21	-0.24	4.00	-0.03
40	4.30	4.37	-0.07	4.15	0.15
41	4.05	4.36	-0.31	4.15	-0.10
42	4.36	4.30	0.05	4.08	0.28
43	4.24	4.45	-0.21	4.24	0.00
44	4.00	4.51	-0.51	4.30	-0.30
45	4.90*	5.41	-0.51	6.00	-1.10
46	4.21	4.43	-0.22	4.22	-0.01
47	4.54	4.58	-0.04	4.38	0.16
48	4.43	4.52	-0.09	4.32	0.11
49	3.91	4.79	-0.88	4.59	-0.68
50	4.15	4.45	-0.30	4.24	-0.09
51	7.34	7.49	-0.15	7.82	-0.48
52	6.80*	7.49	-0.69	7.82	-1.02
53	4.64	4.47	0.16	4.27	0.37
54	4.50	4.47	0.02	4.26	0.24
55	6.17	6.57	-0.40	6.11	0.06
56	5.66	6.38	-0.72	5.90	-0.24
57	4.13	4.62	-0.49	4.42	-0.29
58	4.90	4.74	0.15	4.54	0.36
59	4.32	4.76	-0.44	4.56	-0.24
60	4.92*	5.47	-0.55	6.07	-1.15
61	6.84	6.25	0.58	6.70	0.14

Table 5: Continued

Comp. No.	log1/C(Obs.)	log1/C(Calc.)[a]	Residual	log1/C(Calc.)[b]	Residual
62	6.80	6.25	0.54	6.70	0.10
63	5.61	5.76	-0.15	5.43	0.18
64	7.11	6.80	0.30	7.29	-0.18
65	7.92	7.58	0.33	7.91	0.01
66	7.64	7.58	0.05	7.92	-0.28
67	4.25*	5.49	-1.24	6.09	-1.84
68	5.65*	6.82	-1.17	7.30	-1.65
69	4.87	5.71	-0.84	5.38	-0.51
70	5.94*	6.83	-0.89	7.31	-1.37
71	6.64	6.27	0.36	6.72	-0.08
72	6.32*	7.60	-1.28	7.94	-1.62
73	4.59*	6.88	-2.29	7.37	-2.78
74	6.74	6.19	0.54	6.63	0.11
75	7.47*	6.40	1.06	5.92	1.55
76	7.22*	5.63	1.58	5.30	1.92
77	4.22	4.36	-0.14	4.15	0.07
78	5.78	5.69	0.08	5.37	0.41
79	7.59*	5.74	1.84	5.42	2.17

a: Calculated log1/C values from Eq. 3. b: Calculated log1/C values from Eq. 8, *: Data point not included in calculations from Eq. 8

Table 6: Modeling parameters calculated for the few TIBO derivatives

Comp. No.	TE	DpM	RMSg
25	28.39	2.569	0.177
39	37.08	4.650	0.013
41	30.95	4.670	0.055
43	23.39	4.670	0.034
50	422.98	4.762	0.007
55	26.21	1.316	0.006
65	33.36	1.808	0.024
71	28.18	3.793	0.029
77	24.33	4.694	0.031

TE = Total Energy, DpM = Dipole Moment, RMSg = Root Mean Square gradient

Table 7: Modeling parameters* calculated for the few TIBO derivatives

Comp. No.	NC2	NC6	NC9	ED2	ED6	ED9
25	0.5368	-0.2851	0.1510	3.463	5.285	3.849
39	0.5426	-0.2499	0.1473	3.457	5.250	3.853
41	0.5430	-0.2631	0.1473	3.457	5.263	3.853
43	0.5385	-0.2342	0.1416	3.461	5.234	3.858
50	0.5405	-0.2451	0.1462	3.459	5.245	3.874
55	0.2735	-0.2411	0.1839	3.726	5.241	3.816
65	0.2641	-0.2491	0.1846	3.736	5.249	3.815
71	0.5265	-0.2061	0.1713	3.473	5.205	3.829
77	0.5424	-0.2667	0.1470	3.458	5.267	3.853

NC2 = Net charge on 2nd Carbon atom, NC6 = Net charge on 6th Nitrogen atom, NC9 = Net charge on 9th Carbon atom, ED2 = Electron Density on 2nd C atom, ED6 = Electron Density on 6th N atom, ED9 = Electron Density on 9th C atom (All in a.u.)

Comparison of Eq. 4-8 shows that the model obtained for the set of 60 compounds gives the better statistics and most suitable for the prediction of inhibition activity of the compounds against reverse transcriptase-1. It is obvious that reduction in size of data set increases the regression value, but in present case significant lowering of Se and a large improvement in the F-statistics along with the improvement in the value of R^2_A from Eq. 3-8 justify the improvement in statistics and deletion of the compounds.

At this stage, it is worthy to comment on R^2_A values. We observed that as we pass from the model obtained for 79 compounds (Eq. 1-3) to model obtained for 60 compounds (Eq. 8) there is consistent increase in R^2_A, increasing from 0.7011 to 0.9440, as we pass from (Eq. 3-8). Such an increase in R^2_A values indicates that the deleted compounds have the unfair share in the modeling of

respective activity and also showing exceptional behavior from their parent series. The value of R^2_A will decrease if the deletion of the compounds does not reduce the unexplained variation in the model enough to off set the loss of degree of freedom (Lawtrakul *et al*., 1999; Tanaka, 1991, 1992, 1995). In the second phase of our study based on the category second containing 19 compounds. All these compounds (outliers) taken together resulted into a model according to the following equation:

$$\log 1/C = 22.2389(\pm 5.9919)\ \eta + 7.5472\ (\pm 3.7888)\ J\text{-}1.3837(\pm 0.6223)\ I_R\text{-}43.166 \quad (9)$$

n = 19, Se = 0.8207, R = 0.7422, R^2_A = 0.461, F = 6.131

Model presented in form of Eq. 9 expresses the domination of steric and structural features in comparison with hydrophobic parameter logP for the set of 19 out lairs.

Equation 9 also shows the domination of steric properties over the branching and size specific properties for modeling the activity log1/C for the set of 19 outliers. Equation also demonstrates the high favor of branching and steric property to the biological activity log1/C.

Equation also exhibits the unfavorable presence of acyclic structure at R position for these compounds, just apposite to the parent series.

We have also obtained quantum chemically derived parameters since some of the properties depend strongly on electronic features such as electrophilic regions of the compounds.

In order to carry out quantum computations, we have first carried out the molecular geometry optimizations (Loya *et al*., 1995) to find out the structural behavior of these compounds as a function of attached groups and their positions. The corresponding molecular modeling parameters are shown in Table 6 and 7.

Based on the above study and magnitude of residue from Eq. 8 we have selected compounds viz., 25, 39, 41, 43, 50, 55, 65, 71 and 77 to correlate their modeling parameters with the activities. This we have done to find out which TIBO derivative has the highest correlative and predictive potential for the same category. The molecular modeling is demonstrated in (Fig. 4-12), respectively for compounds 25, 39, 41, 43, 50, 55, 65, 71 and 77. The corresponding molecular modeling parameters are presented in Table 5. In order to resolve our problem of selecting out the TIBO derivative with the best quality and correlation potential; we have carried out further regression analysis using the molecular modeling parameters from Table 6 and 7.

From the modeling parameters significant univariate correlation shown by the net charge on atom 2 (C), electron density at atom 2 (C), electron density on 9th atom, but excellent result obtained from Net charge on 9th atom (C) and the models obtained are shown below.

$$\log 1/C = -9.9796\ (\pm\ 2.8803)\ NC2 + 9.7946 \quad (10)$$

n = 9, Se = 0.9703, R = -0.7948, F = 12.004

$$\log 1/C = 9.9775\ (\pm\ 2.8803)\ ED2\ -30.1140 \quad (11)$$

n = 9, Se = 0.9705, R = 0.7947, F = 11.999

$$\log 1/C = -65.0924\ (\pm\ 13.8025)\ ED9\ -255.2617 \quad (12)$$

n = 9, Se = 0.7823, R = -0.8723, F = 22.24

$$\log 1/C = 80.6404\ (\pm\ 12.3788)\ NC9\ -7.7073 \quad (13)$$

n = 9, Se = 0.6016, R = 0.9265, F = 42.437

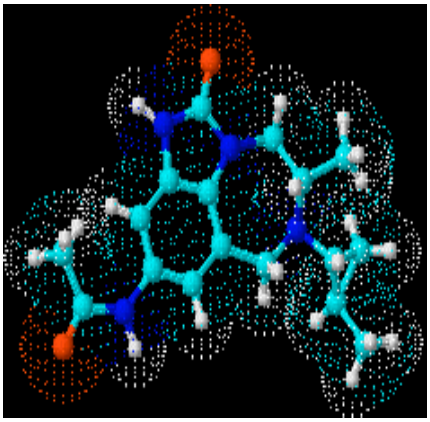

Fig. 4: Opt. structure of comp. 25

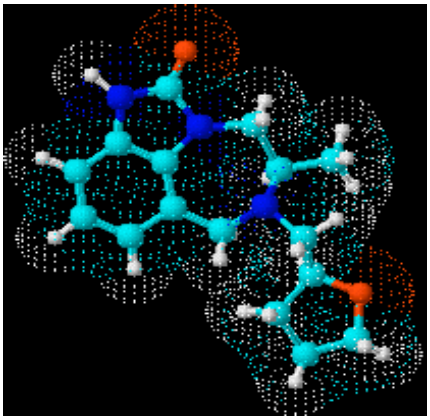

Fig. 5: Opt. structure of comp. 39.

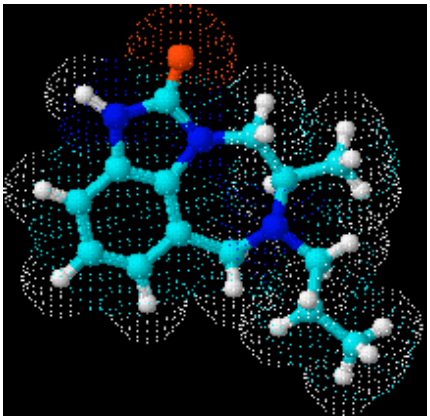

Fig. 6: Opt. structure of comp. 41

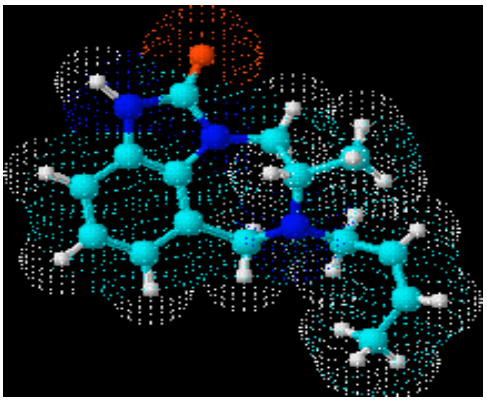

Fig. 7: Opt. structure of comp. 43

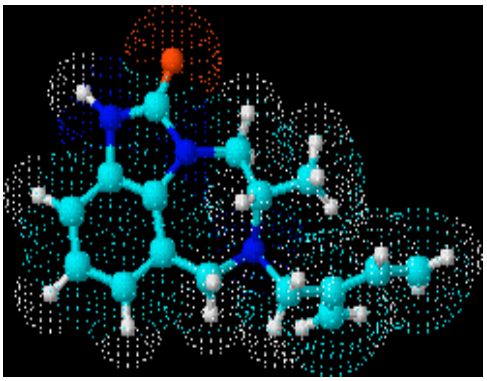

Fig. 8: Opt. structure of comp. 50

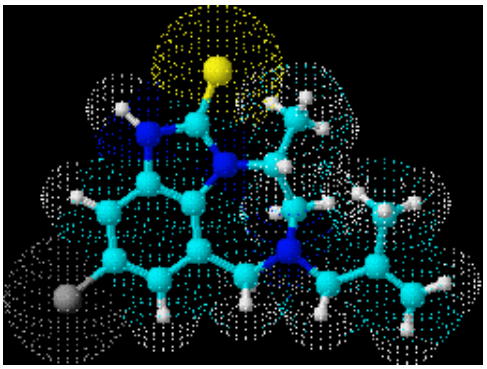

Fig. 9: Opt. structure of comp. 55

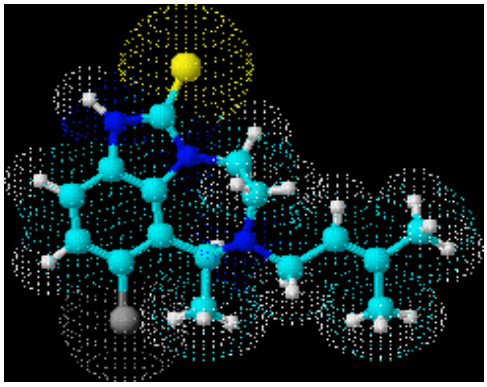

Fig. 10: Opt. structure of comp. 65

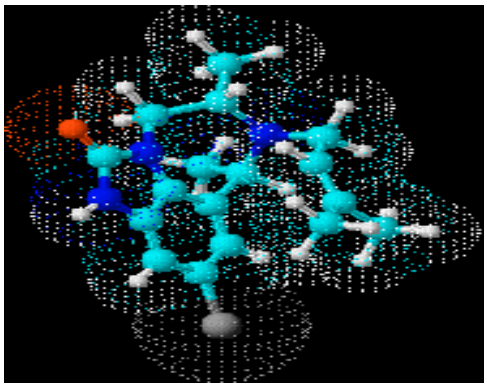

Fig. 11: Opt. structure of comp. 71

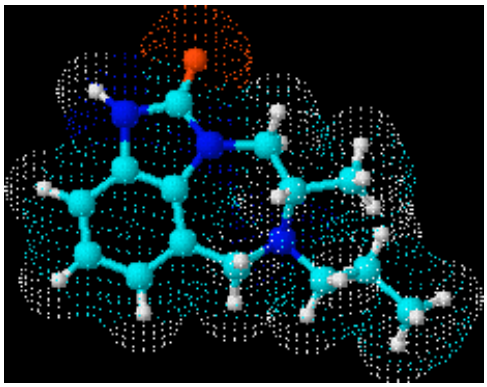

Fig. 12: Opt. structure of comp. 77

Equation 10 demonstrates that the compound having the carbon atom at 2nd position with higher net charge is unfavorable for the inhibition activity against RT's for the TIBO derivatives. At the same time opposite results are shown by the electron density at the same carbon atom in Eq. 11. Both the equations exhibit the significant role of the nature of the C atom at 2nd position in the inhibition of RT1 by TIBO derivatives.

The results obtained from Eq. 12 and 13 express the unfavorable presence of the electron density on carbon atom at 9th position and the presence of the high net charge, at the carbon atom on 9th position in positive manner for the inhibition of RT1 by TIBO derivatives. The comparison of all the four equations exhibit the domination and significant role of 9th atom (C) over the atom 2 (C) i.e., the presence of any substitution at C atom at 9th position containing high net charge favors the inhibition activity for RT1 by TIBO derivatives as compare to substitution containing low electron density at the same atom as well as the substitution with low net charge and high electron density at 2nd carbon atom.

CONCLUSION

From the result and discussion made above, we conclude that the hydrophobic parameters can be used successfully for modeling the inhibition activities of reverse transcriptase-1 by TIBO derivatives and that for the present set of TIBO derivatives the hydrophobic parameter logP is found to be the prominent one. The results also indicate that combination of unconventional, classical and hydrophobic parameters and molecular (3D) modeling can be used for the understanding the structural behavior and select the compound with potential activity.

REFERENCES

Balaban, A.T. and T.S. Balaban, 1991. New vertex invariants and topological indices of chemical graphs based on information on distances. J. Math. Chem., 8: 383-397.

Balaban, A.T., 1992. Using real numbers as vertex invariants for third-generation topological indexes. J. Chem. Inf. Comput. Sci., 32: 23-28.

Barre, S.F., J.C. Cherman, F. Rey, M.T. Nugeyre, S. Chamaret, J. Gruest, C. Daguet. C. Axler-Blin, F. Vezinet-Brun and C. Rouzioux, 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science, 220: 868-871.

Chaterjee, S. and A.S. Hadi, 2000. Price, B Regression Analysis by Examples. 3rd Edn., Wiley VCH: New York.

Debyser, Z., R. Pauwls, K. Andries and E. De Clercq, 1992. Specific HIV-1 reverse transcriptase inhibitors. J. Enzyme Inhibition, 6: 47-53.

De Clercq, E., 1992. HIV inhibitors targeted at the reverse transcriptase. AIDS Res. Hum. Retroviruses, 8: 119-134.

Fischl, M.A., D.D. Richman, D.M. Causey, M.H. Grieco, Y. Bryson, D.D. Mildvan, O.L. Laskin, J.E. Groopman, P.A. Volberding and R.T. Shooley, 1989. Prolonged zidovudine therapy in patients with AIDS and advanced AIDS-related complex. AZT Collaborative Working Group. N. Engl. J. Med., 317: 192-197.

Gallo, R.C., S.Z. Salahuddin, M. Popovic, G.M. Shearer, M. Kaplan, B.F. Haynes, T.J. Palker, R. Redfield, J. Oleske and B. Safai, 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science, 224: 500-503.

Garg, R., S.P. Gupta, H.G. Mekapati, S. Babu, A.K. Debnath and C. Hansch, 1999. Comparative quantitative structure-activity relationship studies on Anti-HIV drugs. Chem. Rev., 99: 3525-3601.

Lawtrakul, L., S. Hannongbua, A. Beyer and P. Wolschann, 1999. Molecular calculations on the conformation of the HIV-1 RT inhibitor of 1-(2-Hydroxyethoxy)-methyl)-6 (phenylthio)thymine (HEPT). Monatsh. Chem., 130: 1347-1363.

Loya, S., Y. Kashman and A. Hizi, 1992. The carotenoid halocynthiaxanthin: A novel inhibitor of the reverse transcriptases of human immunodeficiency viruses type 1 and type 2. Arch. Biochem. Biophys., 293: 208-212.

Loya, S., A. Rudi, R. Tal, Y. Kashman, Y. Loya and A. Hizi, 1994. 3,5,8-Trihydroxy-4-quinolone, a novel natural inhibitor of the reverse transcriptases of human immunodeficiency viruses type 1 and type 2. Arch. Biochem. Biophys., 309: 315-322.

Loya, S., M. Bakhanashvili, Y. Kashman and A. Hadi, 1995a. Mechanism of inhibition of HIV reverse transcriptase by toxiusol, a novel general inhibitor of retroviral and cellular DNA polymerases. Biochemistry, 34: 2260-2266.

Loya, S., M. Bakhanashvili, Y. Kashman and A. Hizi, 1995b. Peyssonols A and B, two novel inhibitors of the reverse transcriptases of human immunodeficiency virus types 1 and 2. Arch. Biochem. Biophy., 316: 789-786.

Mitsuya, H. and S. Broder, 1986. Inhibition of the *in vitro* infectivity and cytopathic effect of human T-lymphotrophic virus type III/lymphadenopathy-associated virus (HTLV-III/LAV) by 2',3'-dideoxynucleosides. Proc. Natl. Acad. Sci., USA., 83: 1911-1915.

Reardon, J.E. and W.H. Miller, 1990. Human immunodeficiency virus reverse transcriptase. Substrate and inhibitor kinetics with thymidine 5'-triphosphate and 3'-azido-3'-deoxythymidine 5'-triphosphate. J. Biol. Chem., 265: 20302-20307.

Richman, D.D., M.A. Fischl, M.H. Grieco, M.S. Gotlieb, P.A. Volberding, O.L. Laskin, J.M. Leedom, J.E. Grooman, D. Mildvan and M.S. Hirsch, 1987. The toxicity of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. N. Engl. J. Med., 317: 192-197.

Tanaka, H., M. Baba, H. Kayakawa, T. Sakamaki, T. Miyasaka, M. Ubasawa, H. Takashima, K. Sekiya, I. Nitta, S. Shigeta, R.T. Walker, J. Balzarini and E.S. De Clercq, 1991. A new class of HIV-1-specific 6-substituted acyclouridine derivatives: Synthesis and anti-HIV-1 activity of 5-or 6-substituted analogues of 1-((2-hydroxyethoxy)methyl)-6-(phenylthio)thymine (HEPT). J. Med. Chem., 34: 349-357.

Tanaka, H., H. Takashima, M. Ubasawa, K. Sekiya, I. Nitta, M. Baba, S. Shigeta, R.T. Walker, E. De Clercq and T. Miyasaka, 1992. Structure-activity relationships of 1-((2-hydroxyethoxy)methyl)-6-(phenylthio)thymine analogues: Effect of substitutions at the C-6 phenyl ring and at the C-5 position on anti-HIV-1 activity. J. Med. Chem., 35: 337-345.

Tanaka, H., H. Takashima, M. Ubasawa, K. Sekiya, N. Inouye, M. Baba, S. Shigeta, R.T. Walker, E. De Clercq and T. Miyasaka, 1995. Nonnucleoside HIV-1 reverse transcriptase inhibitors-3D-QSAR study on 6-(1-naphthylmethyl) substituted thymines. J. Med. Chem., 38: 2860-2865.

Wlodawer, A., 2002. Rational approach to AIDS drug design through structural biology. Annu. Rev. Med., 53: 595-614.