Asian Journal of Biochemistry

ISSN 1815-9923

Academic
Journals Inc.

www.academicjournals.com

Asian Journal of Biochemistry 2 (2): 84-100, 2007
ISSN 1815-9923

2D, 3D Modeling of Inhibition Activity of Reverse Transcriptase-1 by HEPT Derivatives

[1]A. Thakur, [2]B.K. Tiwari, [3]M. Thakur, [4]S. Thakur, [2]N.D. Pandey and [2]S.S. Narvi
[1]Department of Chemistry, Rishiraj institute of Technology, Indore, India,
Department of Chemistry, University of Cape Town, South Africa
[2]Department of Chemistry, M.N.N.I.T., Allahabad, India
[3]Department of Chemistry, Softvision institute of Bio-Technology, Indore, India
[4]Department of Physics, Shri Vaishnav Institute of Technology and Science, Indore, India

Abstract: The study describes SAR and QSAR of inhibition of Reverse Transcriptase-1 by HEPT derivatives using both classical and non-conventional physicochemical parameters along with hydrophobic parameter and indicator parameters. The set of HEPT derivatives Studied contains 85 compounds with different substitution at various positions. Application of multiple linear regression analysis indicated that combination of classical physicochemical parameters with indicator parameters yielded statistically significant model for modeling inhibitory activity (log1/C) against Reverse Transcriptase-1. Final selection of the potential HEPT derivative for the inhibition of Reverse Transcriptase-1 is made with the help of molecular modeling parameters.

Key words: QSAR, reverse transcriptase, physicochemical property

Introduction

The Acquired Immuno Deficiency Syndrome (AIDS) is one of the most rapid sprayed diseases, which is caused by infection of human immuno deficiency virus (HIV). During the studies of effective therapies used to inhibits the compassing of HIV, Reverse Transcriptase (RT) has been distinguish as promising target, because Reverse Transcriptase (RT) is not required for normal host cell replication (Clercq, 1995) and Conversion process of the single-stranded viral RNA genome into double-stranded proviral DNA ahead of its adherence into the host genomic DNA is performed by the reverse transcriptase (Young, 1993). However, a serious puzzle with the reverse transcriptase inhibitors specifically HEPT derivatives is the outburst of viral strains that have point mutations in the region encoding HIV-1 RT which check these drugs from inhibiting RT (Hannongbua *et al*., 1996a, b; Kireev *et al*., 1997).

Inspired by the behavior of HEPT derivatives in inhibition of RT-1 and in continuation to our earlier studies (Thakur, 2005; Balaban *et al*., 2005; Thakur *et al*., 2004a-e, 2005), our objective in present study is to made SAR and QSAR analysis of inhibition of reverse transcripase-1 by HEPT derivatives using three different sets of molecular and hydrophobic descriptors consisting of some non-conventional physicochemical parameters like approximate surface area (ASA), surface area grid (SAG) and Hydration Energy (HE), hydrophobic parameter as logP and some classical physicochemical properties like Molar Refractivity (MR), Molar Volume (MV), Parachor (Pc), refractive index (η), Surface Tension (ST) and density (d) in addition to indicator parameters. For QSAR modeling we have used maximum R^2 method and followed step-wise regression analysis (Chaterjee and Hadi, 2000). To model the most potent HEPT derivative we optimized the molecules using molecular mechanics method, applying MM^+ force field. The parent structure of the HEPT derivatives is presented as Fig. 1.

Corresponding Author: A. Thakur, 205, Sudama Nagar Anp. Sec. Indore 452009, India
Tel: 91-731-2792454 Fax: 91-731-2763618

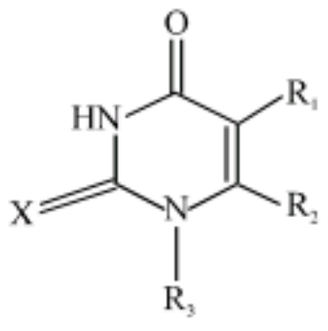

Fig. 1: Parent structure for HEPT derivatives

Materials and Methods

Activity

The inhibitory action against RT's of the HEPT derivatives were adopted from the literature (Hannongbua *et al*., 2001).

Physicochemical Parameters

The classical physicochemical parameters Molar volume(MV), Parachor(Pc), Molar Refractivity (MR), Refractive Index (η), Surface Tension (ST), Density (D) and Polarizability (α), for the set of HEPT derivatives were calculated from ACD Lab software [www.acdlabs.com]. While the non-conventional physicochemical parameters, ASA, SAG and hydration energy were calculated with the help of Hyperchem7 (Demo-version).

Indicator Parameters

These 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.

Molecular Modeling

Molecular optimization and modeling were performed applying MM^+ force field for this purpose and for the calculation of modeling parameters Hyperchem7 (Demo-version) [www.hyper.com] were used.

Statistical Analysis

Maximum R^2 method together with step-wise 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 (QSAR). Multiple linear regression is used to develop these models.

Results and Discussion

The set of 85 HEPT derivatives and their adopted inhibition values for HIV-1 RT expressed as log1/C are presented in Table 1. Non-conventional physicochemical parameters are recorded in Table 2 while the classical physicochemical parameters are shown in Table 3. Hydrophobic parameter logP along with indicator parameters are recorded in Table 4. The inter correlation of classical and non-conventional physicochemical parameters and hydrophobic parameter logP are presented in Table 5 in form of correlation matrix.

The inter correlatedness among molecular descriptors as well as with the activity shows that the mutual correlation exists between classical physicochemical parameters, while this is not so with the

Table 1: Sub-stituents and Biological activity (Observed)[] of HEPT derivatives used in present study

S. No.	X	R_1	R_2	R_3	log1/C
1	O	Me	SC_6H_{11}	$CH_2OCH_2CH_2OH$	5.085
2	O	Me	CH_2Ph	$CH_2OCH_2CH_2OH$	4.637
3	O	Me	C=CPh	$CH_2OCH_2CH_2OH$	4.853
4	O	I	SPh	$CH_2OCH_2CH_2OH$	5.443
5	O	SPh	SPh	$CH_2OCH_2CH_2OH$	4.677
6	O	$COCHMe_2$	SPh	$CH_2OCH_2CH_2OH$	4.920
7	O	COPh	SPh	$CH_2OCH_2CH_2OH$	4.885
8	O	CH_2Ph	SPh	$CH_2OCH_2CH_2OH$	4.637
9	O	$CH=CPh_2$	SPh	$CH_2OCH_2CH_2OH$	6.075
10	O	C=CMe	SPh	$CH_2OCH_2CH_2OH$	4.720
11	O	C=CPh	SPh	$CH_2OCH_2CH_2OH$	5.468
12	O	CH=CHPh	SPh	$CH_2OCH_2CH_2OH$	5.221
13	O	$CH=CH_2$	SPh	$CH_2OCH_2CH_2OH$	5.958
14	O	Me	SPh(2-Me)	$CH_2OCH_2CH_2OH$	4.148
15	O	Me	SPh(2-OMe)	$CH_2OCH_2CH_2OH$	4.720
16	O	Me	SPh(3-Me)	$CH_2OCH_2CH_2OH$	5.584
17	O	Me	SPh(3-Et)	$CH_2OCH_2CH_2OH$	5.568
18	O	Me	SPh(3-*t*-Bu)	$CH_2OCH_2CH_2OH$	4.920
19	O	Me	SPh(3-Cl)	$CH_2OCH_2CH_2OH$	4.885
20	O	Me	SPh(3-Br)	$CH_2OCH_2CH_2OH$	5.243
21	O	Me	SPh(3-I)	$CH_2OCH_2CH_2OH$	4.999
22	O	Me	SPh(3-NO_2)	$CH_2OCH_2CH_2OH$	4.468
23	O	Me	SPh(3-OH)	$CH_2OCH_2CH_2OH$	4.085
24	O	Me	SPh(3-OMe)	$CH_2OCH_2CH_2OH$	4.657
25	O	Me	SPh(3,5-Me_2)	$CH_2OCH_2CH_2OH$	6.584
26	O	Me	SPh(3,5-Cl_2)	$CH_2OCH_2CH_2OH$	5.885
27	S	Me	SPh(3,5-Me_2)	$CH_2OCH_2CH_2OH$	6.656
28	O	Me	SPh(3-COOMe)	$CH_2OCH_2CH_2OH$	5.102
29	O	Me	SPh(3-COMe)	$CH_2OCH_2CH_2OH$	5.136
30	O	Me	SPh(3-CN)	$CH_2OCH_2CH_2OH$	4.999
31	O	$CH_2CH=CH_2$	SPh	$CH_2OCH_2CH_2OH$	5.601
32	O	COOMe	SPh	$CH_2OCH_2CH_2OH$	5.180
33	O	CONHPh	SPh	$CH_2OCH_2CH_2OH$	4.744
34	S	Et	SPh	$CH_2OCH_2CH_2OH$	6.957
35	S	Et	SPh(3,5-Me_2)	$CH_2OCH_2CH_2OH$	8.106
36	S	i-Pr	SPh(3,5-Me_2)	$CH_2OCH_2CH_2OH$	8.300
37	S	Et	SPh(3,5-Cl_2)	$CH_2OCH_2CH_2OH$	7.365
38	O	Pr	SPh	$CH_2OCH_2CH_2OH$	5.468
39	O	Me	SPh	CH_2OMe	5.677
40	O	Me	SPh	CH_2OPr	5.443
41	O	Me	SPh	CH_2OBu	5.327
42	O	Me	SPh	CH_2OCH_2Ph	7.054
43	S	Et	SPh(3,5-Me_2)	CH_2OEt	8.355
44	S	Et	SPh (3,5- Cl_2)	CH_2OEt	7.885
45	S	Et	SPh	CH_2-*i*-Pr	6.656
46	S	Et	SPh	CH_2OCH_2-c-Hex	6.455
47	S	Et	SPh	CH_2OCH_2Ph	8.106
48	S	Et	SPh (3,5-Me_2)	CH_2OCH_2Ph	8.160
49	S	Et	SPh	CH_2OCH_2Ph(4-Me)	7.107
50	S	Et	SPh	CH_2OCH_2(4-Cl)	7.919
51	S	Et	SPh	$CH_2OCH_2CH_2Ph$	7.040
52	S	i-Pr	SPh	CH_2OEt	7.852
53	S	i-Pr	SPh	CH_2OCH_2Ph	8.179
54	S	c-Pr	SPh	CH_2OEt	7.021
55	O	Et	SPh	CH_2O-*i*-Pr	6.467
56	O	Et	SPh	CH_2O-c-Hex	5.397
57	O	Et	SPh	CH_2OCH_2-c-Hex	6.346
58	O	Et	SPh	$CH_2OCH_2CH_2Ph$	7.016
59	O	Me	SPh	Et	5.657
60	O	Me	SPh	Bu	5.920
61	O	Et	CH_2Ph	$CH_2OCH_2CH_2OH$	6.455

Table 1: Continued

S. No.	X	R_1	R_2	R_3	log1/C
62	O	Et	$CH_2Ph(3,5\text{-}Me)$	$CH_2OCH_2CH_2OH$	7.885
63	O	Et	CH_2Ph	CH_2OEt	7.386
64	O	i-Pr	CH_2Ph	$CH_2OCH_2CH_2OH$	7.199
65	O	i-Pr	$CH_2Ph(3,5\text{-}Me)$	$CH_2OCH_2CH_2OH$	8.567
66	O	i-Pr	CH_2Ph	CH_2OEt	8.375
67	O	i-Pr	$CH_2Ph(3,5\text{-di-}Me)$	CH_2OEt	9.220
68	O	i-Pr	CH_2Ph	Bu	7.374
69	O	Et	CH_2Ph	CH_2CH_2OMe	6.601
70	O	i-Pr	CH_2Ph	CH_2CH_2OMe	7.282
71	O	Me	SPh	$CH_2OCH_2CH_2OH$	5.154
72	O	C=CH	SPh	$CH_2OCH_2CH_2OH$	4.744
73	O	Me	SPh(3-F)	$CH_2OCH_2CH_2OH$	5.481
74	S	*i*-Pr	SPh	$CH_2OCH_2CH_2OH$	7.228
75	O	Me	SPh	$CH_2OCH_2CH_2OMe$	5.060
76	O	Me	SPh	CH_2OEt	6.480
77	S	Et	SPh	CH_2OEt	7.584
78	O	Et	SPh	CH_2OEt	7.720
79	O	Et	$SPh(3,5\text{-}Me_2)$	CH_2OEt	8.266
80	O	Et	$SPh(3,5\text{-}Me_2)$	CH_2OCH_2Ph	8.493
81	O	i-Pr	SPh	CH_2OEt	7.919
82	O	c-Pr	SPh	CH_2OEt	6.999
83	O	Et	$CH_2Ph(3,5\text{-}Me_2)$	$EtOCH_2$	7.199
84	O	Et	CH_2Ph	Bu	6.677
85	S	Me	SPh	$CH_2OCH_2CH_2OH$	6.008

Table 2: Non conventional physicochemical parameters used in present study

Comp. No.	HE	ASA	SAG
1	-10.34	425.87	490.27
2	-10.32	419.53	488.47
3	-10.70	502.28	506.27
4	-11.03	452.33	507.00
5	-13.03	518.02	608.77
6	-10.18	502.81	560.57
7	-12.47	484.74	593.59
8	-11.84	482.76	586.52
9	-12.32	499.93	651.15
10	-10.48	480.94	536.47
11	-11.68	440.15	556.48
12	-11.59	442.69	557.02
13	-11.68	431.49	503.34
14	-9.60	451.11	510.18
15	-10.71	463.58	525.41
16	-9.10	461.24	512.01
17	-8.60	488.23	537.02
18	-7.32	537.97	557.18
19	-9.61	453.10	508.76
20	-9.61	462.10	515.23
21	-9.70	475.88	524.23
22	-19.22	462.98	519.97
23	-16.28	437.75	501.12
24	-11.73	479.53	528.22
25	-8.09	493.65	533.16
26	-9.39	490.82	528.40
27	-9.65	502.95	548.46
28	-10.74	514.25	560.50
29	-10.38	525.46	557.40
30	-13.53	468.84	518.03
31	-11.49	461.27	538.06
32	-11.63	471.37	535.79
33	-14.48	464.41	578.83
34	-11.28	458.41	532.64
35	-8.97	522.67	574.85

Table 2: Continued

Comp. No.	HE	ASA	SAG
36	-8.86	522.02	590.56
37	-10.27	520.09	568.32
38	-9.57	484.93	542.41
39	-5.13	398.96	452.96
40	-4.18	439.57	498.53
41	-3.77	472.97	527.64
42	-5.39	395.51	525.19
43	-3.21	513.88	565.94
44	-4.65	512.49	561.69
45	-4.41	469.70	530.97
46	-4.15	511.77	611.98
47	-6.67	507.11	607.38
48	-4.53	585.82	662.97
49	-5.69	563.67	639.65
50	-6.54	556.76	635.45
51	-5.38	551.30	555.08
52	-4.69	472.69	532.64
53	-6.24	545.93	634.84
54	-5.17	420.55	527.87
55	-3.81	451.89	515.20
56	-4.31	306.92	519.14
57	-2.99	502.14	589.75
58	-5.07	483.85	604.41
59	-4.63	396.93	456.79
60	-3.80	431.67	488.03
61	-9.92	431.06	508.86
62	-8.57	532.03	575.37
63	-3.37	421.76	495.94
64	-9.32	441.19	522.44
65	-7.31	510.39	568.84
66	-2.86	424.47	505.76
67	-1.14	493.02	551.59
68	-2.24	463.48	530.17
69	-5.15	440.85	508.12
70	-4.80	457.33	520.21
71	-10.31	411.41	485.90
72	-10.93	445.48	497.27
73	-10.38	446.13	500.65
74	-10.92	484.56	550.65
75	-6.02	468.88	517.17
76	-4.62	417.14	478.09
77	-5.22	446.20	521.14
78	-4.25	438.10	502.19
79	-2.22	506.99	549.66
80	-4.24	595.25	653.91
81	-3.83	469.44	524.79
82	-4.31	411.33	512.79
83	-2.40	564.02	596.98
84	-2.62	446.77	518.89
85	-11.86	430.01	509.17

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

Table 3: Classical physicochemical properties* of HEPT derivatives used in present study

Comp. No.	MR	MV	Pc	η	ST	D	α
1	83.24	252.9	683.7	1.572	53.3	1.24	33.00
2	76.33	232.0	617.1	1.571	50.0	1.25	30.26
3	79.52	223.3	634.5	1.630	65.1	1.34	31.52
4	88.56	225.0	665.1	1.716	76.2	1.86	35.10
5	108.36	278.1	815.7	1.707	73.9	1.44	42.95
6	94.24	267.6	752.8	1.622	62.6	1.36	37.36
7	105.09	277.3	808.5	1.682	72.2	1.43	41.66

Table 3: Continued

Comp. No.	MR	MV	Pc	η	ST	D	α
8	104.97	280.4	803.3	1.671	67.3	1.37	41.61
9	134.00	348.3	1001.1	1.695	68.2	1.35	53.12
10	87.53	237.2	686.9	1.659	70.3	1.40	34.70
11	107.62	279.9	820.7	1.695	73.9	1.40	42.66
12	109.51	289.9	831.0	1.679	67.4	1.36	43.41
13	84.66	232.9	658.7	1.646	63.8	1.37	33.56
14	84.88	237.2	667.7	1.634	62.7	1.35	33.65
15	86.62	243.3	688.1	1.630	63.9	1.39	34.34
16	80.26	221.4	629.5	1.644	65.2	1.39	31.81
17	89.51	253.4	707.8	1.624	60.8	1.32	35.48
18	98.78	286.2	784.2	1.606	56.3	1.27	39.16
19	85.09	232.3	666.6	1.653	67.7	1.47	33.73
20	87.98	234.2	680.5	1.674	71.2	1.65	34.88
21	93.18	240.8	703.4	1.700	72.7	1.80	36.94
22	86.29	232.7	686.5	1.663	75.7	1.51	34.21
23	81.79	218.4	644.7	1.672	75.9	1.48	32.42
24	86.62	243.3	688.1	1.630	63.9	1.39	34.34
25	89.51	253.4	707.8	1.624	60.8	1.32	35.48
26	89.91	243.2	703.7	1.661	70.0	1.55	35.64
27	96.91	251.4	745.2	1.697	77.2	1.56	38.42
28	91.37	256.4	733.3	1.631	66.9	1.42	36.22
29	89.63	250.3	713.0	1.634	65.7	1.39	35.53
30	84.83	230.8	677.1	1.656	73.9	1.44	33.62
31	89.29	249.2	698.8	1.635	61.8	1.34	35.39
32	86.75	240.6	695.1	1.640	69.5	1.46	34.39
33	108.74	284.8	836.7	1.689	74.5	1.45	43.11
34	91.89	245.8	711.0	1.670	69.9	1.67	36.42
35	101.13	277.3	787.6	1.649	65.0	1.32	40.09
36	105.74	294.1	825.6	1.638	62.0	1.29	41.92
37	101.54	267.6	785.3	1.683	74.1	1.52	40.25
38	89.52	253.9	709.6	1.622	61.0	1.32	35.49
39	74.09	207.8	572.4	1.631	57.5	1.33	29.37
40	83.35	240.3	652.5	1.610	54.3	1.27	33.04
41	87.98	256.5	692.6	1.601	53.1	1.24	34.88
42	98.81	266.9	746.2	1.662	61.0	1.32	39.17
43	99.60	280.0	770.5	1.629	57.3	1.25	39.48
44	100.01	270.2	768.3	1.661	65.3	1.44	39.64
45	93.22	259.2	711.7	1.638	56.8	1.23	36.95
46	111.44	311.2	862.9	1.634	59.0	1.25	44.18
47	110.43	291.3	827.8	1.682	65.2	1.31	43.78
48	119.68	322.7	904.3	1.663	61.6	1.27	47.44
49	115.06	307.0	866.1	1.672	63.3	1.29	45.61
50	115.26	302.1	864.9	1.688	67.1	1.38	45.69
51	115.07	307.4	867.9	1.671	63.4	1.29	45.61
52	94.96	265.3	732.0	1.634	57.9	1.26	37.64
53	115.04	308.0	865.8	1.669	62.3	1.29	45.60
54	92.92	246.5	702.6	1.677	65.9	1.35	36.83
55	87.96	257.1	690.5	1.599	52.0	1.24	34.87
56	99.81	286.9	781.3	1.612	54.9	1.25	39.57
57	104.45	303.1	821.4	1.605	53.9	1.23	41.40
58	108.07	299.3	826.4	1.641	58.1	1.27	42.84
59	72.35	201.6	552.0	1.636	56.1	1.30	28.68
60	81.61	234.2	632.2	1.613	53.0	1.23	32.35
61	80.95	249.6	657.2	1.561	48.0	1.21	32.09
62	90.60	282.1	733.7	1.555	45.7	1.17	35.91
63	79.41	252.1	640.2	1.542	41.5	1.14	31.48
64	85.48	264.1	695.2	1.560	47.9	1.20	33.89
65	95.13	296.7	771.8	1.554	45.7	1.16	37.71
66	83.95	266.6	678.2	1.542	41.8	1.13	33.28
67	93.60	299.2	754.8	1.538	40.4	1.10	37.10
68	86.84	276.8	697.9	1.539	40.4	1.08	34.42

Table 3: Continued

Comp. No.	MR	MV	Pc	η	ST	D	α
69	79.41	252.1	640.2	1.542	41.5	1.14	31.48
70	79.41	252.1	640.2	1.542	41.5	1.14	31.48
71	80.26	221.4	629.5	1.644	65.2	1.39	31.81
72	82.70	222.6	649.0	1.665	72.2	1.43	32.78
73	80.37	226.0	636.8	1.629	63.0	1.44	31.86
74	96.49	262.6	749.0	1.655	66.1	1.34	38.25
75	85.10	246.4	672.9	1.606	55.5	1.30	33.73
76	78.72	224.0	612.4	1.620	55.8	1.30	31.21
77	90.35	248.4	694.0	1.647	60.8	1.29	35.81
78	83.35	240.3	652.5	1.610	54.3	1.27	33.04
79	92.60	271.8	729.1	1.596	51.7	1.23	36.71
80	112.68	314.5	862.8	1.635	56.6	1.26	44.67
81	87.96	257.1	690.5	1.599	52.0	1.24	34.87
82	85.92	238.3	661.1	1.640	59.1	1.33	34.06
83	89.06	284.6	716.7	1.538	40.1	1.11	35.30
84	82.30	262.2	659.9	1.540	40.0	1.09	32.62
85	87.26	229.6	670.9	1.684	72.9	1.41	34.59

*MR = Molar Refractivity, MV = Molar Volume, Pc = Parachor, η = Index of refraction, ST = Surface Tension, D = Density, α = Polarizability

Table 4: logP values and indicator parameters for HEPT derivatives used in present study

Comp. No.	logP	I_{TC}	I_{AC}	I_6	I_{SP}	I_{OH}	I_2
1	1.014	0	0	1	0	1	0
2	0.087	0	0	0	0	1	0
3	-0.025	0	0	0	0	1	0
4	-0.137	0	0	1	0	1	0
5	2.397	0	1	1	0	1	0
6	-0.058	0	0	1	0	1	0
7	1.361	0	1	1	0	1	0
8	1.688	0	1	1	0	1	0
9	3.782	0	1	1	0	1	0
10	0.255	0	0	1	0	1	0
11	1.498	0	1	1	0	1	0
12	1.962	0	1	1	0	1	0
13	-0.001	0	0	1	0	1	0
14	0.592	0	0	1	1	1	0
15	-0.041	0	0	1	1	1	0
16	0.592	0	0	1	1	1	0
17	1.111	0	0	1	1	1	0
18	2.206	0	0	1	1	1	0
19	0.820	0	0	1	1	1	0
20	1.082	0	0	1	1	1	0
21	1.467	0	0	1	1	1	0
22	-0.101	0	0	1	1	1	0
23	-0.638	0	0	1	1	1	0
24	-0.019	0	0	1	1	1	0
25	1.011	0	0	1	1	1	0
26	1.467	0	0	1	1	1	0
27	1.526	0	0	1	1	1	0
28	0.041	0	0	1	1	1	0
29	-0.285	0	0	1	1	1	0
30	-0.469	0	0	1	1	1	0
31	0.478	0	0	1	0	1	0
32	-0.752	0	0	1	0	1	0
33	0.400	0	1	1	0	1	0
34	1.095	0	0	1	0	1	0
35	1.933	0	0	1	1	1	0
36	2.523	0	0	1	1	1	0
37	2.389	0	0	1	1	1	0
38	1.116	0	0	1	0	1	0
39	1.060	0	0	1	0	0	0

Table 4: Continued

Comp. No.	logP	I_{TC}	I_{AC}	I_6	I_{SP}	I_{OH}	I_2
40	2.081	0	0	1	0	0	0
41	2.617	0	0	1	0	0	0
42	2.670	1	1	1	0	0	1
43	3.305	0	0	1	1	0	0
44	3.761	0	0	1	1	0	0
45	4.149	0	0	1	0	0	0
46	4.616	1	1	1	0	0	1
47	3.592	1	1	1	0	0	1
48	4.430	1	1	1	1	0	1
49	4.011	1	1	1	0	0	1
50	4.239	1	1	1	0	0	1
51	4.079	1	1	1	0	0	1
52	3.057	0	1	1	0	0	0
53	4.182	1	1	1	0	0	1
54	2.601	0	1	1	0	0	0
55	2.298	0	0	1	0	0	0
56	3.450	1	1	1	0	0	1
57	4.101	1	1	1	0	0	1
58	3.564	1	1	1	0	0	1
59	2.230	0	0	1	0	0	0
60	3.302	0	0	1	0	0	0
61	0.494	0	0	0	0	1	0
62	1.332	0	0	0	1	1	0
63	1.866	0	0	0	0	0	0
64	1.084	0	0	0	0	1	0
65	1.992	0	0	0	1	1	0
66	2.456	0	0	0	0	0	0
67	3.294	0	0	0	1	0	0
68	4.213	0	0	0	0	0	0
69	1.777	0	0	0	0	0	0
70	2.367	0	0	0	0	0	0
71	0.173	0	0	1	0	1	0
72	-0.466	0	0	1	0	1	0
73	0.421	0	0	1	1	1	0
74	1.685	0	0	1	0	1	0
75	0.977	0	0	1	0	0	0
76	1.545	0	0	1	0	0	0
77	2.467	0	0	1	0	0	0
78	1.952	0	0	1	0	0	0
79	2.790	0	0	1	1	0	0
80	3.915	1	0	1	1	0	1
81	2.542	0	0	1	0	0	0
82	2.086	0	1	1	0	0	0
83	2.615	0	0	0	1	0	0
84	3.623	0	0	0	0	0	0
85	0.688	0	0	1	0	1	0

* logP = Partition coefficient (Octenol/Water), I_{TC} = 1 if Cyclic structure at terminal of the chain at R_3 position, 0 otherwise, I_{AC} = 1 if Cyclic structure at alternate atoms, 0 otherwise, I_6 = 1 if S atom attached at R_2 position, 0 otherwise, I_{SP} = 1 if substituted Phenyl ring present at the atom attached on R_2, 0 otherwise, I_{OH} = 1 if OH present in the chain at R_3 position, 0 otherwise, I_2 = 1 if S atom present at X position, 0 otherwise

Table 5: Correlation between various parameters and biological activity in form of correlation matrix

	log1/C	MR	MV	Pc	η	ST	D	α
log1/C	1.0000							
MR	0.29983	1.00000						
MV	0.50189	0.88289	1.00000					
Pc	0.31517	0.98748	0.92660	1.00000				
RI	-0.28950	0.47519	0.00873	0.36867	1.00000			
ST	-0.46866	0.29289	-0.16131	0.21835	0.93569	1.00000		
D	-0.46185	0.05045	-0.34406	-0.02346	0.77589	0.85515	1.00000	
α	0.29975	10.0000	0.88291	0.98750	0.47514	0.29286	0.05043	1.00000

Table 5: Continued

	log1/C	MR	MV	Pc	η	ST	D	α
logP	0.66744	0.58354	0.70953	0.56587	-0.09140	-0.37555	-0.46445	0.58348
HE	0.05049	0.01621	0.04564	0.01026	-0.04797	-0.08599	-0.04726	0.01614
ASA	0.33555	0.59193	0.59990	0.62599	0.14857	0.10564	-0.01365	0.59184
SAG	0.36581	0.90168	0.87911	0.92164	0.27842	0.14528	-0.06809	0.90166
I_2	0.49856	0.46098	0.32866	0.41709	0.36714	0.22933	0.06128	0.46087
I_{AC}	0.30969	0.60325	0.58378	0.58209	0.18620	-0.00285	-0.14792	0.60328
I_{TC}	0.09180	0.72582	0.59569	0.68089	0.41808	0.21552	-0.02868	0.72580
I_6	-0.28742	0.31570	-0.05320	0.23411	0.76221	0.70808	0.54071	0.31574
I_{SP}	0.04447	-0.00140	0.00343	0.04396	0.02314	0.15302	0.21636	-0.00139
I_{OH}	-0.56428	-0.08674	-0.26102	-0.05583	0.31717	0.55735	0.55445	-0.08668

Table 5: Continued

	logP	HE	ASA	SAG	I_2	I_{AC}	I_{TC}	I_6
logP	1.0000							
HE	0.12315	1.00000						
ASA	0.32393	-0.00137	1.00000					
SAG	0.53152	0.01314	0.79073	1.00000				
I_2	0.46387	-0.09503	0.37106	0.39432	1.00000			
I_{AC}	0.59050	0.06969	0.30493	0.56519	0.33266	1.00000		
I_{TC}	0.48706	0.08984	0.11875	0.56524	0.26100	0.62946	1.00000	
I_6	-0.03774	-0.07451	0.05456	0.13323	0.24632	0.18004	0.25436	1.00000
I_{SP}	-0.16374	-0.04957	0.41355	0.14258	-0.00341	-0.15804	-0.36598	0.06247
I_{OH}	-0.73965	-0.15894	0.01808	-0.07891	-0.22642	-0.45091	-0.25301	0.11108

Table 5: Continued

	I_{SP}	I_{OH}
I_{SP}	1.00000	
I_{OH}	0.31748	1.00000

other non-conventional physicochemical parameters used. Furthermore, data presented in Table 5 shows that none of the molecular descriptors, including non-conventional physicochemical parameters, hydrophobic parameters and classical physicochemical parameters correlate well with the activity (log1/C). From this we conclude that these descriptors can be combined to yield statistically significant multi-parametric model for modeling the activity. Initial regression analysis indicated that out of the 12 molecular descriptors Surface Area Grid (SAG), logP, Molar Volume (MV), Parachor (Pc) and index of refraction (η) in combination of indicator parameters plays the dominating role in modeling the activity.

In the case of non-conventional physicochemical parameters, from the perusal of Table 5 non-of the non-conventional physicochemical parameter shows the significant univariate correlation. For the improvement in the modeling potential we test bi, tri and tetra-parametric combinations of non-conventional physicochemical parameters and indicator parameters. Best model obtained from the penta-parametric combination of SAG, indicator parameters I_2, I_6, I_{SP} and I_{OH}. Model obtained from these parameters is as below:

$$\begin{aligned}\log 1/C = {} & 0.0055(\pm 0.0021)\,SAG + 1.2431(\pm 0.2303)\,I_2 - 1.2962(\pm 0.2405)\,I_6 + \\ & 0.5231\,(\pm 0.1906)\,I_{SP} - 1.2389(\pm 0.1884)\,I_{OH} + 4.5839 \qquad (1)\\ n = 85,\ Se = 0.7839,\ R = 0.8109,\ & R^2_A = 0.6359,\ F = 30.340\end{aligned}$$

The statistics obtained from model demonstrate the role of volumetric parameters in the modeling of activity log1/C. Positive coefficient of indicator parameters I_2 and I_{SP} in Eq. 1 exhibits the enhancement in the activity with the substitution at 2nd position and presence of substituted phenyl ring. While the negative coefficient of indicator parameters I_6 and I_{OH} shows the inverse relationship between biological activity log1/C and presence of substitution at 6th position and OH substitution, respectively.

Similarly in case of hydrophobic parameter the best result is obtained from the penta parametric combination of logP and 4 indicator parameters I_2, I_6, I_{SP} and I_{OH}. The model obtained from above variable is shown below:

$$\log 1/C = 0.2748(\pm 0.0991)\log P + 1.1612(\pm 0.2392)\, I_2 - 1.2247(\pm 0.2396)\, I_6 + 0.5572\,(\pm 0.1879) I_{SP} - 0.7246(\pm 0.2695) I_{OH} + 6.74 \quad (2)$$
$$n = 85,\ Se = 0.7797,\ R = 0.8131,\ R^2_A = 0.698,\ F = 30.835$$

Statistics generated by Eq. 2 express the enhancement in the activity log1/C with the increase in hydrophobicity of compounds i.e, increase in lipophilicity in these compounds enhance the inhibitory action. Equation also exhibits the role of indicator parameters similar to the previous Eq. 1.

In case of classical physicochemical parameters only Molar Volume (MV) shows the significant univariate correlation (r = 0.50). In case of bi, tri and tetra-parametric correlation, few significant results are shown by the different combination of MV, Pc, ST, η and MR along with the indicator parameters. The best result is obtained from the combination of Molar Volume (MV), Parachor (Pc) and index of refraction (η) along with indicator parameters I_2 and I_{SP} and the model obtained is as below:

$$\log 1/C = 0.1618(\pm 0.0253)MV - 0.0557(\pm 0.0095)Pc + 22.4781(\pm 6.53)\,\eta + 1.5102\,(\pm 0.2127)\, I_2 + 0.4604(\pm 0.1692) I_{SP} - 32.6826 \quad (3)$$
$$n = 85,\ Se = 0.7089,\ R = 0.8485,\ R^2_A = 0.7022,\ F = 40.619$$

Equation 3 expresses the domination of volumetric parameters in modeling of inhibitory activity of NNRT's against the reverse transcriptase1. The equation also demonstrates the role of steric parameters in modeling the activity. Equation 3 again shows the similar behavior of indicator parameters I_2 and I_{SP} as Eq. 1 and 2.

Comparison of all three equations demonstrates that the highest Regression (R) value is obtained from the Eq. 3 this exhibits significant role of classical physicochemical parameters in modeling log1/C. This also demonstrates the dominating role of substitution at 2nd position and presence of substituted phenyl ring on the parent moiety. Comparison also expresses the domination of volumetric parameters over the hydrophobic and steric parameters in modeling the inhibitory activity of the compounds against RT-1.

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 ten outliers in three different steps (compounds 2, 14, 18, 45, 46, 78, 79, 81, 82 and 83), the deletion of which give the following models with excellent statistics:

$$\log 1/C = 0.1507(\pm 0.0217)MV - 0.0509(\pm 0.0082)Pc + 19.4523(\pm 5.6290)\,\eta + 1.5673\,(\pm 0.1850)\, I_2 + 0.5787(\pm 0.1486) I_{SP} - 28.3634 \quad (4)$$
$$n = 80,\ Se = 0.6046,\ R = 0.8904,\ R^2_A = 0.7788,\ F = 56.630$$

$$\log 1/C = 0.1563(\pm 0.0203)MV - 0.0530(\pm 0.0076)Pc + 19.8637(\pm 5.1838)\,\eta + 1.6766\,(\pm 0.1709)\, I_2 + 0.5976(\pm 0.1396) I_{SP} - 29.0375 \quad (5)$$
$$n = 76,\ Se = 0.5435,\ R = 0.9135,\ R^2_A = 0.8227,\ F = 70.607$$

$$\log 1/C = 0.1508(\pm 0.0200)MV - 0.0511(\pm 0.0075)Pc + 18.1677(\pm 5.1166)\,\eta + 1.6970\,(\pm 0.1669)\, I_2 + 0.5653(\pm 0.1369) I_{SP} - 26.1750 \quad (6)$$
$$n = 75,\ Se = 0.5301,\ R = 0.9174,\ R^2_A = 0.8301,\ F = 73.317$$

Comparison of Eq. 4-6 shows that the model obtained for the set of 75 compounds gives the better statistics and most suitable for the prediction of inhibitory 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. 3-6 justify the improvement in statistics and deletion of the compounds. The suitability of Eq. 6 for the prediction of log1/C expressed graphically in Fig. 2. Also, the observed and calculated log1/C along with residual values are presented in Table 6

At this stage, it is worthy to comment on R^2_A values. We observed that as we passes from the model obtained for 85 compounds (Eq. 1-3) to model obtained for 75 compounds (Eq. 6) there is

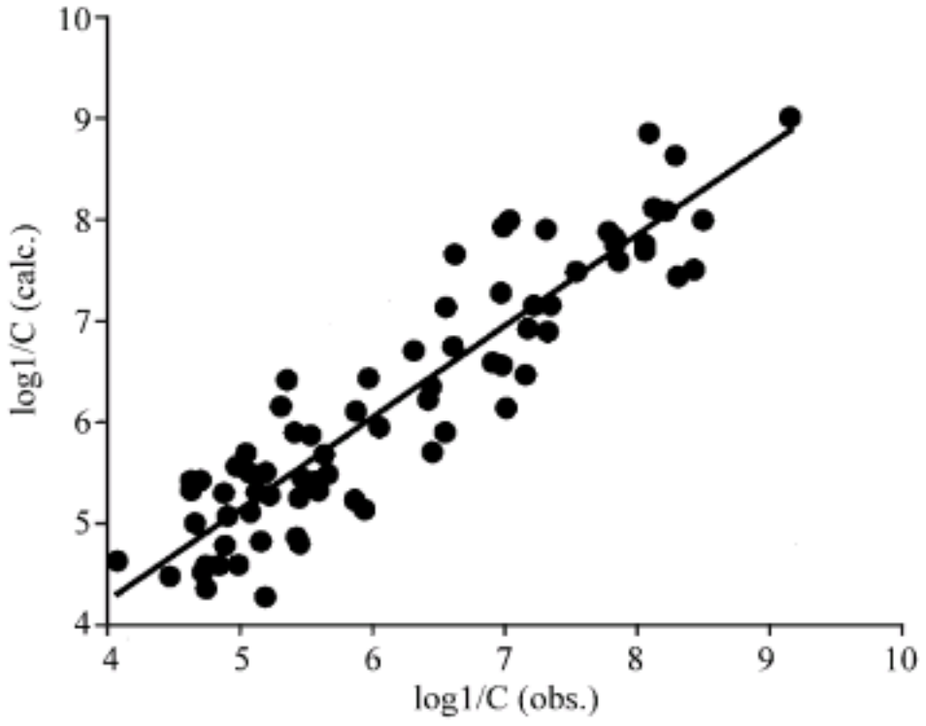

Fig. 2: Graph obtained between Obs. and Calc. log1/C values from Eq. 6

Table 6: Observed and calculated biological activity (log1/C) using Eq. 6 for HEPT derivatives

Comp. No.	log1/C (Obs.)	log1/C (Calc.)	Residual
1	5.085	5.585	-0.500
1	4.637*	5.818	-1.181
3	4.853	4.689	0.164
4	5.443	4.944	0.499
5	4.677	5.090	-0.415
6	4.920	5.179	-0.259
7	4.885	4.885	0.000
8	4.637	5.419	-0.782
9	6.075	5.987	0.088
10	4.720	4.634	0.086
11	5.468	4.890	0.578
12	5.221	5.581	-0.360
13	5.958	5.191	0.767
14	4.148*	5.727	-1.579
15	4.720	5.531	-0.811
16	5.584	5.478	0.106
17	5.568	5.939	-0.371
18	4.920*	6.654	-1.734
19	4.885	5.389	-0.504
10	5.243	5.347	-0.104
11	4.999	5.644	-0.645
12	4.468	4.614	-0.146
13	4.085	4.757	-0.672
14	4.657	5.531	-0.874
15	6.584	5.939	0.645
16	5.885	5.282	0.603
17	6.656	6.749	-0.093
18	5.102	5.215	-0.113
19	5.136	5.387	-0.251

Table 6: Continued

Comp. No.	log1/C (Obs.)	log1/C (Calc.)	Residual
20	5.243	5.347	-0.104
21	4.999	5.644	-0.645
22	4.468	4.614	-0.146
23	4.085	4.757	-0.672
24	4.657	5.531	-0.874
25	6.584	5.939	0.645
26	5.885	5.282	0.603
27	6.656	6.749	-0.093
28	5.102	5.215	-0.113
29	5.136	5.387	-0.251
30	4.999	4.681	0.318
31	5.601	5.400	0.201
32	5.180	4.383	0.797
33	4.744	4.703	0.041
34	6.957	6.597	0.360
35	8.106	7.616	0.490
36	8.300	8.008	0.292
37	7.365	6.889	0.476
38	5.468	5.321	0.147
39	5.677	5.543	0.134
40	5.443	5.969	-0.526
41	5.327	6.200	-0.873
42	7.054	6.137	0.917
43	8.355	8.534	-0.179
44	7.885	7.750	0.135
45	6.656*	8.000	-1.344
46	6.455*	8.043	-1.588
47	8.106	7.707	0.398
48	8.160	8.754	-0.594
49	7.107	7.936	-0.829
50	7.919	7.549	0.370
51	7.040	7.886	-0.846
52	7.852	7.810	0.042
53	8.179	8.048	0.131
54	7.021	7.259	-0.238
55	6.467	6.361	0.106
56	5.397	6.451	-1.054
57	6.346	6.718	-0.372
58	7.016	6.544	0.472
59	5.657	5.741	-0.084
60	5.920	6.141	-0.221
61	6.455	6.242	0.213
62	7.885	7.690	0.195
63	7.386	7.142	0.244
64	7.199	6.468	0.731
65	8.567	7.926	0.641
66	8.375	7.387	0.988
67	9.220	8.881	0.339
68	7.374	7.864	-0.490
69	6.601	7.142	-0.541
70	7.282	7.142	0.140
71	5.154	4.912	0.242
72	4.744	4.478	0.266
73	5.481	5.526	-0.045
74	7.228	6.916	0.312
75	5.060	5.774	-0.714
76	6.480	5.742	0.738
77	7.584	7.439	0.144
78	7.720*	5.969	1.750
79	8.266*	7.116	1.150
80	8.493	7.432	1.061

Table 6: Continued

Comp. No.	log1/C (Obs.)	log1/C (Calc.)	Residual
81	7.919*	6.361	1.558
82	6.999*	5.773	1.226
83	7.199*	8.626	-1.427
84	6.677	7.622	-0.945
85	6.008	6.457	-0.449

*Data point not included in calculation

consistent increase in the value of R^2_A. The values increasing from 0.7022 to 0.8301, as we passes from Eq. 3-6. Such an increase in R^2_A values indicate 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 (Agrawal *et al*., 2001, Khadikar *et al*., 2002a,b; Thakur *et al*., 2003, 2004).

Based on the magnitude of residue from Eq. 6 we have selected compounds 7, 9, 10, 27, 33, 52, 59 and the compound 73 for further molecular modeling. This we have done to find out which HEPT derivative has the highest correlative and predictive potential. We have, therefore, attempted molecular modeling, using Hyperchem software[www.hyper.com]. The molecular modeling is demonstrated in Fig. 3-10, respectively for compounds 7, 9, 10, 27, 33, 52, 59 and 73. The corresponding molecular modeling parameters are presented in Table 7. In order to resolve our problem of selecting out the HEPT derivative with the best quality and desired potential; we have carried out further regression analysis using the molecular modeling parameters from Table 7.

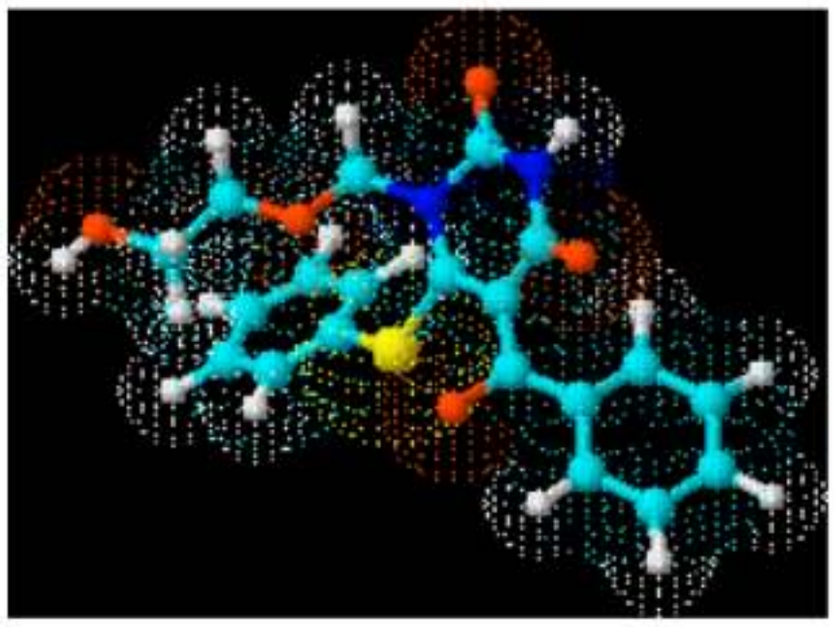

Fig. 3: Optimized structure of Comp. No. 7

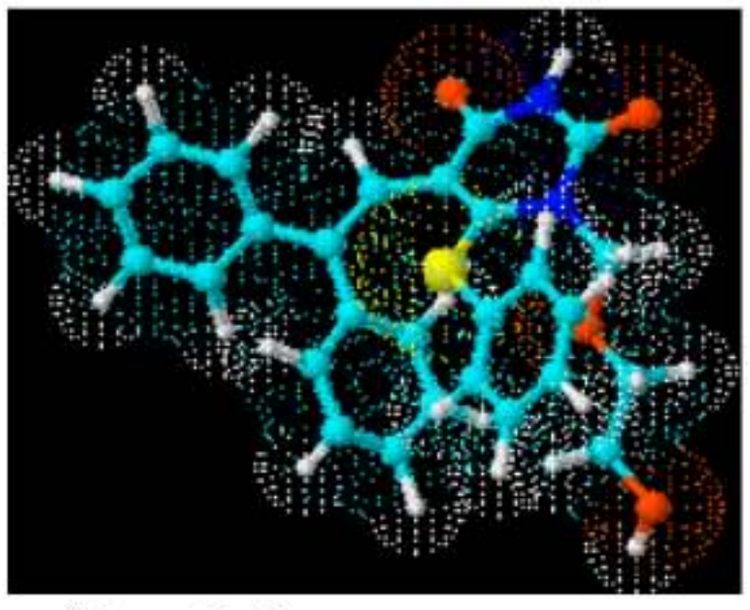

Fig. 4: Optimized structure of Comp. No. 9

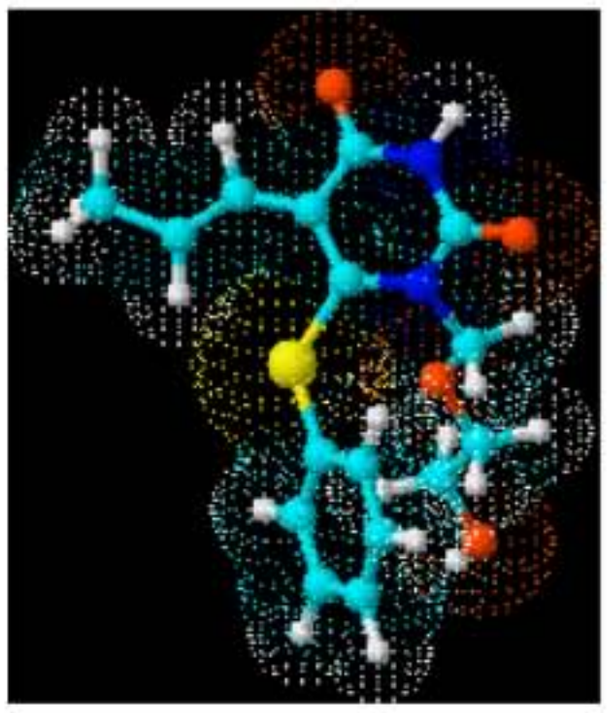

Fig. 5: Optimized structure of Comp. No. 10

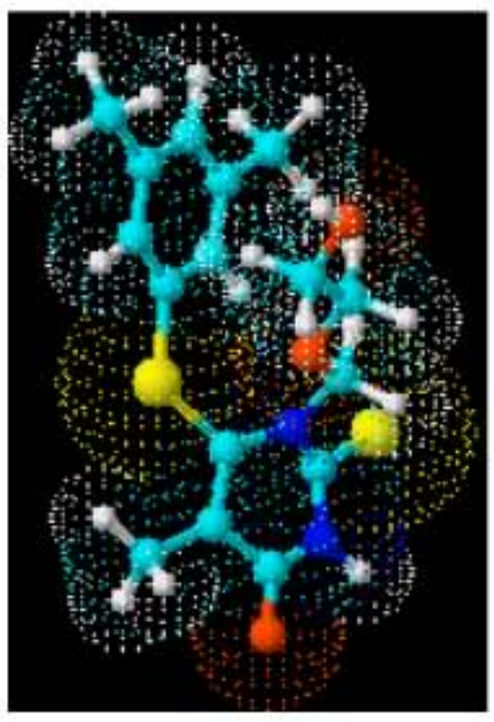

Fig. 6: Optimized structure of Comp. No. 27

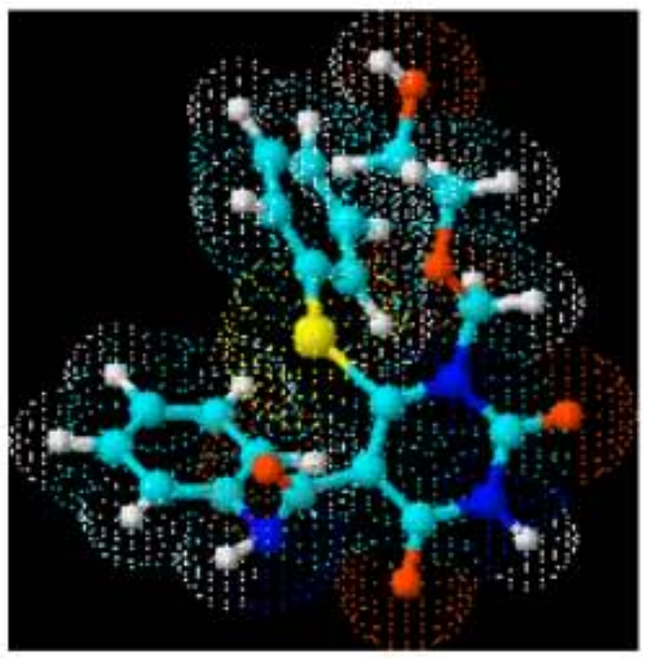

Fig. 7: Optimized structure of Comp. No. 33

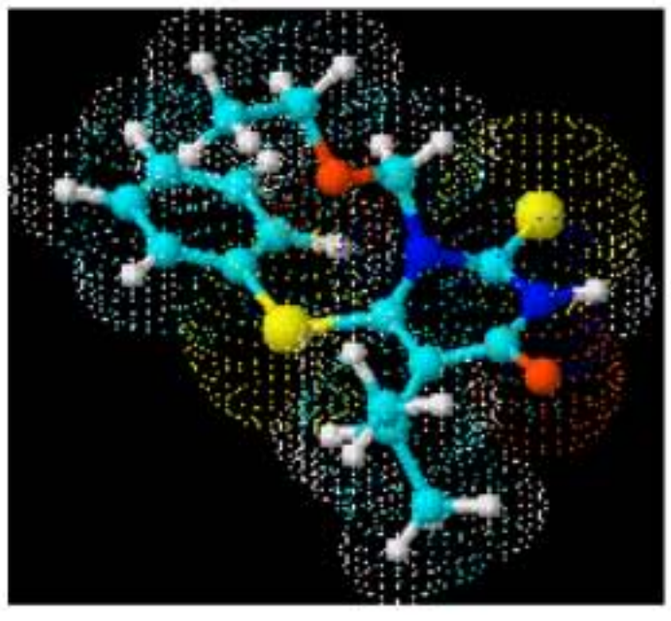

Fig. 8: Optimized structure of Comp. No. 52

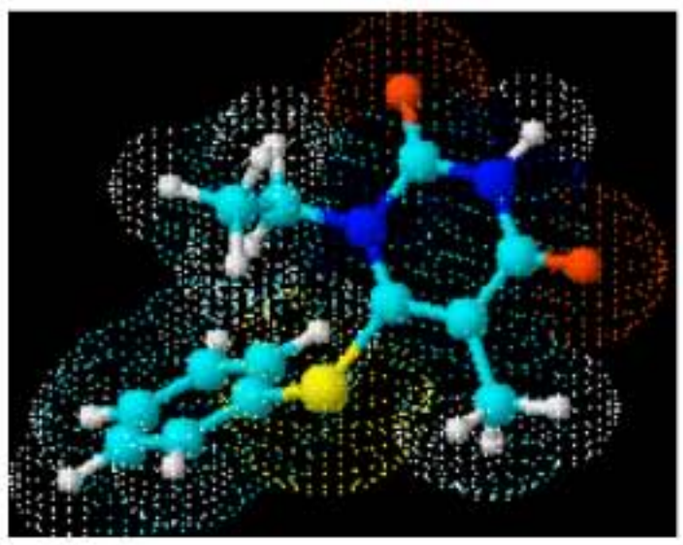

Fig. 9: Optimized structure of Comp. No. 59

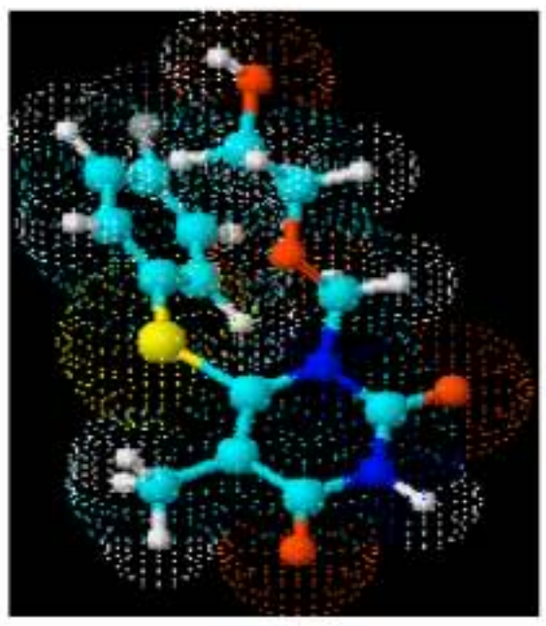

Fig. 10: Optimized structure of Comp. No. 73

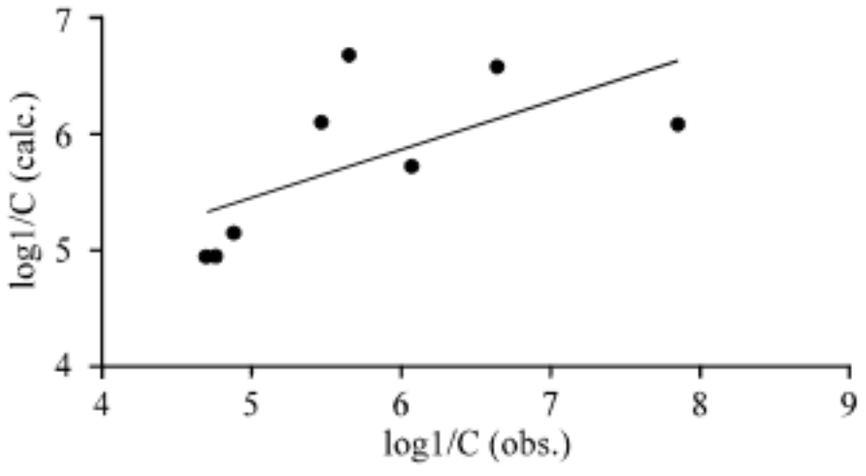

Fig. 11: Graph obtained between calculated and observed log1/C values from Eq. 7

Table 7: Molecular modeling parameters for compounds having minimum residue

Comp. No.	TE	DpM	RMSg
7	33.21	1.479	0.9871
9	41.21	3.238	1.6560
10	18.18	3.438	0.1054
17	3.73	3.384	0.0874
33	17.66	3.481	0.0772
52	7.78	3.560	0.0770
59	8.77	2.950	0.3940
73	7.94	3.186	0.0960

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

Non-of the modeling parameter shows the significant univariate correlation but result obtained from bivariate correlation is significant and model is shown below:

$$\log 1/C = 2.222(\pm 1.4249)\ RMSg - 0.1155\ (\pm 0.0621)\ TE + 6.7910 \quad (7)$$
$$n = 8,\ Se = 0.9846,\ R = 0.6420,\ F = 1.753$$

Equation 7 demonstrates that the compound having the minimum Total Energy (TE) and highest RMSg is favorable for the inhibition activity against RT's for the HEPT derivatives. Graphical representation of the model (Fig. 11) shows that the compound no 27 has the maximum predictive potential and most suitable for the modeling.

Conclusions

From the result and discussion made above we conclude that the classical physicochemical parameters can be used successfully for modeling inhibition of reverse transcriptase-1 by HEPT derivatives and that for the present set of HEPT derivatives MV is find to be prominent. The results also indicate that combination of classical physicochemical parameters and molecular (3D) modeling can be used for select the compound with potential activity.

Acknowledgment

Author's thanks are due to Prof. Padmakar V Khadikar for his excellent guidance to perform this study.

References

ACD-Lab software for calculating the referred physicochemical parameters: Chemsketch 3.O, www.acdlabs.com

Agrawal, V.K., R. Sohgaura, P.V. Khadikar and A. Phadnis, 2001. QSAR study on inhibition of brain 3-hydroxy-anthranilic acid dioxygenase (3-HAO): A molecular connectivity approach. Bioorg. Med. Chem., 9: 3295-3299.

Balaban, A.T., P.V. Khadikar, C.T. Supuran, T. Abhilash and T. Mamta, 2005. Study on supramolecular complexing ability vis-à-vis estimation of pKa of substituted sulfonamides: Dominating role of Balaban index (J). Bioorg. Med. Chem. Let., 17: 3966-3973.

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

Clercq, D.E., 1995. Antiviral therapy for human immunodeficiency virus infections. Microbiol. Rev., 8: 200-239.

Hannongbua, S., L. Lawtrakul and J. Limtrakul, 1996a. Structure-activity correlation study of HIV-1 inhibitors: Electronic and molecular parameters. J. Comput.-Aided Mol. Design, 10: 145-152.

Hannongbua, S., L. Lawtrakul, C.A. Sotriffer and B.M. Rode, 1996b. Three-dimensional quantitative structure-activity relationships study on HIV-1 reverse transcriptaseinhibitors1-[(2-HydroxiEthoxy)methyl]6-[5-Phenyl(thiothaimine)]Quant. Struct. Act. Relat., 15: 389-394.

Hannongbua, S., S. Saen-oon, P. Pungpo and P. Wolschann, 2001. Molecular Calculations on the Conformation of the HIV-1 Reverse Transcriptase Inhibitor 4,5,6,7-tetrahydroimidazo-8-chloro-5-methyl-(3-methyl-2-butenyl)imidazo- [(8-chloro-TIBO). Mh. Chem., 132: 1157-1169.

Hyperchem-7 software for calculating the molecular modeling parameters: www.hyper.com

Khadikar, P.V., K.C. Mathuhr, S. Singh, A. Phadnis, A. Shrivastava and M. Mandloi, 2002a. Study on quantitative structure-toxicity relationships of benzene derivatives acting by narcosis. Bioorg. Med. Chem., 10: 1761-1766.

Khadikar, P.V., S. Karmarker, S. Singh and A. Shrivastava, 2002b. Use of the PI index in predicting toxicity of nitrobenzene derivatives. Bioorg. Med. Chem., 10: 3163-3170.

Kireev, D.B., J.R. Chretien, D.S. Grierson and C. Monneret, 1997. A 3D QSAR study of a series of HEPT analogues: The influence of conformational mobility on HIV-1 reverse transcriptase inhibition. J. Med. Chem., 40: 4257-4264.

Thakur, A., T. Mamta and P.V. Khadikar, 2003. Topological modeling of benzodiazepines receptor binding. Bioorg. Med. Chem., 11: 5203-5207.

Thakur, A., T. Mamta and S. Vishwakarma, 2004a. QSAR study of flavonoid derivatives as p56lck tyrosinkinase inhibitors. Bioorg. Med. Chem., 12: 1209-1214.

Thakur, A., T. Mamta, K. Nitika, J. Ashok and G. Ashok, 2004b. Application of topological and physicochemical descriptors: QSAR study of phenylamino-acridine derivatives. ARKIVOC, 10: 36-42.

Thakur, A., A. Agrawal, M. Thakur and S. Thakur, 2004c. Application of physico-chemical descriptors: Modeling of inhibition of DNA replication in chinese hamster ovary cells V79 by some phenol derivatives. Bioinformatics India, 2: 17-23.

Thakur, M., T. Abhilash and V. Khadikar Padmakar, 2004d. QSAR study on phenolic activity: Need of positive hydrophobic term (logP) in QSAR. Bioorganic Med. Chem., 12: 2287-2293.

Thakur, M., T. Abhilash and V. Khadikar Padmakar, 2004e. QSAR studies on psychotomimetic phenylalkylamines. Bioorg. Med. Chem., 12: 825-831.

Thakur, A., 2005. QSAR Study on benzenesulfonamide DNA binding affinity: Physicochemical approach using surface tension. ARKIVOC, 14: 49-53.

Thakur, M., A. Thakur, P.V. Khadikar and C.T. Supuran, 2005. QSAR study on pKa vis-a`-vis physiological activity of sulfonamides: A dominating role of surface tension (inverse steric parameter). Bioorg. Med. Chem. Lett., 15: 203-209.

Young, S.D., 1993. Non-nucleoside inhibitors of HIV-1 reverse transcriptase. Perspects Drug DescoV. Des., 1: 181-192.