

PREDICTION OF ANTI-HIV ACTIVITY OF 1,3-THIAZOLIDIN-4-ONES: QSAR APPROACH

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

^a *Pharmaceutical Chemistry Research Laboratory, Department of Pharmaceutical Sciences, Dr. H. S. Gour Vishwavidyalaya, Sagar (M.P) – 470 003, India*

^b *Government College of Pharmacy, Aurangabad (MH), India*

In the present work, quantitative structure activity relationship studies were performed on a series of 1,3-thiazolidin-4-one derivatives as anti-HIV agents. Stepwise multiple linear regressions analysis was applied to identify the structural and physicochemical requirements for anti-HIV activity, which was further evaluated for statistical significance and predictive power by internal and external validation. The developed QSAR indicates that the increase in surface area of molecule, substitution of electronegative groups and substitution of hydrophobic groups at R₄ and R₆ leads to better anti-HIV activity. And the negative coefficient of Wiener index, principal moments of inertia X and Y, Radius of Gyration, heat of formation and presence of pyrimidin-4-yl ring substituted at 4th position of thiazolidinone nucleus showed that increase of these properties is not favorable for anti-HIV activity. The information generated from the present study may be useful in the design of more potent thiazolidinone derivatives as anti HIV agents.

(Received February 5, 2008; accepted February 18, 2008)

Keywords: QSAR, Anti-HIV, Thiazolidinone derivatives

1. Introduction

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

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

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

* Corresponding author: dragawal2001@yahoo.co.in, phravi75@rediffmail.com

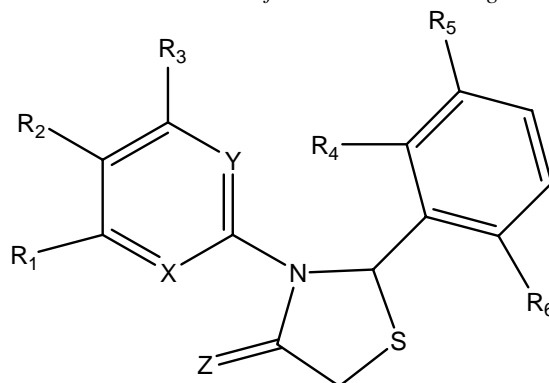
QSAR analyses of HIV-1 reverse transcriptase inhibitors [5], HIV-1 protease inhibitors [6,7] and HIV-1 integrase inhibitors [8] and gp 120 envelope glycoprotein [9] were reported. The present group of authors has developed a few quantitative structure-activity relationship models to predict anti-HIV activity of different group of compounds [10-17]. As a part of ongoing efforts to design novel molecules with potent anti-HIV activity, a QSAR analysis was performed to relate anti-HIV activity of 1,3-thiazolidin-4-one [18-23] derivatives to its physicochemical properties using modeling software WIN CAChe 6.1 (molecular modeling software, a product of Fujitsu private limited, Japan) and statistical software STATISTICA version 6 (StatSoft, Inc., Tulsa, USA).

In the present work we have taken 113 thiazolidinone compounds and their anti-HIV activity from the reported work [18-23]. The compounds which did not show confirmed anti-HIV activity in the above cited literature have not been taken for our study. We carried out QSAR analysis and established QSAR models to guide further structural optimization and predict the potency and physicochemical properties of clinical drug candidates.

2. Modeling

A data set of 113 compounds for anti-HIV activity was used for the present QSAR study. Anti-HIV activity of all the 113 compounds was determined in the same laboratory using same procedure even the compounds have been synthesized by different group of authors. There is high structural diversity and a sufficient range of the biological activity in the selected series of compounds. It insists to select these series of compounds for our QSAR study. The molar concentrations of the thiazolidinone compounds required to produce 50% reduction in the cytopathic effect caused by the virus is stated as the means of at least two experiments were converted to free energy related negative logarithmic values for undertaking the QSAR study. The total set of compounds was initially divided randomly into two groups as training and test set, having 96 compounds in training set and 17 compounds in the test set. Test and training set compounds were chosen manually such that low, moderate and high activity compounds were present approximately in equal proportions in both sets. Training set compounds used to develop QSAR models and the test set compounds used to validate the developed model.

Table 1. Structures of thiazolidinone analogs



Comp. No	X	Y	Z	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆
1	N	N	O	H	H	H	Cl	H	Cl
2	N	N	O	H	H	H	Cl	H	F
3	N	N	O	H	H	Me	Cl	H	F
4	N	N	O	H	H	Me	Cl	H	Cl
5	N	N	O	H	H	Me	F	H	F
6	N	N	O	Me	H	Me	Cl	H	F
7	N	N	O	Cl	H	Me	Cl	H	Cl
8	N	N	O	Me	H	Me	Cl	H	Cl
9	N	N	O	Me	H	Me	F	H	F

10	N	N	O	OMe	H	Me	Cl	H	F
11	N	N	O	OMe	H	Me	Cl	H	Cl
12	N	N	O	OMe	H	Me	F	H	F
13	N	N	O	Me	H	CF ₃	Cl	H	Cl
14	N	N	O	OH	H	Me	Cl	H	Cl
15	N	N	O	Me	H	CF ₃	Cl	H	F
16	N	N	O	Me	H	Ph	Cl	H	Cl
17	N	N	O	Me	H	CF ₃	F	H	F
18	N	N	O	Me	H	Ph	Cl	H	F
19	N	N	O	Me	H	Ph	F	H	F
20	N	N	O	Me	Me	Me	F	H	F
21	N	N	O	Me	Me	Me	Cl	H	Cl
22	N	N	O	Me	Me	Me	Cl	H	F
23	Pyrimidin-4-yl		O		2,6-Dimethyl		Cl	H	Cl
24	Pyrimidin-4-yl		O		2,6-Dimethyl		Cl	H	F
25	Pyrimidin-4-yl		O		2,6-Dimethyl		F	H	F
26	CH	CH	O	H	H	Me	F	H	OMe
27	CH	CH	O	H	H	Me	Cl	Me	F
28	CH	CH	O	H	H	Me	OMe	H	OMe
29	N	CH	O	Me	H	H	F	H	OMe
30	N	CH	O	Me	H	H	Me	H	Me
31	N	CH	O	Me	H	H	F	H	CF ₃
32	N	CH	O	Me	H	H	Cl	H	F
33	N	CH	O	Me	H	H	OMe	H	OMe
34	N	CH	O	H	H	Me	Me	H	Me
35	N	CH	O	H	H	Me	F	H	OMe
36	N	CH	O	H	H	Me	F	H	CF ₃
37	N	CH	O	H	H	Me	Cl	H	F
38	N	CH	O	H	H	Me	OMe	H	OMe
39	N	CH	O	H	H	Br	Me	H	Me
40	N	CH	O	H	H	Br	F	H	CF ₃
41	N	CH	O	H	H	Br	F	H	OMe
42	N	CH	O	H	H	Br	Cl	H	F
43	N	CH	O	H	H	Br	OMe	H	OMe
44	N	CH	O	Me	H	Me	Me	H	Me
45	N	CH	O	Me	H	Me	F	H	OMe
46	N	CH	O	Me	H	Me	F	H	CF ₃
47	N	CH	O	Me	H	Me	Cl	H	F
48	N	CH	O	Me	H	Me	OMe	H	OMe
49	N	N	O	H	H	Me	Me	H	Me
50	N	N	O	H	H	Me	F	H	OMe
51	N	N	O	H	H	Me	F	H	CF ₃
52	N	N	O	H	H	Me	Cl	H	F
53	N	N	O	H	H	Me	OMe	H	OMe
54	CH	CH	O	H	H	H	Cl	H	Cl
55	CH	CH	O	H	H	H	F	H	F
56	N	CH	O	H	H	H	Cl	H	Cl
57	N	CH	O	H	H	H	Cl	H	F
58	N	CH	O	H	H	H	F	H	F
59	N	CH	O	H	Cl	H	Cl	H	Cl
60	N	CH	O	H	Cl	H	Cl	H	F
61	Pyridin-3-yl		O	-	H	H	F	H	F
62	N	CH	O	H	Cl	H	F	H	F
63	N	CH	O	H	Br	H	Cl	H	Cl

64	N	CH	O	H	Br	H	Cl	H	F
65	N	CH	O	H	Br	H	F	H	F
66	N	CH	O	Br	H	H	Cl	H	Cl
67	N	CH	O	Br	H	H	Cl	H	F
68	N	CH	O	Br	H	H	F	H	F
69	N	CH	O	H	H	Me	Cl	H	Cl
70	N	CH	O	H	H	Me	Cl	H	F
71	3-Me-Pyridin-2-yl		O	H	H	H	F	H	F
72	N	CH	O	H	H	Me	F	H	F
73	N	CH	O	H	Me	H	Cl	H	Cl
74	N	CH	O	Me	H	H	Cl	H	Cl
75	N	CH	O	Me	H	H	Cl	H	F
76	N	CH	O	H	Me	H	F	H	F
77	N	CH	O	Me	H	H	F	H	F
78	N	CH	O	Me	H	Me	Cl	H	Cl
79	N	CH	O	Me	H	Me	Cl	H	F
80	N	CH	O	Me	H	Me	F	H	F
81	N	CH	S	Me	H	H	F	H	F
82	N	CH	S	H	H	Me	F	H	F
83	N	N	S	H	H	Me	Cl	H	Cl
84	N	CH	S	Me	H	H	Cl	H	Cl
85	N	H	S	Me	H	Me	F	H	F
86	N	H	S	Br	H	H	F	H	F
87	H	H	O	H	H	Br	Cl	H	Cl
88	H	H	O	H	H	Br	Cl	H	F
89	H	H	O	H	H	Br	F	H	F
90	H	H	O	H	H	Cl	Cl	H	Cl
91	H	H	O	H	H	Cl	Cl	H	F
92	H	H	O	H	H	Cl	F	H	F
93	H	H	O	H	H	Me	Cl	H	Cl
94	H	H	O	H	H	Me	Cl	H	F
95	H	H	O	H	H	Me	F	H	F
96	H	H	O	H	H	Et	Cl	H	Cl
97	H	H	O	H	H	Et	Cl	H	F
98	H	H	O	H	H	Et	F	H	F
99	H	H	O	H	H	OMe	Cl	H	Cl
100	H	H	O	H	H	OMe	Cl	H	F
101	H	H	O	H	H	NO ₂	Cl	H	Cl
102	H	H	O	H	H	NO ₂	Cl	H	F
103	H	H	O	H	H	OMe	F	H	F
104	H	H	O	H	H	NO ₂	F	H	F
105	H	H	O	Cl	H	Cl	Cl	H	Cl
106	H	H	O	Cl	H	Cl	F	H	F
107	H	H	O	F	H	F	Cl	H	Cl
108	H	H	O	Cl	H	Cl	Cl	H	F
109	H	H	O	F	H	F	Cl	H	F
110	H	H	O	F	H	F	F	H	F
111	H	H	O	Me	H	Me	Cl	H	Cl
112	H	H	O	Me	H	Me	Cl	H	F
113	H	H	O	Me	H	Me	F	H	F

All 113 compounds structure were built on workspace of Win CAChe 6.1 (molecular modeling software, a product of Fujitsu private limited, Japan <http://www.cachesoftware.com/contacts/japan.shtml>),

and energy minimization of the molecules was done using Allinger's MM2 force field followed by semi empirical PM3 method available in MOPAC module until the root mean square gradient value becomes smaller than 0.001 kcal/mol Å. Most stable structure for each compound was generated and used for calculating various physico-chemical descriptors like thermodynamic, steric and electronic values of descriptors. Some physico-chemical descriptors were calculated using Molecular modeling pro 6.1.0 (trial version, Cambridge software Corp.) and we have considered some indicator variables also.

All the calculated descriptors and indicator variables (30 descriptors, the complete descriptors data set of all compounds will be provided on request) were considered as independent variable and biological activity as dependent variable. STATISTICA software was used to generate QSAR models. Statistical measures used were n-number of compounds in regression, r-correlation coefficient, r^2 -squared correlation coefficient, F- test (Fischer's value) for statistical significance, SEE- standard error of estimation, q^2 - cross validated correlation coefficient and correlation matrix to show correlation among the parameters. The predictive ability of the generated correlations was evaluated by cross validation method employing a 'leave-one-out' scheme. Validation parameters considered were cross validated r^2 or q^2 , standard deviation based on predicted residual sum of squares (S_{PRESS}) and standard error of prediction (SDEP). The robustness of a QSAR model was checked by Y – randomization test. In this technique, new QSAR models were developed by shuffling the dependent variable vector randomly and keeping the original independent variable as such. The new QSAR models are expected to have low r^2 and q^2 values. If the opposite happens then an acceptable QSAR model can not be obtained for the specific modeling method and data.

3. Results and discussion

A data set of 113 thiazolidinone compounds (Table 1) for anti-HIV activity was used for the present QSAR study. The QSAR studies of the thiazolidinones series resulted in several QSAR equations. The selected descriptors are radius of Gyration (ROG), solvent accessible surface area (SASA), wiener index (WI), principle moments of inertia X coordinate (PMIX), principle moments of inertia Y coordinate (PMIY), heat of formation (HF), highest occupied molecular orbital (HOMO), π value of substituent at R_4 and R_6 [$\pi (R_4 + R_6)$] and INW (if pyrimidin-4-yl group is present at 4th position of thiazolidinone nucleus then INW = 1 otherwise 0) (Table 2).

Table 2. Selected physico-chemical parameters of 1,3,4-thiazolidinone analogs.

Comp No.	SASA	WI	ROG	PMIX	PMIY	HF	HOMO	$\pi (R_4 + R_6)$
1	308.671	744	8.965	1544.828	1996.465	49.451	9.312	1.420
2	287.047	744	8.645	1462.570	1818.282	12.040	9.349	0.850
3	313.157	852	9.082	1698.642	1896.297	2.507	9.326	0.280
4	318.285	852	9.342	1842.429	2028.148	39.955	9.290	0.850
5	307.431	852	9.036	1442.823	1720.772	-35.960	9.373	1.420
6	340.159	964	9.872	1711.858	2269.882	-7.081	9.299	0.280
7	334.897	964	11.055	1897.878	3033.488	30.577	9.344	1.420
8	348.523	964	10.134	1849.153	2424.712	30.375	9.263	0.850
9	338.907	964	9.855	1459.927	2107.224	-45.563	9.347	1.420
10	360.304	1098	10.839	2856.718	1942.071	-37.660	9.359	0.280
11	364.034	1098	10.629	45616.60	1765.954	-180.723	9.262	0.850
12	348.261	1098	10.557	3698.016	1962.895	-75.151	9.400	1.420
13	381.635	1372	12.788	35163.76	2439.568	-114.202	9.467	0.850
14	331.033	964	10.083	1856.353	2435.056	130.389	9.354	1.420
15	372.228	1372	12.497	56142.56	2166.402	-151.639	9.501	0.280
16	454.770	1740	14.847	1999.006	5210.773	65.937	9.268	0.850
17	358.275	1372	12.522	3849.360	2355.752	-190.063	9.542	1.420
18	456.537	1740	14.620	1806.499	4854.010	28.661	9.307	0.280
19	441.924	1740	14.694	1615.294	4641.595	-9.898	9.355	1.420

20	360.853	1092	10.625	1530.916	2439.262	-53.145	9.308	1.420
21	370.055	1092	10.848	1956.012	2726.130	22.763	9.218	0.850
22	373.968	1092	10.617	1720.119	2546.622	-14.700	9.256	0.280
23	350.266	964	10.463	1891.792	2529.057	19.863	9.393	0.850
24	327.930	964	10.517	1662.757	2414.741	-18.078	9.430	0.280
25	324.097	964	10.217	1531.048	2234.470	-55.663	9.457	0.120
26	334.030	966	9.574	32985.15	1892.873	-45.298	8.963	0.850
27	322.709	962	9.884	1875.907	2230.716	-23.124	9.011	-0.040
28	359.719	1086	10.131	1815.543	8635.405	-37.675	8.822	1.120
29	327.569	966	9.752	32189.43	2027.124	-40.136	9.159	0.860
30	324.523	852	9.936	1400.871	2055.246	23.763	9.127	0.120
31	353.824	1328	11.390	64295.70	3124.837	-119.569	9.306	0.850
32	346.838	962	10.577	1529.326	2533.106	-18.624	9.235	-0.040
33	366.778	1086	10.368	1691.376	12487.80	-32.306	8.952	1.120
34	321.884	852	9.400	1570.890	1762.398	23.995	9.123	0.120
35	310.629	966	9.255	47229.80	1688.422	-39.700	9.150	1.020
36	342.801	1328	10.675	2779.373	11671.00	-155.020	9.335	0.850
37	315.381	962	9.594	1852.697	2126.412	-16.987	9.204	-0.040
38	366.976	1086	9.788	1861.741	2372.57	-31.862	8.952	1.120
39	316.539	852	10.561	3487.942	2643.059	40.737	9.213	0.120
40	342.628	1328	11.588	3414.607	17060.64	-142.515	9.528	0.850
41	325.591	966	10.702	7324.804	2173.676	-124.064	9.196	1.020
42	308.467	962	10.792	2012.578	3026.651	-.350	9.329	-0.040
43	368.360	964	10.255	1519.673	2093.781	14.489	9.096	1.120
44	354.452	1087	10.091	1862.102	13044.21	-49.161	9.124	0.120
45	377.313	1083	10.803	1754.571	2529.506	-27.649	9.171	1.020
46	373.166	1476	11.439	3040.749	23010.57	-164.554	9.300	0.850
47	394.894	1216	10.659	33945.03	1985.134	-40.533	9.014	-0.040
48	317.265	852	9.303	1564.315	1768.374	36.142	9.181	1.120
49	323.438	966	9.153	1845.045	19503.44	-27.788	9.188	0.120
50	339.291	1328	10.656	40055.23	2870.548	-142.725	9.429	1.020
51	311.579	962	9.549	1835.465	2144.107	-4.998	9.295	0.850
52	333.760	1086	9.699	1781.850	7961.135	-19.957	8.980	-0.040
53	303.710	744	9.203	1624.216	2067.927	31.182	9.069	1.420
54	275.098	744	8.836	1301.454	1710.582	-44.210	9.074	0.280
55	309.359	744	9.020	1571.236	1981.608	37.004	9.245	1.420
56	288.415	744	8.703	1489.193	1797.619	-.400	9.284	0.850
57	287.136	744	8.653	1248.405	1659.824	-38.862	9.334	0.280
58	272.477	744	8.807	1288.128	1704.820	-36.079	9.380	0.280
59	307.696	866	10.999	1620.014	2718.523	-5.858	9.280	0.850
60	307.379	866	11.062	1410.787	2620.063	-44.341	9.294	0.280
61	319.301	866	11.210	1836.349	12835.12	31.512	9.251	1.420
62	334.383	866	13.085	1868.565	4298.876	46.768	9.327	1.420
63	312.461	866	12.923	1642.586	4124.164	9.391	9.365	0.850
64	312.225	866	13.016	1435.608	4136.303	-29.091	9.418	0.280
65	313.587	852	10.477	2960.410	1966.736	44.444	9.317	1.420
66	310.848	852	10.397	2865.262	1765.108	7.147	9.360	0.850
67	310.778	852	10.383	2848.623	1629.041	-31.394	9.419	0.280
68	314.503	838	9.061	1246.278	1908.720	-41.858	9.301	0.280
69	315.779	852	9.637	1531.023	2193.738	-9.974	9.253	0.850
70	315.504	852	9.640	1277.463	2062.420	-48.431	9.298	0.280
71	322.726	852	10.944	1669.417	2378.101	97.441	9.216	1.420
72	335.431	866	10.004	1757.619	2249.365	27.663	9.165	1.420
73	309.096	866	9.748	1362.220	1996.560	-48.204	9.214	0.280
74	317.348	852	9.139	1707.065	1881.169	-9.491	9.232	0.850

75	311.984	852	9.104	1457.419	1720.688	-47.952	9.262	0.280
76	320.155	852	10.395	1857.652	2009.764	27.934	9.199	1.420
77	359.141	964	10.225	1879.349	2425.988	18.424	9.166	1.420
78	352.354	964	9.971	1733.494	2278.809	-19.000	9.197	0.850
79	343.078	964	9.984	1483.852	2123.123	-57.481	9.222	0.280
80	330.908	852	9.546	1659.654	1838.036	26.119	8.922	0.280
81	342.676	852	9.883	1942.954	2272.483	101.507	8.961	1.420
82	327.053	852	9.798	1910.999	2293.076	113.729	8.638	1.420
83	352.768	964	10.198	1672.194	2197.405	28.551	8.581	0.280
84	322.879	852	10.070	1759.808	2718.851	25.630	8.929	0.280
85	331.893	852	10.765	2929.415	1832.139	43.021	9.048	0.280
86	322.289	852	11.257	2031.226	3263.026	39.687	9.228	1.420
87	316.566	852	11.148	3200.731	1776.182	2.202	9.248	0.850
88	295.516	852	11.121	1623.925	3096.095	-35.666	9.246	0.280
89	317.468	852	10.339	1995.133	2360.711	24.727	9.174	1.420
90	312.653	852	10.128	1744.732	2286.792	-12.777	9.192	0.850
91	290.632	852	10.131	1570.617	2155.448	-50.633	9.184	0.280
92	324.313	852	9.717	1885.334	2136.274	22.066	9.017	1.420
93	319.757	852	9.440	1718.527	2005.365	-15.499	9.031	0.850
94	300.808	852	9.436	1497.112	1857.093	-53.305	9.018	0.280
95	348.800	981	10.157	2085.453	2261.652	17.499	9.025	1.420
96	323.963	981	10.269	1586.495	2199.372	-57.866	9.024	0.850
97	342.855	981	9.947	1863.426	2210.016	-20.091	9.041	0.280
98	331.593	981	10.917	12203.63	1760.639	-6.195	8.980	1.420
99	326.893	981	11.029	26398.90	1537.953	-44.104	8.958	0.850
100	275.612	981	10.647	2998.703	2513.720	-81.711	8.973	0.280
101	314.262	1112	11.919	10658.80	2341.938	-11.914	9.613	0.850
102	314.093	1112	11.530	26894.25	1949.381	-49.466	9.644	0.280
103	323.646	1112	10.805	2975.983	2461.314	25.451	9.540	1.420
104	332.973	964	11.833	2217.132	3196.751	18.728	9.306	1.420
105	326.149	964	11.566	2071.515	3045.202	-18.776	9.327	0.850
106	312.264	964	10.555	1898.922	2620.267	-54.425	9.466	1.420
107	307.576	964	10.252	1756.771	2436.693	-91.950	9.495	0.850
108	305.793	964	11.664	1849.430	2952.378	-56.631	9.327	0.280
109	281.996	964	10.313	1539.361	2316.696	-129.771	9.524	0.280
110	355.855	964	10.494	1909.601	2541.026	12.972	8.974	1.420
111	344.509	964	10.209	1771.720	2370.270	-24.628	8.990	0.850
112	330.361	964	10.252	1547.150	2229.404	-62.458	8.975	0.280
113	294.187	639	8.842	1491.546	2090.252	56.415	9.301	0.280

The following best equations were derived from stepwise multiple linear regression:

$pEC_{50} = -0.199 (\pm 0.087) - 0.219 (\pm 0.065) ROG - 0.00002 (\pm 0.000) PMIX - 0.00008 (\pm 0.000) PMIY - 0.776 (\pm 0.300) INW - 0.002 (\pm 0.0007) WI + 0.013 (\pm 0.003) SASA - 0.008 (\pm 0.001) HF$
(1)

$n = 96, r^2 = 0.467, r^2_{adj} = 0.424, SEE = 0.504, F(7, 88) = 11.01, P < 0.001, q^2 = 0.395, S_{PRESS} = 0.828.$

$pEC_{50} = -5.015 (\pm 3.264) - 0.253 (\pm 0.068) ROG - 0.00001 (\pm 0.000) PMIX - 0.00008 (\pm 0.000) PMIY - 0.876 (\pm 0.305) INW - 0.002 (\pm 0.0007) WI + 0.015 (\pm 0.003) SASA - 0.007 (\pm 0.001) HF + 0.511 (\pm 0.333) HOMO$
(2)

$n = 96, r^2 = 0.481, r^2_{adj} = 0.433, SEE = 0.500, F(8, 87) = 10.07, P < 0.0001, q^2 = 0.409, S_{PRESS} = 0.779.$

$pEC_{50} = -5.508 (\pm 3.194) - 0.263 (\pm 0.067) ROG - 0.00001 (\pm 0.000) PMIX - 0.00008 (\pm 0.000) PMIY - 0.888 (\pm 0.297) INW - 0.002 (\pm 0.0007) WI + 0.016 (\pm 0.003) SASA - 0.006 (\pm 0.001) HF + + 0.537 (\pm 0.326) HOMO + 0.746 (\pm 0.421) [\pi (R_4 + R_6)] - 0.604 (\pm 0.276) [\pi (R_4 + R_6)]^2$
(3)

$n = 96, r^2 = 0.518, r^2_{adj} = 0.461, SEE = 0.487, F(10, 85) = 9.14, P < 0.0001, q^2 = 0.413, S_{PRESS} = 0.732.$

When we considered eleven parameters for developing model, there was no significant improvement in r^2 and q^2 . Equations 2 and 3 were selected for further studies. The compounds 17, 28 and 66 were removed as outliers for Eq. 2, and compounds 17 and 28 were removed as outliers for Eq. 3. Since the residual of observed and calculated activity of the above mentioned compounds were found two times larger than standard deviation of the Eq. 2 and 3, respectively. After removing the outliers we got the following equations:

$$pEC_{50} = -7.805 (\pm 2.821) - 0.248 (\pm 0.058) ROG - 0.00002 (\pm 0.000) PMIX - 0.00009 (\pm 0.000) PMIY - 0.948 (\pm 0.260) INW - 0.002 (\pm 0.0001) WI + 0.017 (\pm 0.003) SASA - 0.008 (\pm 0.001) HF + 0.774 (\pm 0.288) HOMO \quad (4)$$

$$n = 96, r^2 = 0.589, r^2_{adj} = 0.550, SEE = 0.425, F(8, 84) = 15.07, P < 0.0001, q^2 = 0.512, S_{PRESS} = 0.462, SDEP = 0.446.$$

$$pEC_{50} = -7.745 (\pm 2.755) - 0.259 (\pm 0.057) ROG - 0.00002 (\pm 0.000) PMIX - 0.00008 (\pm 0.000) PMIY - 0.959 (\pm 0.254) INW - 0.002 (\pm 0.0007) WI + 0.018 (\pm 0.003) SASA - 0.007 (\pm 0.001) HF + 0.739 (\pm 0.281) HOMO + 0.716 (\pm 0.373) [\pi(R_4 + R_6)] - 0.599 (\pm 0.242) [\pi(R_4 + R_6)]^2 \quad (5)$$

$$n = 96, r^2 = 0.622, r^2_{adj} = 0.577, SEE = 0.416, F(10, 83) = 13.70, P < 0.0001, q^2 = 0.520, S_{PRESS} = 0.460, SDEP = 0.440.$$

Both the models (Eq. 4 and 5) were shown good q^2 and r^2 values. The values given in the parentheses are 92-95% confidence intervals of the regression coefficients. In both equations the contribution of PMIX and PMIY is very less even though we have included those two parameters in our final models. Because, the statistical significance and predictability of the equations were got reduced badly when we removed those two descriptors. Eq. 4 could explain 58.9% and predict 51.2% of the variance of the anti-HIV activity data. Eq. 5 could explain 62.2% and predict 52.0% of the variance of the anti-HIV activity data. The calculated and predicted anti-HIV activity values by Eq. 4 and Eq. 5 are given in Table 3. Inter-correlation between the descriptors is given in Table 4. The predictive ability of the selected models was also confirmed by external r^2_{CVext} method. The Eq. 4 ($r^2_{CVext} = 0.464, r^2_{pred} = 0.549$) had shown less external predictivity than eq. 5 ($r^2_{CVext} = 0.584, r^2_{pred} = 0.742$). So we selected Eq. 5 is the best model to predict anti-HIV activity. According to Tropsha et al., the proposed QSAR model (Eq. 5) is predictive as it satisfies the conditions $r^2_{CVext} > 0.5$ and $r^2_{pred} > 0.6$.

This model showed good correlation coefficient (r) of 0.789 between descriptors and HIV-1 RT inhibitory activity. This model also indicates statistical significance > 99.9% with F value $F(10, 83) = 13.70$. The calculated and predicted activities against observed activities were graphically given in fig. 1 and 2. The robustness of this model was checked by Y-randomization test. The low r^2 and r^2_{CV} values indicate that the good results in our original model are not due to a chance correlation or structural dependency of the training set.

In this QSAR equation, the positive contribution of SASA, HOMO, $[\pi(R_4 + R_6)]$ on the biological activity showed that the increase in surface area of molecule, substitution of electronegative groups and substitution of hydrophobic groups at R_4 and R_6 leads to better anti-HIV activity. The negative coefficient of WI, PMIX, PMIY, ROG, HF and INW showed that increase of these properties is detrimental for the activity. The negative coefficient of $[\pi(R_4 + R_6)]^2$ showed that substitution of highly hydrophobic groups at R_4 and R_6 is not favorable for activity. Meanwhile the positive coefficient of $[\pi(R_4 + R_6)]$ indicates that these hydrophobic groups may interact with amino acid residues (Leu234, Tyr318, His235, Phe227, Trp229, Tyr118) of HIV-1 reverse transcriptase. The negative coefficient of INW indicates that presence of pyrimidin-4-yl at 4th position of thiazolidinone is not favorable for anti-HIV activity.

The negative sign of PMIX and PMIY suggests that the substitution (or mass distribution within the molecules), which does not facilitate the change in state of angular motion of the molecule along X and Y directions, would produce the derivatives with lesser anti-HIV activity. It can be assumed that the interaction between a derivative and HIV-1 reverse transcriptase under that state would be the minimum. HOMO is the highest energy level in the molecule that contains electrons. When a molecule acts as a Lewis base in bond formation, the electrons are supplied from the molecules HOMO. Molecules with high HOMOs are more able to donate their electrons and hence they are relatively reactive compared to molecules with low lying HOMOs, thus the HOMO descriptor should measure the nucleophilicity of a molecule.

Table 3. Observed, calculated and predicted anti-HIV activity.

Compd. No	Exp.Act. pEC ₅₀ (μM)	pEC ₅₀ (μM)			
		Eq.4		Eq.5	
		Cal.Act.	Pred.Act.	Cal.Act.	Pred.Act.
Training set					
1	0.644	0.377	0.356	0.218	0.175
2	0.236	0.437	0.452	0.578	0.612
3	1.131	0.609	0.583	0.737	0.711
5	0.409	0.893	0.917	0.627	0.649
6	1.420	0.676	0.648	0.811	0.778
9	0.767	1.030	1.045	0.776	0.777
10	1.252	0.820	0.796	0.923	0.899
12	0.578	1.007	1.037	0.710	0.728
13	0.000	0.317	0.363	0.333	0.383
15	-0.114	0.259	0.372	0.201	0.307
16	-0.785	-1.034	-1.144	-0.780	-0.778
18	-0.591	-0.578	-0.573	-0.383	-0.296
19	-0.806	-0.474	-0.367	-0.686	-0.642
20	0.398	0.011	-0.199	-0.252	-0.676
23	-0.634	-0.418	-0.304	-0.239	-0.008
24	-0.634	-0.462	-0.373	-0.378	-0.238
25	-1.407	*	*	*	*
26	-0.065	0.242	-0.373	0.363	0.416
27	-0.060	0.288	0.303	0.208	0.242
28	-0.281	0.025	0.056	0.058	0.097
29	0.056	0.197	0.208	0.303	0.326
31	-0.326	-0.317	-0.314	-0.320	-0.318
32	0.492	0.644	0.650	0.588	0.601
33	-0.522	-0.191	-0.146	-0.147	-0.090
34	-0.188	0.361	0.377	0.451	0.489
35	0.541	-0.184	-0.351	-0.162	-0.330
36	-0.480	0.190	0.298	0.152	0.265
37	-0.134	*	*	*	*
39	-0.801	-0.194	-0.170	-0.108	-0.056
40	0.182	-0.473	-0.631	-0.511	-0.685
42	0.457	-0.195	-0.228	-0.290	-0.400
43	1.347	0.750	0.703	0.847	0.801
44	-0.639	-0.116	-0.056	-0.083	0.001
45	0.695	0.889	0.907	0.989	1.020
46	-0.769	-0.721	-0.698	-0.742	-0.728
47	0.631	0.493	0.470	0.455	0.402
48	-0.053	0.251	0.265	0.320	0.342
49	-0.533	-0.862	-1.010	-0.822	-0.971
50	0.274	0.299	0.303	0.187	0.170
51	0.037	0.259	0.275	0.383	0.415
52	0.114	-0.276	-0.312	-0.343	-0.426
53	0.043	0.187	0.192	0.008	0.006
54	0.397	0.444	0.448	0.494	0.501

55	-0.362	0.427	0.468	0.257	0.305
56	0.750	0.499	0.486	0.630	0.621
57	0.556	0.859	0.874	0.924	0.944
58	0.068	0.577	0.611	0.622	0.661
59	-0.250	-0.027	-0.019	0.087	0.106
60	-0.328	0.290	0.323	0.340	0.380
62	-0.857	-0.624	-0.567	-0.806	-0.792
63	-0.179	-0.605	-0.689	-0.490	-0.558
64	-0.093	-0.275	-0.319	-0.226	-0.261
65	-0.686	-0.105	-0.076	-0.292	-0.262
66	0.565	0.225	0.215	0.362	0.351
67	1.194	0.600	0.577	0.671	0.646
68	1.523	1.016	0.992	1.101	1.078
69	0.833	0.539	0.533	0.676	0.670
70	1.004	0.898	0.894	0.971	0.969
72	-0.395	0.372	0.396	0.208	0.240
73	-0.149	0.671	0.694	0.735	0.765
74	1.357	0.694	0.673	0.840	0.812
75	1.276	0.967	0.954	1.041	1.030
77	1.086	0.586	0.561	0.431	0.384
78	1.046	0.879	0.870	1.043	1.043
79	1.377	1.068	1.048	1.155	1.139
80	1.229	*	*	0.485	0.441
81	-0.814	-0.212	-0.154	-0.278	-0.223
82	-0.638	-0.810	-0.853	-0.873	-0.934
83	0.553	-0.029	-0.129	0.183	0.113
85	-0.808	-0.047	-0.014	0.141	0.205
86	0.237	-0.279	-0.308	-0.462	-0.520
87	0.208	0.084	0.079	0.219	0.220
88	0.190	-0.053	-0.069	-0.008	-0.022
89	-0.121	0.027	0.030	-0.168	-0.170
90	0.514	0.328	0.324	0.455	0.453
91	0.215	0.267	0.269	0.301	0.306
92	-0.158	0.220	0.230	0.044	0.055
93	1.097	0.543	0.527	0.690	0.669
94	0.670	0.534	0.525	0.593	0.588
95	0.162	0.302	0.306	0.135	0.133
96	0.807	0.476	0.455	0.561	0.540
97	0.676	0.579	0.576	0.711	0.712
98	0.107	-0.137	-0.150	-0.360	-0.405
99	-0.149	-0.153	-0.153	-0.057	-0.045
100	-0.517	-0.341	-0.279	-0.373	-0.321
101	0.136	-0.436	-0.520	-0.370	-0.451
102	-0.124	-0.229	-0.243	-0.223	-0.239
104	-0.740	-0.229	-0.205	-0.441	-0.417
105	0.553	0.058	0.040	0.166	0.145
106	0.013	0.511	0.531	0.200	0.220
107	0.893	0.853	0.851	0.865	0.863
109	0.607	0.745	0.765	0.658	0.665
110	0.268	0.349	0.353	0.188	0.182
111	0.721	0.562	0.555	0.714	0.714
112	1.066	0.622	0.586	0.692	0.661
113	0.618	0.297	0.268	0.484	0.468
Test set					
4	1.357	0.228	-	0.424	-

7	-1.714	-0.115	-	-0.288	-
8	1.770	0.340	-	0.554	-
11	1.620	1.085	-	1.091	-
14	-1.516	-0.675	-	-0.748	-
17	-0.699	1.042	-	0.643	-
21	1.523	0.270	-	0.495	-
22	1.699	0.744	-	0.912	-
30	-0.758	0.196	-	0.283	-
38	1.301	0.758	-	0.817	-
41	1.469	0.902	-	0.864	-
61	-1.978	-1.183	-	-1.271	-
71	-1.728	-0.641	-	-0.760	-
76	-0.710	0.025	-	-0.164	-
84	-1.060	-0.101	-	0.075	-
103	-1.072	-0.317	-	-0.516	-
108	0.498	-0.069	-	-0.020	-

Where Cal.Act - calculated activity, Pred.Act (LOO) – predicted activity by leave one out method.

Table 4. Inter-correlation matrix of physico-chemical descriptors.

	SASA	WI	ROG	PMIX	PMIY	HF	HOMO	INW	$\pi (R_4 + R_6)$	$[\pi (R_4 + R_6)]^2$
SASA	1									
WI	0.629	1								
ROG	0.541	0.626	1							
PMIX	0.174	0.352	0.174	1						
PMIY	0.267	0.415	0.177	-0.09	1					
HF	0.004	-0.39	-0.08	-0.41	-0.34	1				
HOMO	-0.09	0.183	0.282	0.108	0.024	-0.36	1			
INW	0.089	0.028	0.019	-0.06	-0.04	0.015	0.150	1		
$\pi (R_4 + R_6)$	0.158	0.019	0.092	-0.01	-0.04	0.252	-0.048	0.043	1	
$[\pi (R_4 + R_6)]^2$	0.160	-0.01	0.068	-0.06	-0.06	0.321	-0.089	0.039	0.968	1

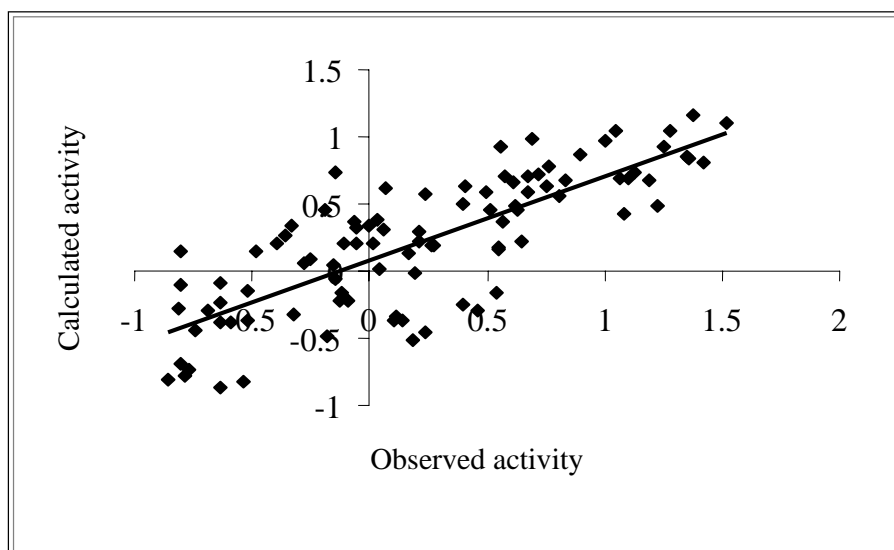


Fig. 1. Observed Vs calculated anti-HIV activity of training set compounds by Eq. 5.

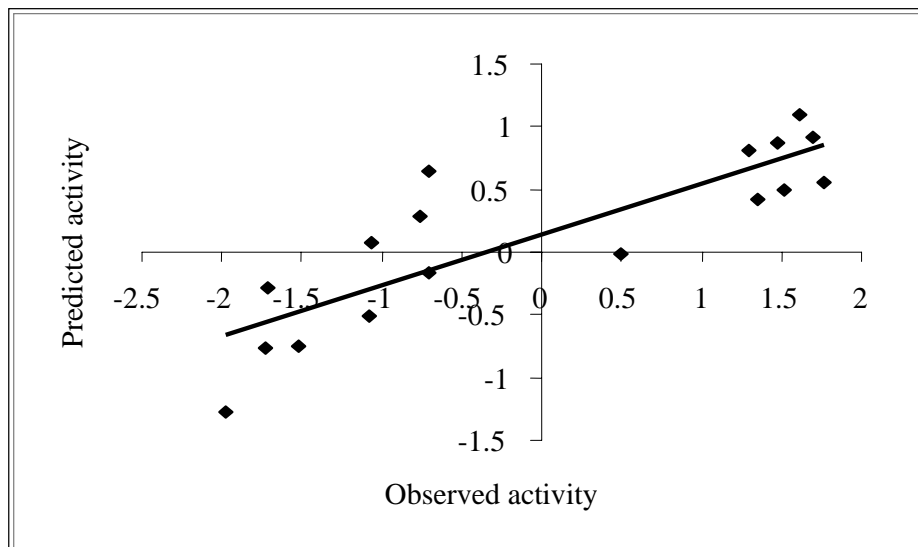


Fig. 2. Observed Vs predicted anti-HIV activity of test set compounds by Eq. 5.

4. Conclusions

Therefore, obtained data by adequate designed QSAR studies allow observing aspects and essential structural characteristics of thiazolidinone related to the increased biological activity, suggesting certain structural requirements for an increased anti-HIV potential. Our results open very interesting perspectives regarding thiazolidinone derivatives with anti-HIV activity.

Acknowledgments

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

References

- [1] R. C. Gallo, S. Z. Salahuddin, M. Popovic, G. M. Shearer, M. Kaplan, B. F. Haynes, T. J. Palker, R. Redfield, J. Oleske, B. Safai, *Science* **224**, 500 (1984).
- [2] E. D. Clercq, *J. Med. Chem.* **38**, 2491 (1995).
- [3] J. Milton, M. J. Slater, A. J. Bird, D. Spinks, G. Scott, C. E. Price, S. Downing, D. V. S. Green, Madar, R. Bethell, D. K. Stammers, *Bioorg. Med. Chem. Lett.* **8**, 2623 (1998).
- [4] S. D. Young, S. F. Britcher, L. O. Tran, L. S. Payne, W. C. Lumma, T. A. Lyle, J. R. Huff, P. S. Anderson, D. B. Olsen, S. S. Carroll, *Antimicrob. Agents. Chemother.* **39**, 2602 (1995).
- [5] P. Pungpo, S. Hannongbua, *J. Mol. Graphics & Modell.* **18**, 581 (2000).
- [6] A. C. Nair, P. Jayatilleke, X. Wang, S. Miertus, W. J. Welsh, *J. Med. Chem.* **45**, 973 (2002).
- [7] P. R. N. Jayatilleke, A. C. Nair, R. Zauhar, W. J. Welsh, *J. Med. Chem.* **43**, 4446 (2000).
- [8] K. Raghavan, J. K. Buolamwini, M. R. Fesen, Y. Pommier, K. W. Kohn, *J. Med. Chem.* **38**, 890 (1995).
- [9] A. K. Debnath, S. Jiang, N. Strick, K. Lin, P. Haberfield, *J. Med. Chem.* **37**, 1099 (1994).
- [10] V. Ravichandran, R. K. Agrawal, *Bioorg. Med. Chem. Lett.* **17**, 2197 (2007).
- [11] V. Ravichandran, V. K. Mourya, R. K. Agrawal, *Arkivoc* **XIV**, 204 (2007).
- [12] V. Ravichandran, V. K. Mourya, R. K. Agrawal, *Internet Electron. J. Mol. Des.* **6**, 363 (2007).
- [13] V. Ravichandran, B. R. Prashanthakumar, S. Sankar, R. K. Agrawal, *Med. Chem. Res.* (Inpress) (2007).
- [14] V. Ravichandran, P. K. Jain, V. K. Mourya, R. K. Agrawal, *Med. Chem. Res.* (Inpress) (2007).
- [15] V. Ravichandran, V. K. Mourya, R. K. Agrawal, *Digest J. Nanomat. Biostruct.* **3**, 9 (2008).

- [16] K. K. Sahu, V. Ravichandran, V. K. Mourya, R. K. Agrawal, *Med. Chem. Res.* **15**, 418 (2007).
- [17] K. K. Sahu, V. Ravichandran, V. K. Mourya, R. K. Agrawal, *Acta. Chim. Slov. (Inpress)* (2007).
- [18] A. Rao, J. Balzarani, A. Carbone, A. Chimirri, E. D. Clercq, A. M. Monforte, P. Monforte, C. Pannecouque, M. Zappala, *Antiviral Res.* **63**, 79 (2004).
- [19] R. K. Rawal, R. Tripathi, S. B. Katti, C. Pannecouque, E. D. Clercq, *Bioorg. Med. Chem.* **15**, 3134 (2007).
- [20] A. Rao, J. Balzarani, A. Carbone, A. Chimirri, E. D. Clercq, A. M. Monforte, P. Monforte, C. Pannecouque, M. Zappala, *Il Farmaco* **59**, 33 (2004).
- [21] M. L. Barreca, J. Balzarani, A. Chimirri, E. D. Clercq, L. D. Luca, A. M. Monforte, A. Rao, P. Monforte, C. Pannecouque, M. Zappala, *J. Med. Chem.* **45**, 5410 (2002).
- [22] A. Rao, A. Carbone, J. Balzarani, A. Chimirri, E. D. Clercq, A. M. Monforte, P. Monforte, C. Pannecouque, M. Zappala, *Il Farmaco* **57**, 747 (2002).
- [23] A. Rao, A. Carbone, J. Balzarani, A. Chimirri, E. D. Clercq, A. M. Monforte, P. Monforte, C. Pannecouque, M. Zappala, *Il Farmaco* **58**, 115 (2003).