QSAR Modeling with the Electrotopological State: TIBO Derivatives

Jarmo Huuskonen†

Division of Pharmaceutical Chemistry, Department of Pharmacy, POB 56, FIN-00014 University of Helsinki, Finland

Received October 20, 2000

Quantitative structure−activity relationships (QSAR), based on the atom level E-state indices and calculated molecular properties (log *P*, *MR*), have been developed for the affinity of a large set of TIBO derivatives against HIV-1 reverse transcriptase (HIV-1 RT) utilizing multiple linear regression techniques. A model with five descriptors, including four atom level E-state indices (carbon atoms 2, 4, 8, and 9) and calculated log *P*, showed good statistics both in the regression ($r^2 = 0.85$ and $s = 0.52$) and leave-one-out cross-validation ($q^2 = 0.80$ and $s_{PRESS} = 0.56$) for the training set of 41 compounds. The statistics for the prediction of anti-HIV activity in the test set of 24 TIBO derivatives were $r^2 = 0.80$ and $s = 0.64$, respectively. The model descriptors indicate the importance of lipophilic and electronic contributions toward HIV-1 RT inhibition of TIBO derivatives used in this study.

INTRODUCTION

During the past decade, several compounds with a wide variety of structures have been synthesized and tested for the treatment of acquired immunodeficiency syndrome (AIDS), which is caused by the human immunodeficiency virus (HIV). One of the most potent and selective inhibitors displaying high activity against HIV-1 reverse transcriptase (HIV-1 RT) is tetrahydroimidazo[4,5,1-jk][1,4]benzodiazepinone or TIBO (Figure 1), discovered by Pauwels and coworkers.[1] Some successful conventional and 3D QSAR approaches in the predicting affinity of designed HIV-1 inhibitors have been reported.[2,3] However, the experimental investigations show that the mechanism of inhibition of TIBO is not yet properly understood.[4,5] The steric and electronic character of molecules play an important role in drug−receptor interaction. The steric effects relate directly to the molecular geometry of interacting molecules at the active site. In a similar manner the electronic characters relate to the electronic distribution of interacting molecules at the receptor site. Hence a correlation study based on steric and electronic properties of molecules could provide insight into the anti-HIV mechanism of TIBO derivatives.

Among different methods for quantifying electronic characteristics of the molecules, the electrotopological state indices based on the chemical graph theory have been found useful in several QSAR[6−13] studies. Using these structural descriptors the data set worth by Hannongbua et al.[2] was examined, the affinity of a large set of TIBO derivatives against HIV-1 RT, and allowed a direct comparison to conventional and 3D QSAR where quantum chemical calculations were used to derive structural descriptors.

METHODS

The affinity of 46 TIBO derivatives against HIV-1 RT along with their chemical structures were the same as used by Hannongbua et al.[2] The biological activity was expressed as log(1/C), where C is the effective concentration of a compound to achieve 50% (IC_{50}) protection of MT-4 cell against the cytopathic effect of HIV-1. In addition, a set of 24 TIBO derivatives was used to test the validation of the model. The log(1/C) values in the training set of 46 derivatives were in the range from 3.74 to 8.30 and in the test set of 24 derivatives from 4.0 to 8.52, respectively. The chemical structures and affinities used in this study against HIV-1 RT are presented in Table 1.

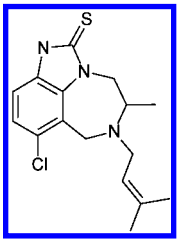


Figure 1. Example structure of a high active TIBO derivative, 8Cl-TIBO.

The physicochemical parameters investigated include the lipophilic parameter log *P* (octanol−water partition coefficient) and the molar refractivity (*MR*) for the steric effects and/or dispersion interaction due to the substituents. The log *P* and *MR* values were calculated using the CLOGP program,[14] taking advantage of the additive-constitutional nature of both parameters.

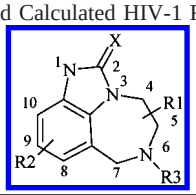
Two-dimensional structures of the molecules under study were constructed using the ChemDraw program running on a 486 PC Pentium workstation and were saved as MDL mol files and imported into the Molconn-Z program[15] for calculation of the atom level E-state indices.

Using the set of 46 TIBO derivatives, multiple linear regression models were developed based on regression algorithms in the SPSS package.[16] The quality of the model was considered as statistically satisfactory on the basis of squared correlation coefficient (r^2), standard deviation (s), and F-statistics (F) when all the parameters in the model

† Corresponding author phone: 358 9 19159170; fax: 358 9 19159556; e-mail: jarmo.huuskonen@helsinki.fi.

10.1021/ci0001435 CCC: $20.00
Published on Web 02/20/2001

Table 1. Structure of TIBO Derivatives and Experimental and Calculated HIV-1 RT Inhibitory Affinities



no.	R1	X	R2	R3	log(1/C)$_{exp}$	log(1/C)$_{calc}$[a]	resid	press[e]
1	H	S	8-Cl	DMA[b]	7.340	7.684	−0.344	−0.470
2	H	S	9−Cl	DMA	6.790	6.790	0.000	0.000
3	5-Et	O	H	2-MA[c]	4.300	5.090	−0.790	−0.849
4	5-*i*-Pr	O	H	2-MA	5.000	5.257	−0.257	−0.278
5	5-*i*-Pr	O	H	DMA	5.000	5.544	−0.544	−0.620
6	5,5-di-Me	O	H	2-MA	4.640	4.859	−0.219	−0.235
7	4-Me	O	H	2-MA	4.490	3.672	0.818	0.956
8	4-Me	S	9-Cl	2-MA	6.170	5.703	0.467	0.549
9	4-Me	S	9-Cl	$CH_2CH(CH_2)_2$	5.660	5.722	−0.062	−0.072
10	4-*i*-Pr	O	H	*n*-Pr	4.130	4.113	0.017	0.018
11	4-*i*-Pr	O	H	2-MA	4.900	4.168	0.732	0.804
12	4-*n*-Pr	O	H	*n*-Pr	3.740	4.379	−0.639	−0.689
13	4-*n*-Pr	O	H	2-MA	4.320	4.333	−0.013	−0.014
14	7-Me	O	H	*n*-Pr	4.080	4.715	-0.635	−0.719
15	7-Me	O	H	DMA	4.920	5.067	-0.147	−0.157
16	7-Me	O	8-Cl	DMA	6.840	6.758	0.082	0.118
17	7-Me	O	9-Cl	DMA	6.790	5.845	0.945	1.081
18	7-Me	S	H	*n*-Pr	5.610	5.977	−0.367	−0.413
19	7-Me	S	H	DMA	7.110	6.330	0.780	0.848
20	7-Me	S	8-Cl	DMA	7.920	8.022	−0.102	−0.137
21	7-Me	S	9-Cl	DMA	7.640	7.106	0.534	0.594
22	4,5-di-Me (cis)	O	H	DMA	4.250	4.140	0.110	0.122
23	4,5-di-Me (cis)	S	H	DMA	5.650	5.401	0.249	0.278
24	4,5-di-Me (trans)	S	H	$CH_2CH(CH_2)_2$	4.870	5.136	−0.266	−0.296
25	4,5-di-Me (trans)	S	H	DMA	4.840	5.401	−0.561	−0.627
26	4-keto-5-Me	S	9-Cl	*n*-Pr	4.300	4.422	−0.122	−0.190
27	4,5-di-benzo	S	H	$CH_2CH(CH_2)_2$	5.000	5.473	−0.473	−0.680
28	5,7-di-Me (trans)	S	H	DMA	7.380	6.549	0.831	0.930
29	5,7-di-Me (cis)	S	H	DMA	5.940	6.549	−0.609	−0.680
30	5,7-di-Me (R,R; trans)	O	9-Cl	DMA	6.640	6.066	0.574	0.685
31	5,7-di-Me (R,R; trans)	S	9-Cl	DMA	6.320	7.327	−1.007	−1.183
32	5,7-di-Me (S,S; trans)	O	9-Cl	DMA	5.300	6.066	−0.766	−0.913
33	4,7-di-Me (trans)	S	H	DMA	4.590	5.420	-0.830	−0.926
34	5,6-$CH_2C(=CHCH_3)CH_2$ (S)	S	9-Cl		5.420	6.448	-1.028	−1.195
35	6,7-$(CH_2)_4$	S	9-Cl		5.700	7.083	−1.383	−1.635
36	5-Me (S)	S	8-Cl	DMA	8.300	8.014	0.286	0.388
37	5-Me (S)	O	9-Cl	DMA	6.740	5.846	0.894	1.022
38	5-Me (S)	S	9-Cl	DMA	7.370	7.111	0.259	0.288
39	5-Me (S)	S	9-Cl	$CH_2CH(CH_2)_2$	7.470	6.838	0.632	0.695
40	5-Me (S)	S	H	$CH_2CH(CH_2)_2$	7.220	6.065	1.155	1.292
41	5-Me	O	H	*n*-Pr	4.220	4.718	-0.498	−0.563
42	5-Me	S	H	*n*-Pr	5.780	5.983	-0.203	−0.228
43	5-Me	O	H	2-MA	4.460	4.783	−0.123	−0.135
44	5-Me	S	H	DMA	7.010	6.335	0.675	0.732
45	5-Me (S)	O	H	DMA	5.480	5.071	0.409	0.440
46	5-Me (S)	S	H	2-MA	7.580	6.045	1.535	1.669

test set

no.	R1	X	R2	R3	log(1/C)$_{exp}$
T1	H	O	H	DMA	4.900
T2	H	O	H	2-MA	4.330
T3	H	O	H	*n*-Pr	4.050
T4	H	O	H	2-EA[d]	4.430
T5	5-Me (S)	S	H	DMA	7.355
T6	5-Me (S)	O	H	Allyl	4.154
T7	5-Me (S)	O	H	*n*-Bu	3.999
T8	5-Me (S)	S	8-F	DMA	8.235
T9	5-Me (S)	O	8-Br	DMA	7.324
T10	5-Me (S)	S	8-Br	DMA	8.521
T11	5-Me (S)	S	8-Me	DMA	7.865
T12	5-Me (S)	S	8-OMe	DMA	7.468
T13	5-Me (S)	S	9,10-di-Cl	DMA	7.592
T14	5-Me (S)	O	8-CN	DMA	5.940
T15	5-Me (S)	S	8-CN	DMA	7.250
T16	5-Me (S)	O	8-Me	DMA	6.000
T17	5-Me (S)	S	10-OMe	DMA	5.330
T18	5-Me (S)	O	10-OMe	DMA	5.180
T19	5-Me (S)	S	10-Br	DMA	5.970
T20	5-Me (S)	S	8-CHO	DMA	6.730
T21	5-Me (S)	O	8-I	DMA	7.060
T22	5-Me (S)	S	8-I	DMA	7.320
T23	5-Me (S)	O	8-C°CH	DMA	6.360
T24	5-Me (S)	S	8-C°CH	DMA	7.530

[a] Calculated by eq 1. [b] DMA = 3,3-dimethylallyl. [c] 2-MA = 2-methylallyl. [d] 2-EA = 2-ethylallyl. [e] Residuals in leave-one-out prediction.

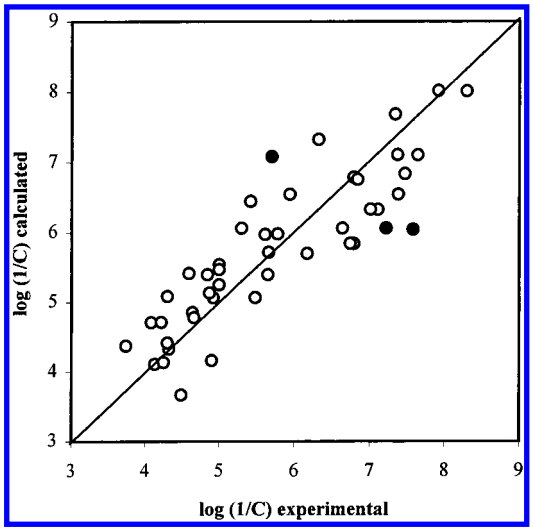


Figure 2. Experimental vs calculated log(1/C) in the training set. Key: five outliers (●) and the remaining (○).

were significant at 95% confidence level ($p < 0.05$). An analysis of predictive ability was carried out in two ways. The predictive ability in the training set was carried out using the leave-one-out cross-validation. For a reliable model, the squared predictive correlation coefficient (q^2) should be > 0.60.[17] In addition, 24 TIBO derivatives (T1−T24) were kept out to test the actual prediction of the model. The actual prediction ability was evaluated using this test set of 24 compounds not included in the training set, and predictive r^2 and s are considered.

One way to improve the overall quality of the regression model is using the jackknife r^2 (r_j^2) values as suggested by Dietrich et al.[18] and Cornish-Bowden and Wang.[19] The jackknife r_j^2 values are simple and useful indicators of the tendency of any particular point in a data set of n cases. For any given compound, C_j, its corresponding r_j^2 can be determined by deleting this compound from the regression analysis and computing the resulting squared correlation coefficient, r^2, from the original model using $n - 1$ data points. The unduly high r_j^2 values might indicate outliers and/or biases, and those with low r_j^2 values might be considered the influential points in the data set, respectively. Hence, the leave-one-out method was employed in two ways, i.e., to test the predictive ability in the training set (q^2 and s_{PRESS}) and to determine possible outliers and/or biases and influential points in the given data set (r_j^2), respectively.

RESULTS AND DISCUSSION

The atom level E-state indices for atoms X, C2, N1, N3, C4, C5, N6, C7, C8, C9, C10, C11, C12, and C13 common for all TIBO derivatives studied were calculated along with lipophilicity (log P) and molar refractivity (MR) parameters. The correlation study showed that the atom level E-state indices X, N1, N3, C10, C11, and C12 and MR had a high pairwise correlation ($r > 0.70$) with each other and with other parameters. After excluding them all the possible regression models of three, four, five, and six parameters were examined. The following five parameters yielded a satisfactory statistical model

$$\log(1/C) = 0.499(\pm 0.169)\log P + 1.165(\pm 0.264)C2 + 2.003(\pm 0.378)C4 - 1.046(\pm 0.273)C8 - 0.305(\pm 0.170)C9 + 4.519(\pm 1.196) \quad (1)$$

$$(n = 46,\ \ r^2 = 0.737,\ \ s = 0.683,\ \ F = 22.40,\ \ q^2 = 0.663,\ \ s_{PRESS} = 0.729)$$

where n is the number of compounds used in the fit, F is the overall F-statistics for the addition of each successive term, and values in parentheses are the 95% confidence limit of each coefficient. The calculated log(1/C) are presented in Table 1, and a plot of experimental log(1/C) versus calculated log(1/C) is given in Figure 2. An examination of possible outliers showed that only two of the residuals are higher than $2 \times s$ (compounds 35 and 46). However, the jackknife test (see Table 2) indicated five data points higher than the limits of the mean r_j^2 values (0.737 ± 0.120). After

Table 2. Structural Parameters in the Models with the Jackknife Results

log(1/C)	log P	C2	C4	C8	C9	r_j^2
7.34	4.159	0.660	0.874	0.804	1.929	0.730
6.79	4.159	0.660	0.890	2.039	0.733	0.733
4.30	4.165	−0.122	0.718	2.101	1.986	0.738
5.00	4.564	−0.125	0.706	2.105	1.989	0.736
5.00	5.093	−0.114	0.723	2.123	2.003	0.739
4.64	3.955	−0.141	0.659	2.090	1.978	0.733
4.49	3.636	−0.132	0.139	2.087	1.976	0.742
6.17	4.349	0.647	0.300	2.028	0.719	0.740
5.66	4.149	0.685	0.369	2.078	0.758	0.737
4.13	4.249	−0.093	0.211	2.130	2.008	0.726
4.90	4.564	−0.117	0.158	2.105	1.989	0.742
3.74	4.579	−0.079	0.257	2.136	2.013	0.727
4.32	4.694	−0.104	0.204	2.111	1.994	0.729
4.08	3.321	−0.108	0.747	2.124	2.000	0.732
4.92	4.165	−0.122	0.716	2.117	1.995	0.734
6.84	4.878	−0.176	0.656	0.709	1.844	0.733
6.79	4.878	−0.176	0.672	1.965	0.637	0.747
5.61	3.321	0.725	0.955	2.216	2.093	0.739
7.11	4.165	0.712	0.924	2.209	2.088	0.739
7.92	4.878	0.657	0.865	0.801	1.937	0.718
7.64	4.878	0.657	0.880	2.058	0.730	0.728
4.25	4.684	−0.124	0.122	2.111	1.994	0.728
5.65	4.684	0.709	0.330	2.204	2.087	0.738
4.87	3.955	0.736	0.384	2.236	2.111	0.735
4.84	4.684	0.709	0.330	2.204	2.087	0.739
4.30	3.475	0.318	−0.007	1.908	0.630	0.729
5.00	5.983	0.681	0.091	2.255	2.125	0.739
7.38	4.684	0.709	0.910	2.216	2.093	0.738
5.94	4.684	0.709	0.910	2.216	2.093	0.743
6.64	5.397	−0.179	0.658	1.972	0.635	0.740
6.32	5.397	0.654	0.866	2.065	0.728	0.753
5.30	5.397	−0.179	0.658	1.972	0.635	0.746
4.59	4.684	0.710	0.346	2.216	2.093	0.743
5.42	3.416	0.669	0.906	2.049	0.743	0.750
5.70	3.570	0.700	1.184	2.077	0.762	0.769
8.30	4.878	0.657	0.860	0.801	1.931	0.712
6.74	4.878	−0.177	0.667	1.953	0.638	0.746
7.37	4.878	0.657	0.876	2.046	0.730	0.728
7.47	4.149	0.684	0.927	2.078	0.758	0.732
7.22	3.436	0.739	0.971	2.229	2.106	0.751
4.22	3.321	−0.109	0.742	2.112	1.995	0.732
5.78	3.321	0.725	0.951	2.204	2.087	0.738
4.66	3.636	−0.133	0.694	2.087	1.976	0.732
7.01	4.165	0.711	0.920	2.197	2.082	0.738
5.48	4.165	−0.122	0.711	2.105	1.990	0.739
7.58	3.636	0.700	0.902	2.179	2.068	0.762

Table 3. Correlation Matrix for the Parameters in Eq 1

	log(1/C)	log *P*	C2	C4	C8	C9
log(1/C)	1.000					
log *P*	0.205	1.000				
C2	0.521	−0.094	1.000			
C4	0.620	−0.216	0.332	1.000		
C8	−0.449	−0.191	0.008	−0.156	1.000	
C9	−0.282	−0.171	−0.053	−0.122	0.027	1.000

Table 4. Comparison of Multiple Linear Regression and CoMFA Models To Predict Anti-HIV Affinity of TIBO Derivatives

	training set						test set		
model	r^2	*s*	*F*	q^2	s_{PRESS}	*n*	r^2	*s*	*n*
eq 1	0.74	0.68	22.4	0.66	0.73	46	0.80	0.65	24
eq 2[a]	0.85	0.53	40.0	0.80	0.56	41	0.80	0.64	24
MLR[b]	0.75	0.66	24.6	0.68	0.72	46			
CoMFA[b]	0.94	0.31	195	0.77	0.61	41	0.87	0.54	24

[a] After excluding compounds with high r_j^2 values. These data points are in bold font in Table 2. [b] The results of Hannongbua et al. (ref 2). Partial charges for atoms C2, C4, C8, C9, and C13 were used as structural parameters in MLR, and electrostatic and steric fields were employed in CoMFA analysis.

excluding these data points (compounds 31, 34, 35, 40, and 46) the following equation was obtained

$$\log(1/C) = 0.524(\pm 0.144)\log P + 1.172(\pm 0.208)C2 + 2.121(\pm 0.305)C4 - 1.053(\pm 0.213)C8 - 0.595(\pm 0.142)C9 + 4.850(\pm 0.968) \quad (2)$$

$$(n = 41,\quad r^2 = 0.851,\quad s = 0.525,\quad F = 40.01,\quad q^2 = 0.802,\quad s_{PRESS} = 0.560)$$

Cross-correlation analysis showed that all pairwise correlations were < 0.332 in this equation indicating a low collinearity as well (see Table 3). Results for the prediction

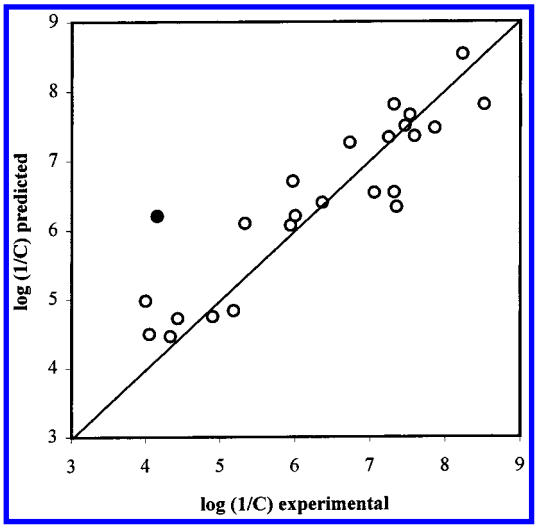


Figure 3. Experimental vs predicted log(1/C) in the test set. Key: one outlier (●) and the remaining (○).

in test set of 24 compounds were $r^2 = 0.800$ and $s = 0.649$ by eq 1 and $r^2 = 0.800$ and $s = 0.644$ by eq 2, respectively. There was one compound with a large estimation error for both equations (compound T6), and when this compound was excluded the standard deviation of predictions were $s = 0.515$ for eq 1 and $s = 0.516$ for eq 2, respectively. Hence, these results are in good agreement with the results obtained for the training sets. Interestingly, the exclusion of outliers and/or biases detected by the jackknife test did not improve the prediction ability outside the model as it did in leave-one-out cross-validation for the training set. Development of the regression model according to the jackknife test results is summarized in Table 4. The experimental and predicted log(1/C) in the test set are presented in Table 5 along with the used structural parameters. In the present study the atom level E-state indices for carbon atoms C2, C4, C8, and C9

Table 5. Parameters and Comparison of Predictive Ability of Multiple Linear Regression and CoMFA Models in the Test Set

no.	log(1/C)	log *P*	C2	C4	C8	C9		eq 1[a]		CoMFA[b]
1	4.90	3.446	−0.119	0.725	2.098	1.985		4.75		5.66
2	4.33	2.917	−0.130	0.708	2.080	1.971		4.46		4.41
3	4.05	2.802	−0.106	0.756	2.105	1.990		4.50		4.04
4	4.43	3.446	−0.126	0.712	2.090	1.979		4.73		4.42
5	7.36	4.165	0.711	0.920	2.197	2.082		6.34		7.25
6	4.15	4.664	−0.118	0.711	1.236	2.073		6.21		4.80
7	4.00	3.850	−0.105	0.746	2.122	2.002		4.98		4.03
8	8.24	4.308	0.549	0.752	−0.171	1.499		8.54		7.61
9	7.32	5.028	−0.139	0.690	1.049	1.990		6.54		7.50
10	8.52	5.028	0.695	0.898	1.141	2.082		7.81		8.27
11	7.87	4.664	0.716	0.919	1.329	2.166		7.47		7.34
12	7.47	4.084	0.667	0.867	0.917	1.999		7.50		7.26
13	7.59	5.471	0.580	0.816	1.951	0.542		7.35		6.94
14	5.94	3.598	−0.220	0.605	0.630	1.767		6.07		6.05
15	7.25	3.598	0.613	0.813	0.722	1.86		7.33		6.63
16	6.00	4.664	−0.118	0.711	1.236	2.073		6.21		6.73
17	5.33	3.884	0.650	0.867	2.141	1.999		6.10		6.35
18	5.18	3.884	−0.184	0.659	2.048	1.906		4.84		5.91
19	5.97	5.028	0.689	0.898	2.187	2.082		6.71		6.85
20	6.73	3.518	0.607	0.807	0.736	1.835		7.27		6.55
21	7.06	5.288	−0.121	0.707	1.206	2.06		6.54		7.02
22	7.32	5.288	0.712	0.916	1.299	2.152		7.81		7.12
23	6.36	4.889	−0.145	0.680	1.088	2.01		6.40		6.29
24	7.53	4.889	0.688	0.888	1.181	2.103		7.66		6.49
							r^2	0.80	(0.87)[c]	0.87
							s	0.65	(0.52)	0.54
							n	24	23	24

[a] This study. [b] Results from ref 2. [c] One compound (T6) is excluded, and the results for the remaining 23 compounds are expressed in parentheses.

were found to be the most important in the same way the atomic charges calculated by the AM1 method for the same set of carbons in the study of Hannongbua et al. were.[2] In their best conventional QSAR model the atomic charge for C13 was included. The statistics for this model were for the regression $r^2 = 0.754$ and $s = 0.662$ and for the leave-one-out prediction $q^2 = 0.677$ and $s_{PRESS} = 0.724$, respectively. In the present study eq 1 gives almost corresponding statistics. Unfortunately, only the CoMFA model was used to predict log(1/C) in the test set of 24 TIBO derivatives. They found the best CoMFA model by excluding five compounds (33, 46, 39, 15, and 40). For the remaining 41 compounds the results for the prediction in the training set was $q^2 = 0.77$and $s_{PRESS} = 0.61$, and in the test set $r^2 = 0.87$ and $s = 0.54$, respectively. In the present study the statistics for the prediction in the training set by eq 2 were $q^2 = 0.80$ and $s_{PRESS} = 0.56$, and in the test set (after excluding T6) $r^2 = 0.87$ and $s = 0.52$, respectively. Hence, we could conclude that the results for the training set and test set are comparable to those by Hannongbua et al.[2] The results of this study clearly show that regression models constructed can be used in the prediction of new TIBO derivatives, and more importantly the prediction of high active compounds is accurate.

The parameters used in regression may be examined for their structural information. In eqs 1 and 2 three parameters (C2, C4, and log *P*) have a positive sign, and two parameters (C8 and C9) have a negative sign. Atom level index for atom C2 increases when a carbonyl group is changed to a thionyl group, and the biological activity increases as well. Atom level index C4 reflects the substitution pattern in the seven-ring system. The anti-HIV activity of TIBO derivatives increases when lipophilicity, log *P*, increases. In fact, the influence of log *P* is parabolic and has an optimal value of log $P_o = 4.20$. The atom level indices for atoms C8 and C9 have a negative sign and highlight the substitution pattern in an aromatic ring system. Electron withdrawing substituents, like halogens, decrease E-state values for C8 and C9, and hence TIBO derivatives with a halogen substituent in one of these positions are more active than derivatives without a substituent in these positions.

CONCLUSIONS

For the past decade, a method for quantifying electronic characteristics of the molecules, the electrotopological state (E-state) indices based on the chemical graph theory and introduced by Kier and Hall,[20] has been found useful in several QSAR studies. E-state indices represent valuable tools in QSAR since they can be computed for any arbitrary molecule and the calculations are made in a clearly described and reproducible way. In addition, these parameters are usually weakly redundant and error free. The chemical interpretation for factors influencing the biological activity of compounds can also be given. The results of the present study, the prediction of anti-HIV affinity of a large set of TIBO derivatives against HIV-1 RT, gives new evidence of the importance of atom level E-state indices as descriptors in QSAR studies.

REFERENCES AND NOTES

(1) Pauwels, R.; Andries, K.; Desmyter, J.; Schols, D.; Kukla, M. J.; Breslin, H. J.; Raeymaechers, A.; Gelder, J. V.; Woestenborgs, R.; Heykants, J.; Schellekens, K.; Janessen, M. A. C.; De Clercq, E.; Janssen, P. A. J. Potent and Selective Inhibition of HIV-1 Replication in vitro by a Novel Series of TIBO Derivatives. *Nature* **1990**, *343*, 470−474.

(2) Hannongbua, S.; Pungpo, P.; Limtrakul, J.; Wolschann, P. Quantitative Structure−Activity Relationships and Comparative Molecular Field Analysis of TIBO Derivatised HIV-1 Reverse Transcriptase Inhibitors. *J. Comput.-Aided. Mol. Des.* **1999**, *13*, 563−577.

(3) Barreca, M. L.; Carotti, A.; Carrieri, A.; Chimirri, A.; Monforte, A. M.; Calace, M. P.; Rao, A. Comparative Molecular Field Analysis (CoMFA) and Docking Studies of Nonnucleoside HIV-1 RT Inhibitors. *Bioorg. Med. Chem*. **1999**, *7*, 2283−2292.

(4) Das, K.; Ding, J.; Hsiou, Y.; Clark Jr., A. D.; Moereels, H.; Koymans, L.; Andries, K.; Pauwels, R,; Janssen, R. A. J.; Boyer, P. L.; Clark, P.; Smith Jr., R. H.; Smith, M. B. K.; Michejda, C. J.; Hughes, S. H.; Arnold, E. Crystal Structure of 8-Cl and 9-Cl TIBO Complexed with Wild-type HIV-1 RT and 8-Cl TIBO Complexed with the Tyr181Cys HIV-1 RT Drug-resistant Mutant. *J. Mol. Biol.* **1996**, *264*, 1085−1100.

(5) Tantillo, C.; Ding, J.; Jacobo-Molina, A.; Nanni, R. G.; Boyer, P. L.; Hughes, S. H.; Pauwels, R.; Andries, K.; Janssen, P. A. J.; Arnold, E. Locations of Anti-AIDS Drug Binding Sites and Resistance Mutations in the Three-dimensional Structure of HIV-1 Reverse Transciptase. Implications for Mechanisms of Drug Inhibition and Resistance. *J. Mol. Biol*. **1994**, *243*, 369−387.

(6) Huuskonen, J.; Salo, M.; Taskinen, J. Neural Network Modeling for Estimation of Aqueous Solubility of Structurally Related Drugs. *J. Pharm. Sci*. **1997**, *86*, 450−454.

(7) Abou-Shaaban, R. R.; al-Khamees, H. A.; Abou-Auda, H. S.; Simonelli, A. P. Atom Level Electrotopological State Indices in QSAR: Design and Testing Antithyroid Agents. *Pharm. Res.* **1996**, *13*, 129−136.

(8) Bualamwini, J. K.; Raghavan, K.; Fesen, M. R.; Pommier, Y.; Kohn, K. W.; Weinstein, J. N. Application of the Electrotopological State Index in QSAR. Analysis of Flavone Derivatives as HIV-1 Integrase Inhibitors. *Pharm. Res.* **1996**, *13*, 1892−1895.

(9) De Gregorio, C.; Kier, L. B.; Hall, L. H. QSAR Modeling with the Electrotological State Indices: Corticosteroids. *J. Comput.-Aided. Mol. Des*. **1998**, *12*, 557−561.

(10) Gough, J. D.; Hall, L. H. Modeling Antileucemic Avtivity of Carboquinenes with Electrotopological State and Chi Indices. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 356−361.

(11) Gough, J. D.; Hall, L. H. Modeling the Toxicity of Amide Herbicides Using the Electrotopological State. *Environ. Toxic. Chem*. **1999**, *18*, 1069−1075.

(12) Huuskonen, J.; Vuorela, H.; Hiltunen, R. Prediction of Activity for a Set Flavonoids Against HIV-1 Integrase. In *Molecular Modeling and Prediction of Bioactivity*; Gundrtofte, K., Jörgensen, F. S., Eds.; Kluwer Academic: Plenum Publishers: New York, 2000.

(13) Kier, L. B.; Hall, L. H. *Molecular Structure Description: The Electrotopological State*; Lemont, B., Ed.; Academic Press: New York, 1999.

(14) CLOGP program; Daylight.com. see http://www.daylight.com/demos/ClogP.

(15) Molconn-Z, vs 3.0; Hall Associated Consulting: 2 Davies Street, Quincy, MA.

(16) SPSS, vs 9.0; SPSS Inc.: Chicago, IL.

(17) Wold, S. Validation of QSAR's. *QuantumStruct-Act. Relat*. **1991**, *10*, 191−193.

(18) Dietrich, S. W.; Dreyer, N. D.; Hansch, C.; Bentley, D. L. Confidence Interval Estimators for Parameters Associated with Quantitative Structure−Activity Relationships. *J. Med. Chem*. **1980**, *23*, 1201−1205.

(19) Cornish-Bowden, A.; Wong, J. T. Evaluation of Rate Constants for Enzyme-catalysed Reactions by the Jackknife Technique. *Biochem. J*. **1978**, *175*, 969−976.

(20) Kier, L.-B.; Hall, L. H. An Electrotopological State Index for Atoms in Molecules. *Pharm. Res*. **1990**, *7*, 801−807.

CI0001435