Indian Journal of Chemistry
Vol. 47B, April 2008, pp. 586-591

QSAR studies on 4-thiazolidinones and 2-azetidinones bearing benzothiophene nucleus as potential anti-tubercular agents

A S Narute , P B Khedekar & K P Bhusari*

Sharad Pawar College of Pharmacy, Wanadongri, Hingna Road, Nagpur 441 110, India

E-mail: kp_bhusari@rediffmail.com

Received 31 March 2006; accepted (revised) 10 January 2008

Quantitative structure-activity relationships (QSAR) study on a series of (substituted 1, 2-dihydro)4–thiazolidinones and 2-azetidinones bearing benzothiophene nucleus with anti-tubercular activity has been carried out using a combination of various physicochemical descriptors. Several significant equations with good co-efficient of correlation (>0.860) have been obtained. The two models are selected using internal predictive power discerned by cross-validated coefficient q^2. Both models highlight some common important feature, *i.e.,* bulky substitution and the high nucleophilicity nature of the molecules, favorable for anti-tubercular activity.

Keywords: Anti-tubercular activity, QSAR, benzothiaphene, anti-tubercular agents

Tuberculosis is a chronic grannulomatous disease[1]. It is estimated that today one-third to one-half of the world population is infected with tuberculosis leading to approximately 6% of all deaths worldwide[2,3]. The causative moiety of the disease is *Mycobacterium tuberculosis*[4]. Despite of the development of several types of synthetic anti-tubercular agents, the incidences of tuberculosis is still increasing in large parts of the world due to the development of resistance in Mycobacterium to the available drugs. Thus, there is an urgent need for novel anti-tubercular agents. With modes of action and chemical structures different from the currently used compounds, it is planned to study the quantitative structure activity relationships[5,6] (QSAR), of some 4-thiazolidinones[7-9] and 2-azetidinones[10-12], which have played an important role in medicinal chemistry. Moreover, they have been studied extensively because of their ready accessibility, diverse chemical reactivity and broad spectrum of biological activity. The focus of the present investigations is the QSAR analysis of thiazolidinones and azetidinones nucleus as potent anti-tubercular agents.

Materials and Methods

The Dataset and Parameters

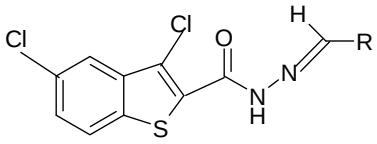
Quantitative structure activity relationship (QSAR) studies of anti-tubercular activity of newly reported thiazolidinones and azetidinones derivatives against *Mycobacterium tuberculosis* reported by Joshi *et al*[13] were performed using linear free energy relation of Hansch. Some of the compounds reported in the original paper were excluded in the present study because of their non-graded quantitative activity data or non-availability of parametric values. Anti-tubercular activity of remaining compounds are given in **Table I**. The biological activity values [BA] reported in the literature were converted to molar units and then further to -log scale and subsequently used as the response variable for the QSAR analysis. The molar anti-tubercular activities were then subjected to multiple regression analysis on different physicochemical parameters and indicator variables (QSAR). The relationships between the activities were also studied to explore the selectivity in terms of structural requirements. The congeneric series possesses one region of structural variation. **Figure 1** shows effect of R substitution on 2-(substituted-benzal- hydrazinocarbonyl)-3,5-chlorobenzo(b)thiophene, **Figure 2** shows effect of R substitution on 2-aryl-5*H*-3-(3′,5′-dichloro-2′-benzo(b)thiophenylamino)-4-thiazolidin- ones and **Figure 3** shows effect of R substitution on the 4-aryl-3-chloro-1-(3′,5′-dichloro-2′-benzo(b)-thiophenylamino)-2- azetidinones.

All the computations in the present study were performed on PIV workstation. The molecular structures of the training set were sketched using

Table I — Anti-tubercular activity data for 4-thiazolidinones and 2-azetidinones benzothiophene derivatives used in this study

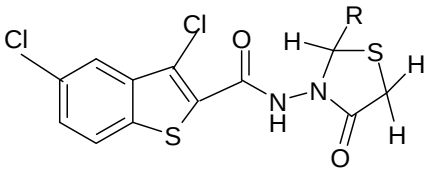
Compd	R	BA	-logBA
1a	3-Br-C_6H_4	07	1.12
1b	4-OCH_3- C_6H_4	14	0.78
1c	3-OCH_3 4-OH- C_6H_3	08	1.06
1d	4-N, N- $(CH_3)_2$- C_6H_4	12	0.86
1e	3-OC_6H_5- C_6H_4	06	1.19
1f	4-S- CH_3- C_6H_4	07	1.12
2a	3-Br-C_6H_4	20	0.6
2b	3-Cl- C_6H_4	15	0.75
2c	3,4-$(O CH_3)_2$- C_6H_3	01	1.99
2d	4-OCH_3- C_6H_4	18	0.65
2e	3-OCH_3 4-OH- C_6H_3	10	0.95
2f	4-N, N- $(CH_3)_2$- C_6H_4	11	0.9
2g	3-OC_6H_5- C_6H_4	18	0.65
2h	4-S- CH_3- C_6H_4	06	1.19
2i	3,4,5-$(O CH_3)_3$- C_6H_2	20	0.6
3a	3-Br-C_6H_4	19	0.62
3b	3-Cl- C_6H_4	26	0.45
3c	3,4-$(O CH_3)_2$- C_6H_3	28	0.41
3d	4-OCH_3- C_6H_4	02	1.69
3e	3-OCH_3 4-OH- C_6H_3	10	0.65
3f	4-N, N- $(CH_3)_2$- C_6H_4	15	0.75
3g	3-OC_6H_5- C_6H_4	20	0.6
3h	4-S- CH_3- C_6H_4	22	0.54
3i	3,4,5-$(O CH_3)_3$- C_6H_2	10	0.95

Chem Draw Ultra module of CS Chem Office 2001 molecular modeling software ver. 6.0, supplied by Cambridge Software Company[14]. The sketched structures were exported to Chem3D module in order to create its 3D model. Each model was "cleaned up" and energy minimization was performed using Allinger's MM2 force field by fixing Root Mean Square Gradient (RMS) to 0.1 Kcal/molÅ. Further, geometry optimization was done using semiemperical AM1 (Austin Model) Hamiltonian method, closed shell restricted wave function available in the MOPAC module until the RMS value becomes smaller than 0.001 Kcal/molÅ. The low energy conformers obtained from the aforementioned procedure was used for the calculation of the descriptors. The descriptors include physicochemical, thermodynamic, electronic and spatial descriptors available in the 'Analyze' option of the Chem. 3D package **(Table II**). The descriptors calculated for the present study accounts for four important properties of the molecules: physicochemical, thermodynamic, electronic and steric, as they represent the possible molecular interactions between the receptor and thiadiazinoacridines.



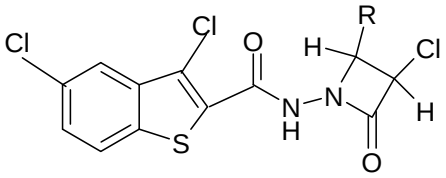
1 a-f

Figure 1 — 2-(Substituted-benzalhydrazinocarbonyl)-3,5-chlorobenzo(b)thiophen



2a-i

Figure 2 — 2-Aryl-5*H*-3-(3′,5′-dichloro-2′-benzo(b)thiophenylamino)-4-thiazolidinones



3a-i

Figure 3 — 4-Aryl-3-chloro-1-(3′,5′-dichloro-2′-benzo(b)thiophenylamino)-2-azetidinones

Multivariate Regression Analysis

The regression analyses were carried out using SYSTAT[15] version 10.2. The statistical quality of equation was judged by the parameters like correlation coefficient (R), standard deviation, standard error of estimation (SEE), variance ratio (f), at specified degree of freedom (df), and 't' values of the regression constant (*i.e.,* the constant term of the regression equation: regression coefficients and intercepts). The use of more than one variable in multivariate equation was justified by autocorrelation study. All the accepted equations have regression constant and f-ratio significant at 95% and 99% level, respectively, if not stated otherwise.

Predicted Residual Analysis

QSAR models can be cross-validated by predicted residual leave one out (LOO) analysis. Each compound of the list is deleted once from the data set

Table II — Descriptors calculated for the QSAR study

Sr No	Descriptor	Type
1	Heat of Formation (HF)	Thermodynamic
2	Boiling Point (BP)	Thermodynamic
3	Critical Pressure (CP)	Thermodynamic
4	Critical Temperature (CT)	Thermodynamic
5	Critical Volume (CV)	Thermodynamic
7	Henry's Law Constant (HLC)	Thermodynamic
8	Ideal Gas Thermal Capacity (IGTC)	Thermodynamic
9	Log P	Thermodynamic
10	Melting Point (MP)	Thermodynamic
11	Molar Refractivity (MR)	Thermodynamic
12	Standard Gibbs Free Energy (SGFE)	Thermodynamic
13	Connolly Accessible Area (CAA)	Steric
14	Connolly Molecular Area (CMA)	Steric
15	Connolly Solvent–Excluded Volume (CSEV)	Steric
16	Ovality (OVA)	Steric
17	Principal Moment of Inertia – X (PMI–X)	Steric
18	Principal Moment of Inertia – Y (PMI–Y)	Steric
19	Principal Moment of Inertia – Z (PMI–Z)	Steric
20	Dipole Moment (D)	Electronic
21	Dipole Moment –X Axis (DX)	Electronic
22	Dipole Moment –Y Axis (DY)	Electronic
23	Dipole Moment –Y Axis (DZ)	Electronic
24	Electronic Energy (EE)	Electronic
25	HOMO Energy (HOMO)	Electronic
26	LUMO Energy (LUMO)	Electronic
27	Repulsion Energy (RE)	Electronic
28	Bend Energy (E_b)	Thermodynamic
29	Charge–Charge Energy (CCE)	Thermodynamic
30	Charge–Dipole Energy (CDE)	Thermodynamic
31	Dipole–Dipole Energy (DDE)	Thermodynamic
32	Non–1, 4 VDW Energy (E_v)	Thermodynamic
33	Stretch Energy (SE)	Thermodynamic
34	Stretch–Bend Energy (SBE)	Thermodynamic
35	Torsion Energy (E_t)	Thermodynamic
36	Total Energy (E)	Thermodynamic
37	Van der Waals e 1,4 Energy (VDWE)	Thermodynamic
38	VDW 1,4 Energy (VDWE)	Thermodynamic
39	Partition coefficient	Thermodynamic

and corresponding regression equation is found out to calculate predicted activity value and predicted residual (press) of deleted compound. The PRESS (predicted residual sum of squares) statistics provides the relations between the observed activity and calculated value (according to PRESS equation). The stability and predictive capacity of the equation were cross validated from PRESS statistics obtained by running VALSTAT[16] programs using "leave-one-out" technique.

Table III — Calculated descriptor values for the given series of compounds

Compd	PMI-X	D	HOMO	E_t
1a	2730.62	2.4079	-8.7059	-5.2672
1b	2160.68	3.5584	-8.6642	-5.3606
1c	2288.34	5.1355	-8.7057	-6.1963
1d	2343.45	6.0749	-8.26	-7.5633
1e	3051.31	3.2713	-8.7667	-11.004
1f	2252.2	3.7861	-8.2024	-7.967
2a	2699.54	3.7577	-8.8652	1.76074
2b	2527.28	3.414	-8.6927	2.42927
2c	3154.93	3.8617	-8.7583	4.29722
2d	2425.3	4.4814	-8.8369	1.78119
2e	2947.99	3.5869	-8.6887	0.47622
2f	2754.77	4.5916	-8.3607	2.07256
2g	2353.77	3.0341	-8.7454	-2.6077
2h	2945.17	2.6997	-8.1723	-1.7454
2i	3297.32	4.3984	-8.8235	3.87878
3a	2998.69	4.1765	-8.8799	5.07532
3b	1826.95	4.1287	-8.9326	5.83832
3c	3052.83	4.3274	-8.8261	4.15428
3d	2659.45	4.5517	-8.8324	5.93828
3e	2281	4.6457	-8.667	3.66351
3f	3068.56	7.1458	-8.4923	2.65777
3g	3224.63	5.5351	-8.9489	-1.2288
3h	2523.64	6.2512	-8.6158	3.63627
3i	3644.38	4.2893	-8.7998	6.50846

Results and Discussion

Biological activity data and various physico-chemical parameters were taken as dependent and independent variables, respectively, and correlations were established using sequential multiple regression analysis. The descriptors selected for modeling anti-tubercular activity of thiazolidinones and azetidinones derivatives are summarized in **Table III**.

The quarter parametric models were obtained and these models are significant for anti-tubercular activity.

Model I

-logBA = 4.48858(± 2.70123) + 0.00018139(± 0.000147211) PMI-X - 0.0606539(± 0.0559619) D +

0.453772(±0.305151)HOMO-0.0246623(± 0.014462) E_t

n =22, R=0.860104, Variance=0.0183792, SD =0.13557, F=12.0823

Model II

-logBA = 4.45369(± 2.67143) + 0.000217417(± 0.000135005) PMX- 0.0646233(± 0.0528988) D + 0.458231(±0.301616)HOME- 0.0232831(± 0.014025) E_t

n=24, R =0.854129, variance=0.0183422, SD =0.135433, F=12.8125

The study of model I and model II revels those thermodynamic parameters like torsion energy (E_t), steric parameters like principal moment of inertia X-axis (PMI-X) and electronic parameters like dipole moment (D) and highest occupied molecular orbit (HOMO) are associated with anti-tumor activity.

In model I, dipole moment, an electronic parameter and is important in case when dipole–dipole interactions are involved in ligand–receptor interactions and torsion energy (E_t) is the thermodynamic parameter, which represents the energy associated with deforming torsion angles in the molecules from their ideal values. The negative coefficients of descriptors suggest presence of conjugation and bulky substituents tolerable for activity, whereas principal moment of inertia, X-axis is a spatial descriptor, which explains the significance of orientation and conformation rigidity of the molecule. The positive coefficient of these descriptor suggest the presence of bulky substituents oriented towards X-axis of the molecules will give better activity and highest occupied molecular orbital. HOMO is an electronic parameter and is the highest energy level in the molecule that contains electrons. It is crucially important in governing the molecular reactivity and properties. When a molecule acts as an electron pair donor in bond formation, the electrons are supplied form the molecule's HOMO. HOMO descriptor denotes nucleophilicity of the molecule and this term was correlated positively. The model suggests that PMI-X and HOMO is of significance having high value of t-test indicating statistical significance of calculated regression coefficient. To confirm this result the value of -logBA was estimated using LOO and correlated with observed value of -logBA. The value of bootstrapping r^2, chance and q^2 in randomized biological activity indicates statistical significance of model as follows.

Bootstrapping r^2 = 0.813112, q^2 = 0.571626, Spress = 0.173942, S_{DEP} = 0.152903

In model II principal moment of inertia and highest occupied molecular orbital positively contribute to biological activity whereas dipole moment and torsion energy negatively contribute to biological activity. The model suggests that principal moment of inertia and highest occupied molecular orbital is of significance having high value if t-test indicating statistically significant calculated regression coefficient. Leave one out cross validation method was used for predictivity of model II. The value of bootstrapping r^2, chance and q^2 in randomized biological activity indicates the statistical significance of the model as follows.

Bootstrapping r^2 = 0.777145, q^2 = 0.586244, Spress = 0.167511, S_{DEP} = 0.149044

The correlation matrix shows model II to be more significant than model I. In model I, all independent parameters (PMI-X, D, E_t and HOMO) have poor (independent) correlation with each other as expected in QSAR analyses but in model-II independent parameters (PMI-X, D, E_t and HOMO) have dependent correlation.

The correlation matrix and predicted activity data for model I and model II are shown in **Table IV, V** and **Table VI, VII,** respectively. **Figure 4** and **Figure 5** show a plot of observed *vs* predicted activities of compounds of model I and model II, respectively. The comparison of model I and model II, the model II was more significant than model I, having good correlation coefficient (R), cross-validated (q^2) value (reflects predictive power of model) bootstrapping (r^2) value (reflect accuracy of the model), and independent correlation between

Table IV — Correlation matrix for parameters in model I

Parameters	PMI-X	D	HOMO	E_t
PMI-X	1.000000			
D	0.055336	1.000000		
HOMO	0.170548	0.077995	1.000000	
E_t	0.256476	0.222781	0.402842	1.000000

Table V — Correlation matrix for parameters in Model II

Parameters	PMI-X	D	HOM0	E_t
PMI-X	1.000000			
D	0.014297	1.000000		
HOMO	0.208958	0.090494	1.000000	
E_t	0.319568	0.247327	0.405912	1.000000

Table VI — Predicted activity data of model I

Compd	Observed-log BA	Predicted -logBA	Calculated -logBA
1a	1.12	0.991225	1.01723
1b	0.78	0.881516	0.865298
1c	1.06	0.717722	0.794575
1d	0.86	1.05483	0.983557
1e	1.19	1.09521	1.13694
1f	1.12	1.15091	1.14194
2a	0.6	0.692616	0.684136
2b	0.75	0.733646	0.735515
2d	0.65	0.597315	0.602833
2e	0.95	0.842528	0.851327
2f	0.9	0.856629	0.864789
2g	0.65	0.855464	0.827379
2h	1.19	1.19709	1.19375
2i	0.6	0.74133	0.720402
3a	0.62	0.625229	0.624585
3b	0.45	0.31487	0.372207
3d	0.41	0.702978	0.67239
3e	0.65	0.588495	0.597353
3f	0.75	0.658391	0.692664
3g	0.6	0.754452	0.707332
3h	0.54	0.575164	0.567922
3i	0.95	0.643861	0.735868

Compd **2c** and **3c** are outlier

Table VII — Predicted activity data of Model II

Compd	Observed -log BA	Predicted -logBA	Calculated -logBA
1a	1.12	1.00262	1.02507
1b	0.78	0.860242	0.848092
1c	1.06	0.698692	0.77437
1d	0.86	1.01338	0.961711
1e	1.19	1.11086	1.14472
1f	1.12	1.12779	1.1255s9
2a	0.6	0.703618	0.694477
2b	0.75	0.741783	0.742706
2c	1.99	0.595105	0.870313
2d	0.65	0.864143	0.876491
2e	0.95	0.871481	0.822645
2f	0.9	0.84961	1.2154
2g	0.65	1.23674	0.752844
2h	1.19	0.772968	0.648548
2i	0.6	0.651867	0.354963
3a	0.62	0.289744	0.696674
3b	0.45	0.723234	0.5926
3c	0.41	0.583873	0.705752
3d	1.69	0.684702	0.72504
3e	0.65	0.776434	0.565729
3f	0.75	0.57102	0.78498
3g	0.6	0.735432	0.565729
3h	0.54	0.57102	0.78498
3i	0.95	0.735432	0.870313

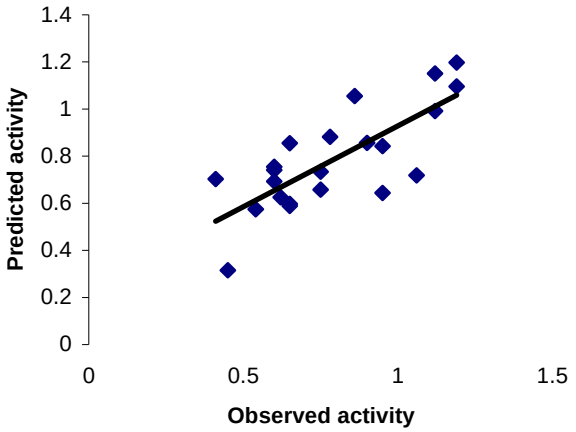


Figure 4 — Graph between observed activity and predicted activity of model I

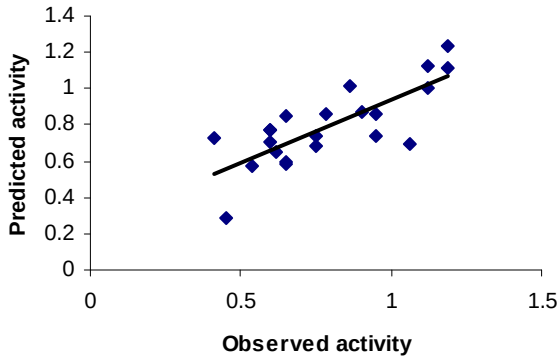


Figure 5 — Graph between observed activity and predicted activity of model II

parameters as expected in QSAR analyses. These results show that such models can be helpful for theoretical prediction of anti-tubercular activity of new molecules.

Conclusion

QSAR analysis was performed on a series of anti-tubercular activity of thiazolidinones and azetidinones derivatives using molecular modeling program Chemoffice 2001. QSAR models were proposed for anti-tubercular activity of the thiazolidinones and azetidinones using descriptors employing sequential multiple regression analysis method. The predictive power of each model was estimated with bootstrapping r^2 method and leave one out cross validation method. It was observed from the selected models that biological activity of thiazolidinones and azetidinones derivatives is governed by thermodynamic and steric properties of the molecules. The models also provide valuable

insight into the mechanism of action of these compounds. The result of the study suggests involvement of partition coefficient in the mechanism of anti-tubercular action of thiazolidinones and azetidinones. The study will be helpful in the design of better anti-tubercular analogs of thiazolidinones and azetidinones derivatives for anti-tubercular activity.

Acknowledgements

One of the authors, Ashok Narute, is grateful to the All India Council for Technical Education (AICTE) for providing fellowship. The authors wish to especially thank Dr. P. Trivedi and Director, Department of Pharmacy, Shri G.S. Institute of Technology and Sciences for providing software for study and Principal, Sharad Pawar College of Pharmacy, Nagpur for providing the necessary facilities for undertaking this research work.

References

1 Tripathi K D, *Essentials of Medical Pharmacology*, 5th Edn (Jaypees Brothers Medical Publishers Pvt Ltd, New Delhi), **2003**, 689.
2 Marwick C, *JAMA*, 267, **1992**, 174.
3 Daniel Y M, *Tuberculosis* in *Harrison's Principles of Internal Medicine*, 12th Edn, edited by Jean Wilson D (McGraw-Hill, New York), **1991**.
4 Williams D A & Lemke T L, *Foye's Principle of Medicinal Chemistry*, 5th Edn (Lippincott Wiliams and Wilkins, New York), **2002**, 905.
5 Hansch C, *Comprehensive Medicinal Chemistry* (Paragon Press, New York), 4, **1990,** 579.
6 Smith H J & Williams H, *Introduction to Principle of Drug Design* (John Wright and Sons Ltd, Bristol), **1983**, 216.
7 Sharma R C & Kumar D, *J Inst Chem Soc*, 77, **2000,** 492.
8 Joshi H, Upadhay P S & Baxi A J, *Indian J Chem*, 39B, **2000**, 967.
9 Ingle S, Sawale A R, Ingle R D & Mane R A, *Indian J Chem*, 40B, **2001**, 124.
10 Kagathara P, Upadhay T, Doshi R & Parekh H H, *Indian J Het Chem*, 10, **2000**, 9.
11 Matsui N, *Jpn Kokai Tokyo JP*, 07, **2000**, 652; *Chem Abstr,* 132, **2000**, 64109u.
12 Desai K R, *Asian J Chem*, 132, **2000**, 279145.
13 Joshi H S, Thaker K M & Kachhadia V V, *Indian J Chem,* 42B**, 2003**, 1544.
14 CS Chem Office, version 6.0, Cambridge Soft Corporation, software publisher Association, 1730 M Street, NW, Suite 700, Washington DC, 20036 (202), 452-1600, USA.
15 SYSTAT 10.2 version supplied by Systat Software Inc.
16 Gupta A K, Babu M A & Kaskhedikar S G, VALSTAT: Validation Program for Quantitative Structure Activity Relationship Studies, *Indian J Pharm Sci*, 66, **2004**, 396.