

QSAR Studies of Antibacterial Activity of 4-Thiazolidone Derivatives

Anita K.¹, Sarita Shrivastava², Shweta Sharma³

^{1,3} Dept. of Chemistry, Career College, Bhopal, Madhya Pradesh, India

² Dept. of Chemistry, Govt. MVM, Bhopal, Madhya Pradesh, India

Abstract: In the present work efforts have been made to study the effect of 4-thiazolidone derivatives on the antibacterial activity by QSAR method based on multiple regression analysis. The QSAR models obtained are subjected to cross validation by leave one out procedure to get well predictive statistically significant QSAR models which help to explore some expectedly potent compounds.

Key words: Antibacterial activity, QSAR, Regression analysis, Cross validation

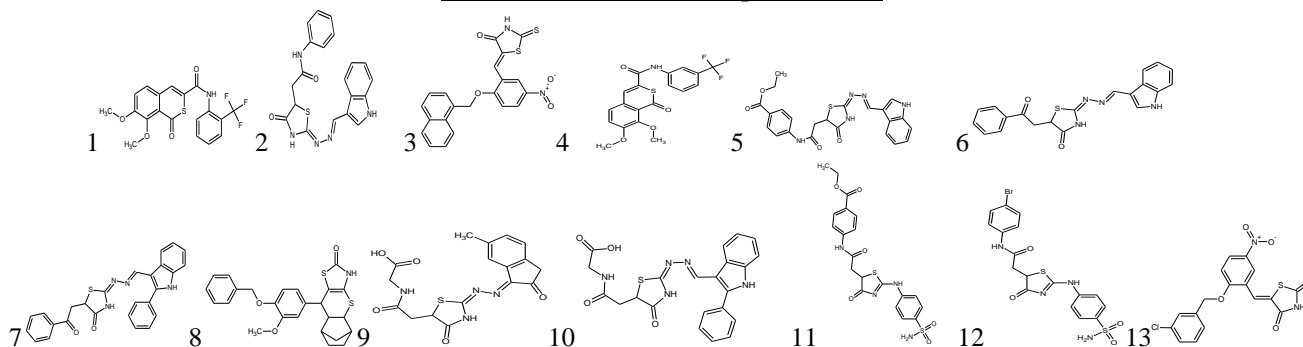
Introduction

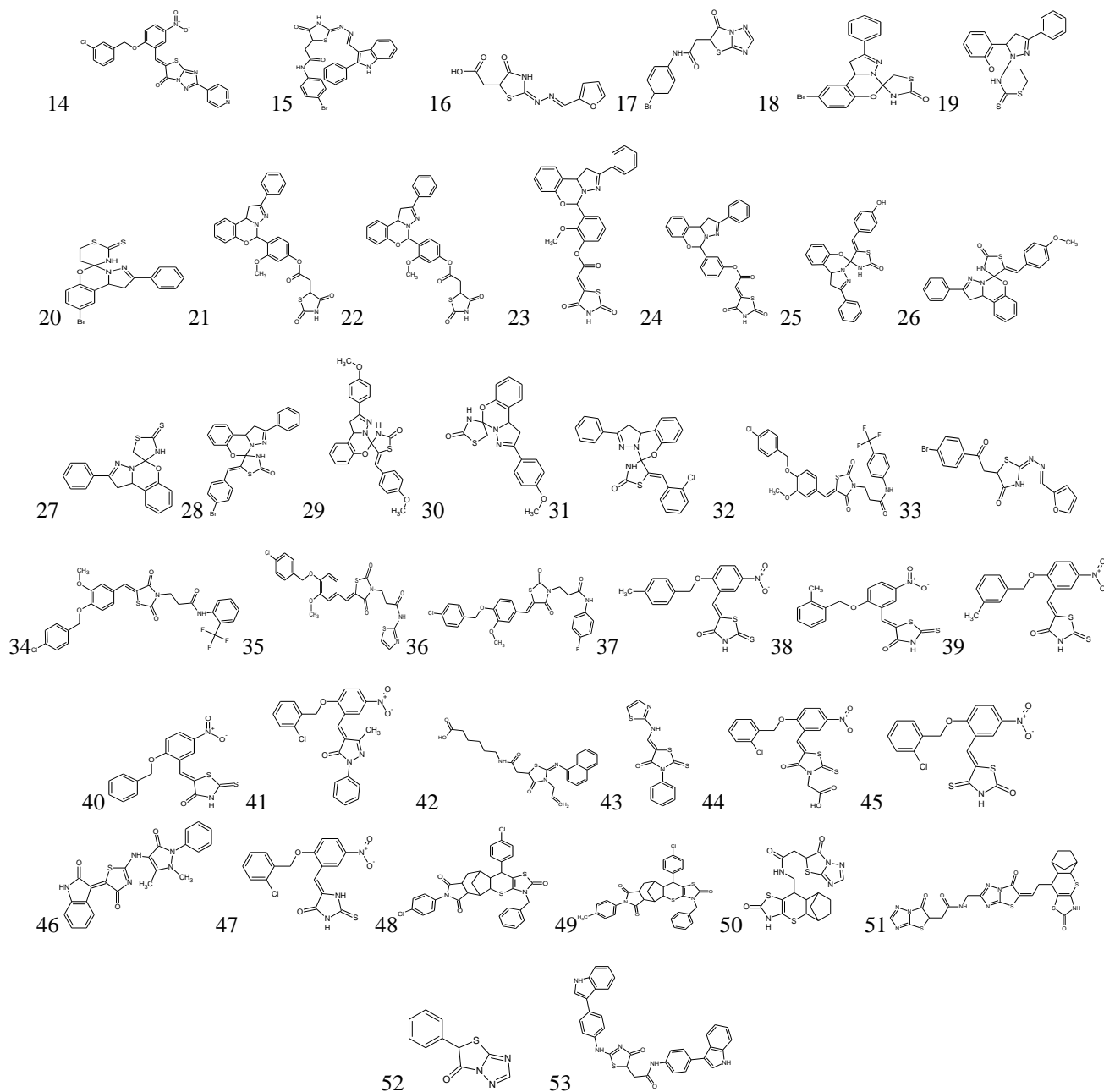
Heterocyclic compounds are one of the major groups of modern synthetic drugs. Thiazolidones, especially penicillin, have a special place in the group of antibacterial drugs. Thiazolidones and their derivatives are related to antimicrobial, antiviral, antitumor, antihypertensive, anti-inflammatory, hepatoprotective, and sedative agents [1-2]. Searching of efficient antibacterial agents is one of the main tasks of clinical practice due to the appearance of antibiotic resistant strains. Quantitative structure activity relationship (QSAR) analysis of organic compounds was included in the development of novel antibacterial and antiviral agents [3, 7]. Computer technologies based on model development provide preliminary screening and molecular design of new organic substances with predicted set of useful properties (including antibacterial activity). Thus the aim of present work is to study the effect of 4-thiazolidone derivatives on the antibacterial activity by QSAR method based on multiple regression analysis.

Materials and methods

Quantitative structure activity modeling establishes a quantitative correlation between chemical structure and biological activity [8]. In the present study, 53 thiazolidone derivatives are considered having biological activity (BA) in terms of log values. The compounds are taken from the literature [9]. Dragon 6 software is used for the calculation of descriptors. The structural details of the compounds used in the present study are given in Table 1. The clinical isolates of the pathogenic bacteria *C. albicans* has been used as the test microorganisms. The biological activities of the thiazolidone derivatives are indicated as log BA.

Table 1. Structures of compounds used





The calculated descriptors are reported in table 2. From the descriptors calculated, useful descriptors are selected by variable selection procedure of multiple regression analysis using NCSS software [10]. The models obtained are subjected to cross validation by leave one out procedure [11].

Table 2. Calculated values of different parameters of 4-thiazolidone derivatives along with their antimicrobial activity values

Mol. ID	X5A	GGI9	RDF060m	Mor03u	Mor12m	G1e	G3s	C-042	log BA
M-1	0.073	0.259	4.654	-4.466	1.261	0.169	0.172	0	0.785
M-2	0.083	0.196	5.964	-4.28	1.222	0.154	0.172	0	0.908
M-3	0.076	0.267	13.267	-5.595	2.26	0.156	0.171	0	0.708

M-4	0.073	0.297	6.121	-4.148	1.154	0.156	0.172	0	1.061
M-5	0.081	0.282	13.764	-6.285	1.271	0.148	0.174	0	0.708
M-6	0.081	0.156	6.087	-4.546	1.213	0.156	0.174	0	0.785
M-7	0.076	0.229	14.359	-7.046	1.493	0.149	0.165	0	0.763
M-8	0.063	0.162	14.496	-9.528	0.68	0.147	0.186	0	1.049
M-9	0.078	0.238	5.218	-6.486	0.994	0.156	0.174	0	0.785
M-10	0.077	0.264	13.085	-5.874	0.967	0.15	0.167	0	0.908
M-11	0.079	0.362	9.659	-6.676	0.594	0.149	0.167	0	0.944
M-12	0.081	0.322	11.008	-5.285	0.634	0.156	0.183	0	0.851
M-13	0.078	0.199	7.381	-4.233	1.78	0.161	0.198	0	0.929
M-14	0.072	0.356	12.813	-5.22	2.124	0.152	0.164	0	0.978
M-15	0.077	0.316	27.397	-5.611	2.196	0.147	0.163	0	0.716
M-16	0.089	0.089	3.06	-1.945	0.45	0.174	0.193	0	0.785
M-17	0.079	0.073	9.368	-3.554	1.19	0.171	0.219	1	1.190
M-18	0.062	0.071	19.241	-5.862	1.768	0.159	0.177	0	0.708
M-19	0.063	0.064	12.826	-6.334	0.658	0.156	0.177	0	0.929
M-20	0.062	0.098	20.568	-6.897	1.902	0.156	0.175	0	0.792
M-21	0.068	0.362	16.199	-9.02	2.048	0.144	0.16	0	1.061
M-22	0.068	0.362	16.199	-9.02	2.048	0.144	0.16	0	0.708
M-23	0.068	0.386	19.464	-8.958	2.014	0.145	0.16	0	0.813
M-24	0.069	0.307	17.448	-7.487	1.598	0.147	0.17	0	0.785
M-25	0.063	0.192	16.607	-6.093	2.226	0.16	0.167	0	0.845
M-26	0.062	0.236	14.539	-6.787	1.737	0.148	0.174	0	0.954
M-27	0.062	0.053	11.975	-6.939	0.807	0.159	0.179	0	0.929
M-28	0.063	0.192	12.452	-7.715	1.252	0.151	0.167	0	0.908
M-29	0.062	0.348	18.586	-5.993	2.623	0.154	0.163	0	0.929
M-30	0.062	0.144	9.742	-5.032	1.611	0.156	0.175	0	0.959
M-31	0.062	0.194	15.79	-5.084	1.727	0.151	0.167	0	1.004
M-32	0.079	0.307	16.829	-7.583	1.753	0.144	0.158	0	0.708
M-33	0.083	0.12	4.978	-3.652	0.967	0.162	0.179	0	0.785
M-34	0.079	0.479	19.933	-8.513	1.629	0.144	0.158	0	0.763
M-35	0.08	0.255	20.369	-6.855	1.341	0.147	0.179	0	0.708
M-36	0.081	0.287	15.107	-6.855	1.457	0.145	0.161	0	0.740
M-37	0.079	0.185	6.894	-4.351	1.973	0.158	0.175	0	0.740
M-38	0.078	0.239	8.172	-4.675	1.54	0.158	0.175	0	0.929
M-39	0.078	0.199	6.86	-5.018	1.851	0.158	0.175	0	0.771
M-40	0.079	0.164	7.096	-3.969	1.824	0.161	0.19	0	0.875
M-41	0.075	0.38	18.638	-5.466	2.01	0.151	0.167	0	0.708
M-42	0.078	0.232	5.664	-11.414	0.673	0.153	0.167	0	0.740
M-43	0.078	0.014	9.097	-3.541	0.465	0.171	0.188	0	0.740
M-44	0.078	0.368	8.154	-5.474	1.995	0.156	0.169	0	0.822
M-45	0.078	0.239	12.411	-4.304	1.726	0.161	0.198	0	0.838
M-46	0.069	0.364	12.261	-5.172	1.554	0.152	0.177	0	1.041
M-47	0.078	0.239	11.619	-4.294	0.865	0.16	0.187	0	0.825
M-48	0.059	0.587	25.421	-10.281	0.748	0.141	0.156	0	1.022

M-49	0.059	0.587	23.485	-10.696	0.783	0.14	0.156	0	1.025
M-50	0.063	0.247	14.126	-10.854	0.022	0.152	0.171	1	1.182
M-51	0.064	0.321	16.314	-13.82	-0.127	0.144	0.158	1	1.179
M-52	0.072	0	2.184	-2.593	0.513	0.183	0.204	1	1.185
M-53	0.074	0.327	11.344	-8.088	2.271	0.142	0.182	0	0.964

Results and discussion

The parameters calculated have been summarized in Table 2 which include ${}^5\chi^A$ (average connectivity index of order 5), GGI9 (topological charge index of order 9), RDF060m (radial distribution function descriptors), Mor03u (3D-Molecule Representation of Structures based on Electron diffraction- un-weighted), Mor12m (3D-Molecule Representation of Structures based on Electron diffraction- weighted), G1e (1st component symmetry directional WHIM index weighted by Sanderson electronegativity), G3s (1st component symmetry directional WHIM index weighted by I- state) and C-042 (atom centered fragments)[12-17]. Correlation matrix showing correlatedness among parameters and activity is presented in table 3. A close look at this table reveals that only C042 shows moderate correlation. But the combination of different parameters gives better result. The data was subjected to regression analysis and statistically significant models obtained have been reported in table 4. During the study it was observed that compound no. 6 and 26 are serious outliers. So these compounds are omitted. After deletion there is a dramatic increase in the value of R^2 is observed. The R^2 value changes from 0.8148 to 0.8629. Different statistically significant models are given below:

Six variable model

$\log BA = 0.2703 \pm (0.0407)C042 - 4.6721 \pm (2.1859)G1e + 5.2739 \pm (1.2786)G3s + 0.4501 \pm (0.1239)GGI9 - 0.0124 \pm (0.0026)RDF060m - 11.2358 \pm (1.6008){}^5\chi^A + 1.5199$
 $N=51, R^2 = 0.7660, R^2_A = 0.7341, F=24.012$

Seven Variable model

$\log BA = 0.3540 \pm (0.0457)C042 - 9.6499 \pm (2.5484)G1e + 4.4989 \pm (1.1925)G3s + 0.4364 \pm (0.1132)GGI9 + 0.0247 \pm (0.0079)Mor030u - 0.0136 \pm (0.0024)RDF060m - 13.0431 \pm (1.5701){}^5\chi^A + 2.7216$
 $N=51, R^2 = 0.8096, R^2_A = 0.7787, F=26.127$

Eight variable model

$\log BA = 0.3290 \pm (0.0397)C042 - 11.5711 \pm (2.2390)G1e + 4.2849 \pm (1.0252)G3s + 0.4577 \pm (0.0973)GGI9 + 0.0382 \pm (0.0075)Mor030u - 0.0118 \pm (0.0021)RDF060m - 13.8056 \pm (1.3612){}^5\chi^A - 0.0660 \pm (0.0163)Mor12m + 3.2607$
 $N=51, R^2 = 0.8629, R^2_A = 0.8368, F= 33.056$

The observed activity has been plotted against estimated activity and is given in figure 1. The predictive power of the model has come out to be 0.8634. Further confirmation is obtained by estimating the log BA values of the compounds using model- 16 (Table 5). The predicted log BA values are in good agreement with the observed values showing that the proposed model is best suited for estimating log BA values of present set of compounds. On the basis of cross validated parameters (Table 6) an eight parametric model discussed above is found to be the best model for which the cross validated R^2 comes out to be 0.8413 (table 7).

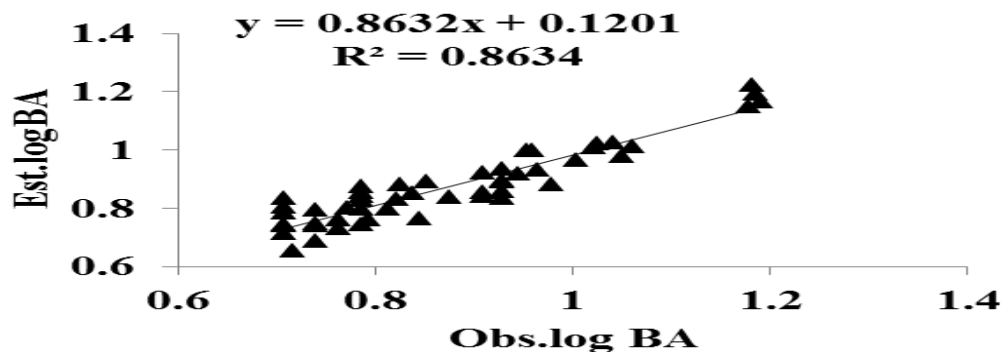


Fig. 1. Correlation between observed and estimated biological activity values using model 16

Table3. Correlation matrix

	log BA	${}^5\chi^A$	GGI9	RDF060m	Mor03u	Mor12m	G1e	G3s	C-042
log BA	1.0000								
${}^5\chi^A$	-0.4412	1.0000							
GGI9	0.0175	-0.1217	1.0000						
RDF060m	-0.0853	0.5060	0.5282	1.0000					
Mor03u	-0.1878	0.5024	-0.5173	-0.5412	1.0000				
Mor12m	-0.3731	0.0304	0.2008	0.2718	0.1970	1.0000			
G1e	0.0582	0.2963	-0.7289	-0.6749	0.7402	-0.2019	1.0000		
G3s	0.2084	0.3149	-0.6602	-0.5352	0.6064	-0.1963	0.7529	1.0000	
C-042	0.6322	-0.1179	-0.2056	-0.1132	-0.1676	-0.4522	0.2781	0.3195	1.0000

Table 4. Regression parameters and quality of correlation

Model No.	Parameters Used	$A_i=(1\dots\dots\dots 8)$	B	R^2	R^2_A	F-ratio
1	C042	0.3322±(0.0570)	0.8518	0.3996	0.3879	33.948
2	C042 ${}^5\chi^A$	0.3091±(0.0510) -6.6612±(1.7381)	1.3378	0.5360	0.5174	28.874
3	C042 RDF060m ${}^5\chi^A$	0.2824±(0.0491) -0.0070±(0.0026) -9.3957±(1,9268)	1.6285	0.5959	0.5712	24.087
4	C042 GGI9 RDF060m ${}^5\chi^A$	0.3023±(0.0455) 0.3586±(0.1136) -0.0117±(0.0028) -10.3578±(1.7974)	1.6670	0.6654	0.6376	23.868
6	C042 G3s GGI9 RDF060m ${}^5\chi^A$	0.2605±(0.0440) 3.9121±(1.2590) 0.5670±(0.1242) -0.0106±(0.0026) -11.7228±(1.7117)	1.0233	0.7225	0.6929	24.469
6	C042 G1e G3s GGI9 RDF060m ${}^5\chi^A$	0.2701±(0.0428) -4.6779±(2.2685) 5.0731±(1.3415) 0.4639±(0.1301) -0.0128±(0.0028) -11.7546±(1.6555)	1.5976	0.7459	0.7128	22.510
7	C042 G1e G3s GGI9 Mor030u RDF060m ${}^5\chi^A$	0.3491±(0.0493) -9.3655±(2.7204) 4.3274±(1.2836) 0.4521±(0.1217) 0.0233±(0.0085) -0.0140±(0.0026) -13.5003±(1.6732)	2.7357	0.7826	0.7488	23.138

8	C042 GIe G3s GGI9 Mor030u Mor12m RDF060m χ^2	0.3286±(0.0466) -10.8441±(2.5947) 4.1303±(1.2002) 0.4713±(0.1138) 0.0337±(0.0088) -0.0518±(0.0187) -0.0127±(0.0025) -14.1721±(1.5805)	3.1642	0.8148	0.7811	24.197
---	--	--	--------	--------	--------	--------

After the deletion of M-6 & M-26

Model No.	Parameters used	Ai=(1.....8)	B	R ²	R ² _A	F-ratio
9	C042	0.3335±(0.0551)	0.8504	0.4278	0.4162	36.640
10	C042 χ^2	0.3127±(0.0499) -6.1376±(1.7218)	1.2978	0.5476	0.5287	29.050
11	C042 RDF060m χ^2	0.2866±(0.0477) -0.0070±(0.0026) -8.9310±(1.9063)	1.5920	0.6105	0.5857	24.558
12	C042 GGI9 RDF060m χ^2	0.3046±(0.0445) 0.3352±(0.1115) -0.0114±(0.0028) -9.8689±(1.7889)	1.6310	0.6745	0.6462	23.832
13	C042 G3s GGI9 RDF060m χ^2	0.2608±(0.0420) 4.1278±(1.2059) 0.5524±(0.1188) -0.0102±(0.0025) -11.2795±(1.6629)	0.9498	0.7418	0.7131	25.851
14	C042 GIe G3s GGI9 RDF060m χ^2	0.2703±(0.0407) -4.6721±(2.1859) 5.2739±(1.2786) 0.4501±(0.1239) -0.0124±(0.0026) -11.2358±(1.6008)	1.5199	0.7660	0.7341	24.012
15	C042 GIe G3s GGI9 Mor030u RDF060m χ^2	0.3540±(0.0457) -9.6499±(2.5484) 4.4989±(1.1925) 0.4364±(0.1132) 0.0247±(0.0079) -0.0136±(0.0024) -13.0431±(1.5701)	2.7216	0.8096	0.7787	26.127
16	C042 GIe G3s GGI9 Mor030u Mor12m RDF060m χ^2	0.3290±(0.0397) -11.5711±(2.2390) 4.2849±(1.0252) 0.4577±(0.0973) 0.0382±(0.0075) -0.0660±(0.0163) -0.0118±(0.0021) -13.8056±(1.3612)	3.2607	0.8629	0.8368	33.056

Table 5. Observed and estimated values of biological activity using model no.16

Compd.no.	Actual Biological	Predicted Biological	Residual Biological
-----------	-------------------	----------------------	---------------------

	Activity(log BA)	Activity(log BA)	Activity(log BA)
1	0.785	0.844	-0.059
2	0.908	0.845	0.064
3	0.708	0.742	-0.034
4	0.061	1.014	0.047
5	0.708		
6	0.785	0.829	-0.043
7	0.763	0.762	0.002
8	1.049	0.981	0.068
9	0.785	0.858	-0.073
10	0.908	0.856	0.053
11	0.944	0.919	0.026
12	0.851	0.895	-0.044
13	0.929	0.894	0.035
14	0.978	0.883	0.095
15	0.716	0.657	0.059
16	0.785	0.746	0.039
17	1.19	1.167	0.023
18	0.708	0.788	-0.08
19	0.929	0.937	-0.007
20	0.792	0.762	0.03
21	1.061		
22	0.708	0.836	-0.128
23	0.813	0.801	0.012
24	0.785	0.878	-0.093
25	0.845	0.767	0.078
26	0.954	1	-0.046
27	0.929	0.896	0.033
28	0.908	0.923	-0.014
29	0.929	0.859	0.071
30	0.959	1.002	-0.043
31	1.004	0.967	0.037
32	0.708	0.717	-0.01
33	0.785	0.8	-0.015
34	0.763	0.732	0.032
35	0.708	0.748	-0.04
36	0.74	0.749	-0.009
37	0.74	0.798	-0.058
38	0.929	0.838	0.091
39	0.771	0.802	-0.031
40	0.875	0.84	0.035
41	0.708	0.806	-0.098
42	0.74	0.688	0.053
43	0.74	0.744	-0.003
44	0.822	0.834	-0.012
45	0.838	0.854	-0.016
46	1.041	1.029	0.011
47	0.825	0.885	-0.059
48	1.022	1.009	0.013
49	1.025	1.025	0

50	1.182	1.224	-0.042
51	1.179	1.151	0.028
52	1.185	1.194	-0.009
53	0.964	0.933	0.031

Table 6. Cross validation parameters for proposed models after deletion of compound no.6 & 26

Model No	Parameters used	PRESS	SSY	PRESS/SSY	R ² CV	S _{PRESS}	PSE
9	C042	0.5484	0.4101	1.3372	-0.3372	0.1058	0.1037
10	C042, χ^A	0.4336	0.5249	0.8261	0.1739	0.0950	0.0922
11	C042, RDF060m, χ^A	0.3733	0.5852	0.6379	0.3621	0.0891	0.0856
12	C042, GGI9, RDF060m, χ^A	0.312	0.6465	0.4826	0.5174	0.0824	0.0782
13	C042, G3s, GGI9, RDF060m, χ^A	0.2475	0.711	0.3481	0.6519	0.0742	0.0697
14	C042, GIe, G3s, GGI9, RDF060m, χ^A	0.2242	0.7343	0.3053	0.6947	0.0714	0.0663
15	C042, GIe, G3s, GGI9, Mor030u, RDF060m, χ^A	0.1824	0.776	0.2351	0.7649	0.0651	0.0598
16	C042, GIe, G3s, GGI9, Mor030u, Mor12m, RDF060m, χ^A	0.1313	0.8271	0.1587	0.8413	0.0559	0.0507

Table 7. Ridge analysis for eight variable model (model no.16)

Independent variables	VIF	T	λ	K
X5A	1.85	0.54	3.82	1.00
GGI9	2.45	0.41	1.67	2.29
RDF060m	2.47	0.40	0.95	4.04
Mor03u	5.20	0.19	0.57	6.69
Mor12m	1.67	0.60	0.36	10.63
GIe	6.28	0.16	0.29	13.07
G3s	2.84	0.3	0.24	15.81
C_042	1.87	0.54	0.09	40.59

VIF = Variance inflation factor

T = Tolerance

 λ = Eigen value

Lowest value of PSE, S_{PRESS} and PRESS/SSY further confirmed the findings. To be a reasonable and significant QSAR model, the ratio PRESS/SSY should be less than 0.4 and the value of this ratio, smaller than 0.1 indicates an excellent model. From Table 7 it is clear that the eight parametric model (model no. 16) have PRESS/SSY less than 0.4 indicating that this model has good predicting capacity. R²_{CV} is the cross validated squared correlation coefficient. The highest R²_{CV} value for model 16 further confirms the predictions. Uncertainty in prediction (S_{PRESS}) and predictive squared error were also calculated. There is no colinearity among the used parameters which has been established by ridge analysis as well as various inflation factors calculated from the model 16 (Table 7, figures 2 and 3). The parameters whose VIF value is greater than 10 will show colinearity. Similarly if

λ (eigen values) is found to be greater than 5, then the model will suffer from colinearity. Another test for colinearity is condition no. (k). if its value is found to be greater than 100, then colinearity exists. Tolerance value (T) equal to 1 or less indicates absence of colinearity. Table 7 indicates that all the above mentioned parameters or models discussed in the study are free from multicollinearity.

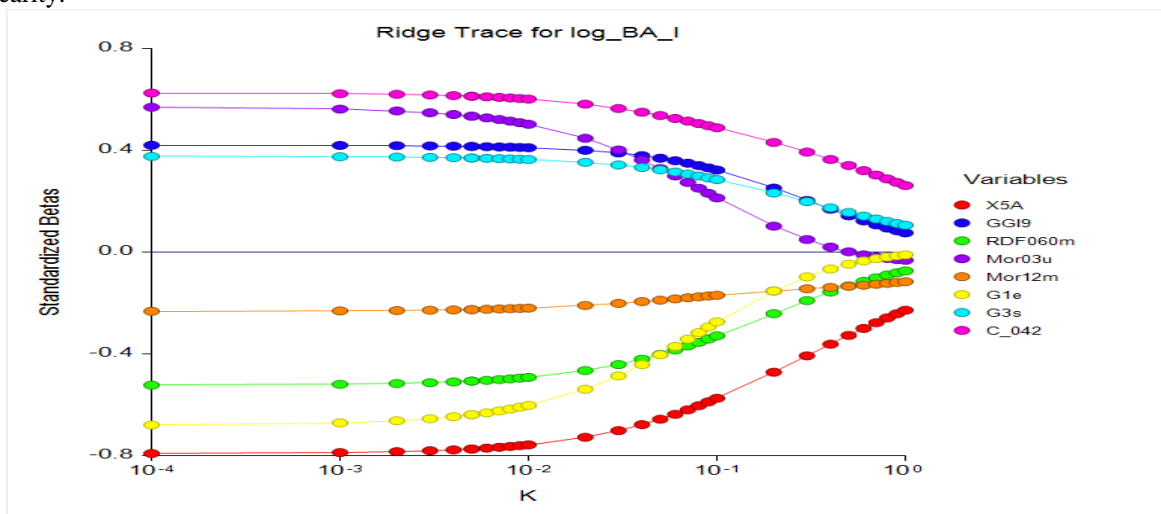


Fig. 2. Ridge Trace for eight variable model(model 16)

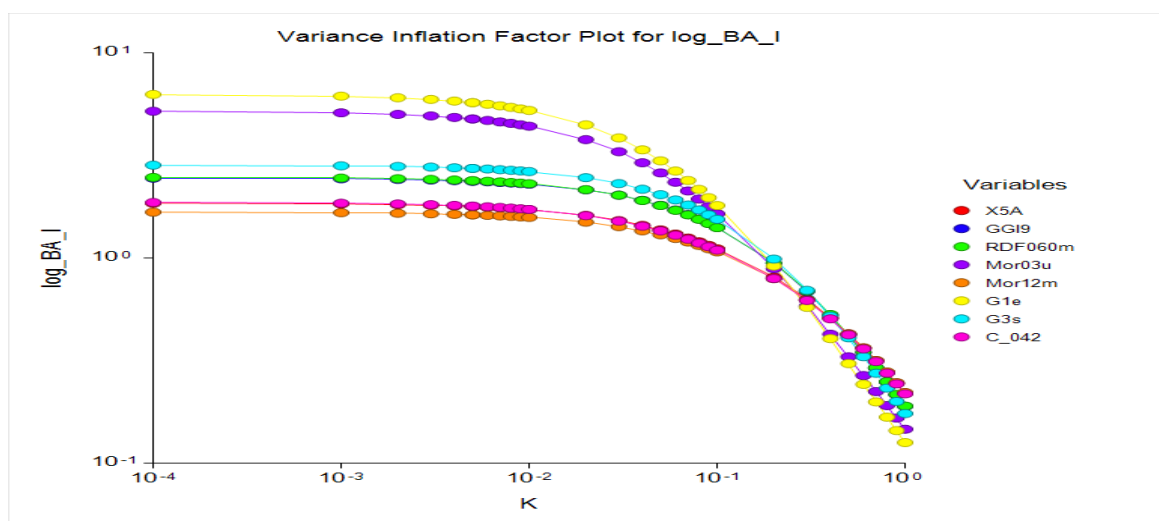


Fig. 3. VIF plot for eight variable model(model 16)

Conclusions

On the basis of above discussion following conclusions can be drawn.

1. C042 is a parameter which can model the anti-bacterial activity of present set compounds along with GGI9, G3s, Mor03u, Mor12m, RDF060m, G1e and $^5\chi^A$.
2. Coefficients for C042, GGI9 and G3s are positive suggesting that a higher value of these parameters will favour the log biological activity.
3. Negative coefficients for Mor12m, RDF060m, G1e and $^5\chi^A$ suggests that it has retarding effect towards the log biological values, hence in future designing of potent compounds their lower values will give better results.

Acknowledgement

The authors are very thankful to Prof. V.K.Agrawal, Director, NITTTR, Bhopal for his kind guidance and help.

REFERENCES

1. Vidal Reference Book, Medicines in Russia, N. B. Nikolaeva, B. R. Al_perovich, and V. N. Sozinov (eds.), Moscow: AstraFarmServis 1996 (In Russian).
2. M. Negwer, H-G. Scharnow, Organic-chemical drugs and their synonyms, Wiley-VCH, Weinheim 2002.
3. H. Yuan, A. L. Parrill, J. Mol. Struct.-Theochem. 2000, 1-3, 273 – 282.
4. R. B. Lesyk, B. S. Zimenkovsky, Curr. Org. Chem. 2004, 8, 1547 – 1577.
5. R. B. Lesyk, B. S. Zimenkovsky, R. V. Kutsyk, D. V. Atamanyuk, H. M. Semenciv, Pharm. J. 2003, 2, 52 – 56 (in Ukrainian).
6. V. E. Kuz_min, A. G. Artemenko, P. G. Polischuk, J. Mol.Mod. 2005, 11, 457 – 467.
7. V. E. Kuz_min, A. G. Artemenko, E. N. Muratov, I. L. Volineckaya, V. A. Makarov, O. B. Riabova, P. Wutzler, M. Schmidtke, J. Med. Chem. 2007, 17, 4205 – 4213.
8. Oleg, A. Costescu, M.V. Diudea, B. Parv, (2006), QSAR modeling of antifungal activity of some heterocyclic compounds, CROATICA CHEMICA ACTA, 79, (3), 483, 17-20.
9. Anatoliy G. Artemenko*, Eugene N. Muratova, b, Dmytro V. Atamanyukc, Victor E. Kuz_mina, Alexander I. Hromova, Roman V. Kutsyk d and Roman B. Lesykc QSAR Analysis of Antimicrobial Activity of 4-thiazolidone Derivatives QSAR Comb. Sci. 28, 2009, No. 2, 194 – 205.
10. NCSS, Kaysville Utah, www.ncss.com
11. S. Chatterjee, A.S. Hadi, B Price, (2000), Regression Analysis by Examples, 3rd Ed. Wiley: New York.
12. L.B. Kier, L.H. Hall, Molecular Connectivity in Structure-Activity Analysis, Research Studies Press - Wiley, Chichester (UK), 1986.
13. J. Gálvez, R. García, M.T. Salabert, R. Soler, J.Chem.Inf.Comput.Sci. 1994, 34, 520-525.
14. Hemmer, M.C., Steinhauer, V. & Gasteiger, J. Vibrat. Spect., (1999), 19, 151 -164.
15. J.H. Schuur, P. Selzer, J. Gasteiger, J. Am. Chem. Soc. 1996, 36, 334-344.
16. R. Todeschini, M. Lasagni, E. Marengo, J. Chemom. 1994, 8, 263-273; R. Todeschini, P. Gramatica, 3D QSAR in Drug Design - Vol. 2, H. Kubinyi, G. Folkers, Y.C. Martin (Eds.), Kluwer/ESCOM, Dordrecht (The Netherlands), 1998, 355-380.
17. V.N. Viswanadhan, A.K. Ghose, G.R. Revankar, R.K. Robins, J.Chem.Inf.Comput.Sci. 1989, 29, 163-172.