Available online at www.derpharmachemica.com

ISSN 0975-413X
CODEN (USA): PCHHAX

Der Pharma Chemica, 2016, 8(6):57-65
(http://derpharmachemica.com/archive.html)

Modeling of Sulmazole analogues as cardiotonic agents

Teena Pareek[a],*, Amrit V. Bajaj[a] and Dheeraj Mandloi[b]

[a]*School of Chemical Sciences, Devi Ahilya Vishwavidyalaya, Indore, M.P., 452001, India*
[b]*Institute of Engineering and Technology, Devi Ahilya Vishwavidyalaya, Indore, M.P., 452001, India*

ABSTRACT

Quantitative Structure-Activity Relationship (QSAR) studies on a series of 16 Sulmazole analogues have been reported in this paper. The study has been made utilizing Topological indices and Connectivity indices. The results have been critically discussed using different statistical parameters. Six estimated models were obtained, out of which the bi-parametric model has been considered to be the best. The robustness of the best estimated model has been evaluated using ridge regression which indicated that multicollinearity is not a problem in this model. Our results indicate that with the decrease in the value of third order valency connectivity index and with the increase in the value of Balaban index G, cardiotonic activity [log (I/IC$_{50}$)] increases. The values of statistical parameters obtained are: R^2 = 0.7049, R^2A = 0.6312, CV = 0.0460, F-ratio = 9.556

Keywords: QSAR, Topological indices, Connectivity indices, Correlation regression analysis.

INTRODUCTION

Quantitative Structure-Activity Relationship (QSAR) is the study of how the physico-chemical properties of compounds affect their biological activity. The compounds taken under consideration in the present paper are Sulmazole analogues [1-5]. Sulmazole are the cardiotonic agents which posses contractile properties due to which they increase the sensitivity of the myofibrils to Ca++ [6]. A number of topological indices [7-17] have been determined for these compounds such as Weiner index [16], Platt's index [18-21], Schultz molecular topological index [22,23], Balaban indices (J, F and G) [24], Zagreb's group indices (M_1 & M_2) [25], Randic's connectivity indices (^0X, ^1X, ^2X and ^3X) [26-29] and Kier and Hall's valency connectivity indices ($^0X^v$, $^1X^v$, $^2X^v$ and $^3X^v$) [30,31].

A series of 16 sulmazole analogues has been taken for consideration. All the above topological indices have been obtained for these compounds. A number of estimated models have been obtained based on the multiple regression [32,33] studies. But after evaluation, the bivariate model has been considered to be the best estimated model as the ridge regression statistics [34-36] for this model shows that multicollinearity is not a problem in this model. Further details are discussed in results and discussion section of this paper.

MATERIALS AND METHODS

The methodology used in the present investigation is QSAR. It is of great importance in modern chemistry and biochemistry. To obtain a significant correlation, it is essential that appropriate descriptors are employed, whether they are theoretical, empirical or derived from readily available experimental characteristics of structures.

QSAR methodology has been helpful in correlating the structure of a large series of sulmazole analogues with their cardiotonic activity [log(I/IC$_{50}$)]. Mathematical models have been formed that correlate molecular structure to the activity. Molecular structure has been encoded through the generation of the descriptions which are numerical values corresponding to topological and connectivity indices.

Computational Section

1. The Cardiotonic Activity:

The Cardiotonic activity of the series of Sulmazole analogues is taken from the work of Khadikar et al [1-5].

2. Topological descriptors:

The structure of compounds is drawn by using ACD-labs Chem-sketch Software [37]. The various topological descriptors used in the present study were calculated from the Hydrogen Suppressed molecular graphs of Sulmazole analogues by using Dragon Software [38].

3. Regression Analysis:

The correlation-regression analysis of the data was done by using NCSS-8 Software [39] as well as Origin-6 software.

RESULTS AND DISCUSSION

In this section, we have done modeling of Sulmazole analogues as cardiotonic agents. The biological activities in the form of $\log(I/IC_{50})$, along with different substituents on the parent compound (Figure-1), are given in Table-1. The structural details of these compounds are given in Table-2. The descriptors used for obtaining statistically significant models, which were calculated using DRAGON software, are summarized in Table-3 and 4; as many as 17 descriptors have been used. Out of these 17 descriptors, the selection of significant descriptors was done using variable selection for multivariate regression technique mentioned in NCSS software. It was shown that the descriptors W, P, M, J, G, T, $^{0}X^{v}$, $^{1}X^{v}$, $^{2}X^{v}$ and $^{3}X^{v}$ were useful descriptors for obtaining statistically significant models. The maximum number of descriptors is three, which we can use in obtaining appropriate model based on rule of thumb. But NCSS results show that the best model can be obtained using maximum two descriptors as the value of R-squared change is the highest in this model. Thus the bi-variate model obtained is:

Bi-variate model

$\log(1/IC_{50}) = 3.6253 - 1.1287 * {}^{0}\chi^{v} + 0.1326 * G$

$R^2 = 0.2787$, Adj $R^2 = 0.1678$, Coefficient of Variation = 0.1363, F-ratio = 2.512

However, this model does not show the appropriate value of R^2. It implies that both these parameters viz., Zero-order valency connectivity index and Balaban Index G are not good for obtaining statistically significant model. However, comparison of estimated activity with the experimental one (Table-6) indicates that there are five compounds viz. 1, 2, 8, 9 and 14, having abnormal residues; therefore, these compounds have been considered as outliers. After deleting these outliers, the selection results obtained are presented in Table-7 and the best estimated models are given in Table-8.

From the perusal of Table-8, it is evident that, as we go on adding a variable, the values of R^2 as well as R^2A go on increasing. However, according to the rule of thumb, the model size should not exceed 2 or 3. To confirm the robustness of these models, we have considered Ridge regression analysis. The results thereof (Table-10 and Table-11) indicate that bi-parametric model is the best model as VIF values and Condition Numbers are well within the range, and hence, multicollinearity is not a problem.

We have also plotted a graph between actual and predicted values of the activity [log (I/IC50)] [Table-9], based on the best model, which also shows that the model obtained is statistically significant.

Best model:

$\log(1/IC_{50}) = 2.6891 - 0.5549 * {}^{3}X^{v} + 3.6782 * G$

$R^2 = 0.7049$, $R^2A = 0.6312$, CV = 0.0460, F-ratio = 9.556

This equation shows that, after deletion of outliers, the quality of regression model has improved to a much greater extent, as the value of R^2 has increased from 0.2787 to 0.7049. The value of $R^2 = 0.7049$ indicates that the model explains 70.49% variation in the activity.

Figure 1: Parent structure of Sulmazole

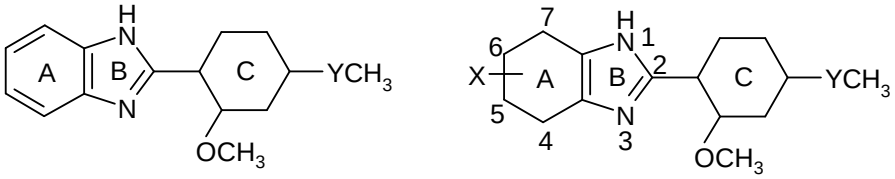
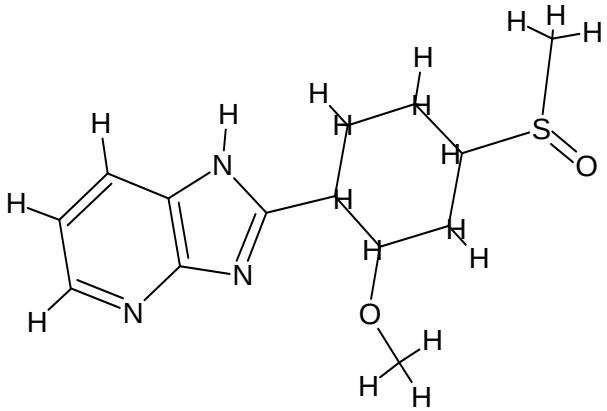


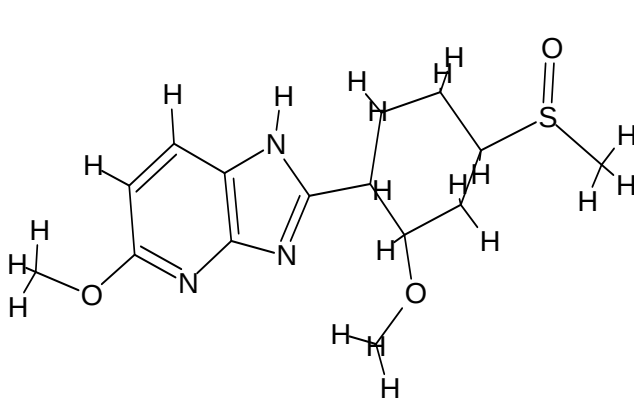
Table-1: Different Substituents on the Parent Compound and Their log $1/IC_{50}$ Values

Compound No.	X	Y	log I/IC_{50}
1	H	S(O)	4.70
2	5-OCH_3	S(O)	3.00
3	6-OCH_3	S(O)	3.63
4	7-OCH_3	S(O)	4.10
5	6-NH_2	O	4.64
6	6-NH_2	S(O)	3.91
7	6-Cl	S(O)	3.76
8	6-CH_3	S(O)	3.00
9	6-OH	S(O)	4.80
10	H	O	4.01
11	H	S(O)	3.61
12	4-OCH_3	S(O)	3.73
13	6-OCH_3	S(O)	3.75
14	7-OCH_3	S(O)	5.08
15	7-NH_2	S(O)	3.72
16	H	O	4.01

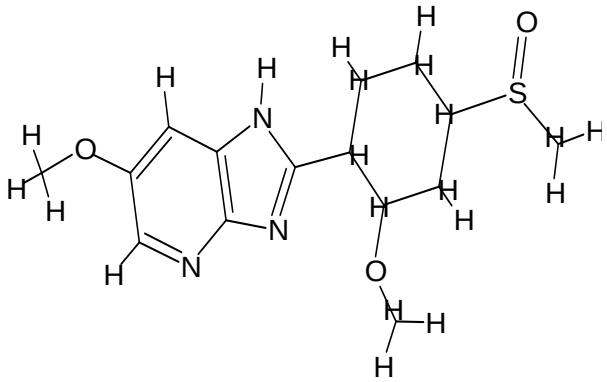
Table-2: Structural Details of Sulmazole Analogues



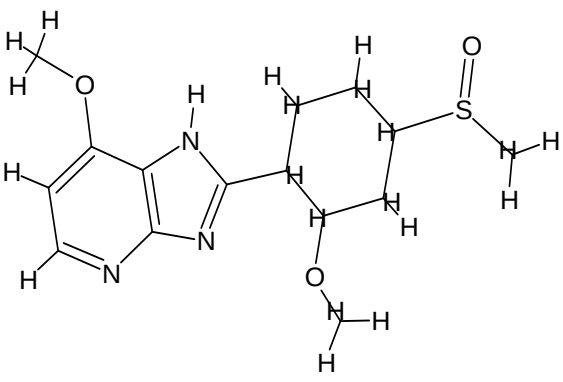
Compound No. 1



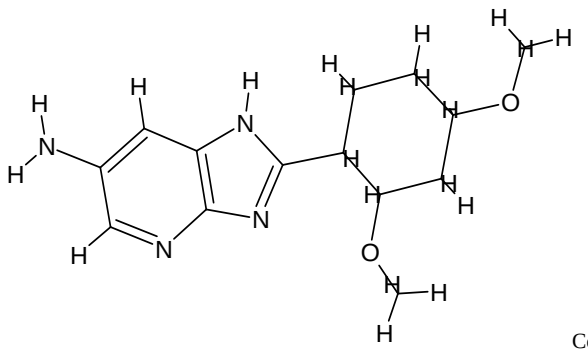
Compound No. 2



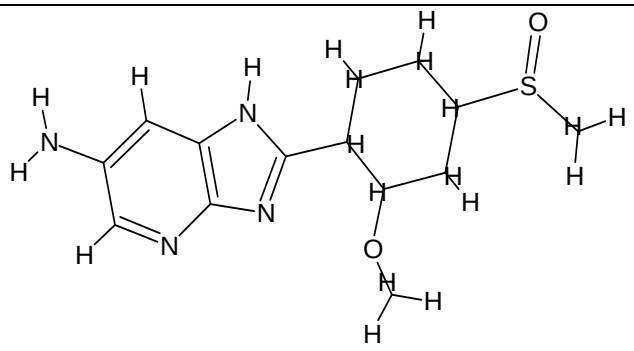
Compound No. 3



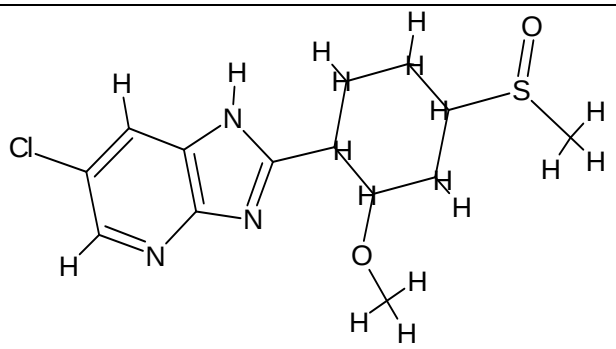
Compound No. 4



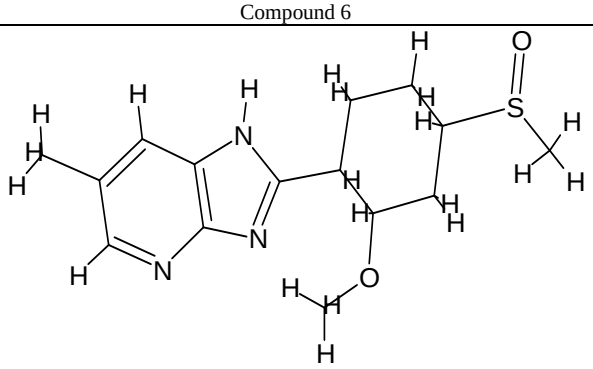
Compound 5



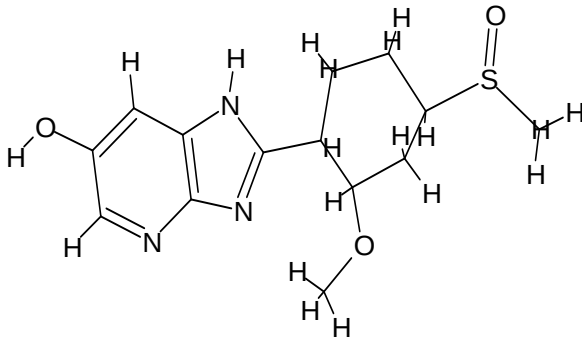
Compound 6



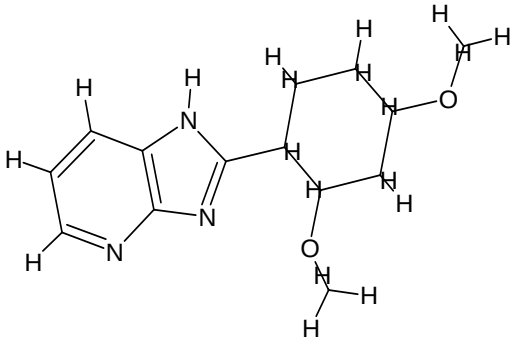
Compound 7



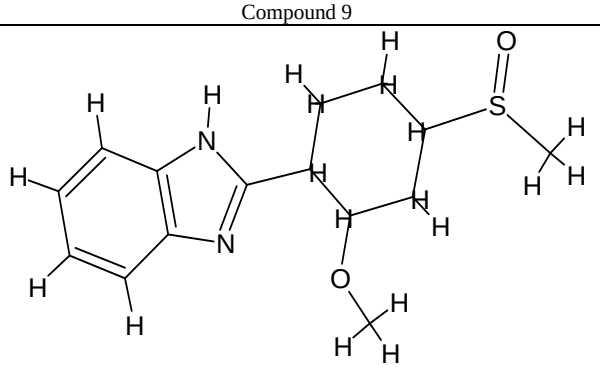
Compound 8



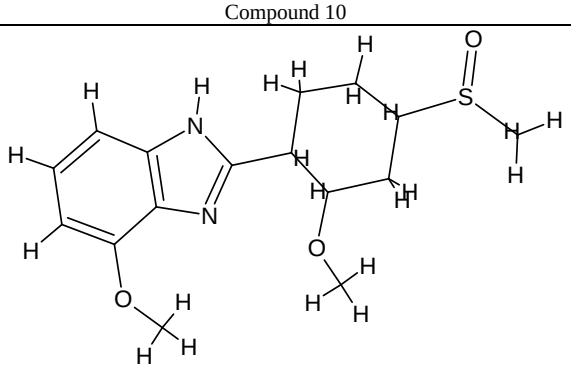
Compound 9



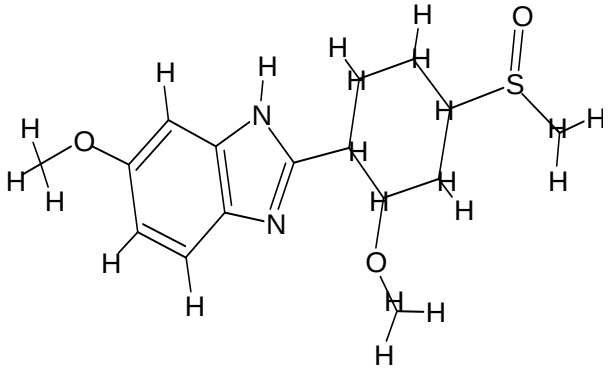
Compound 10

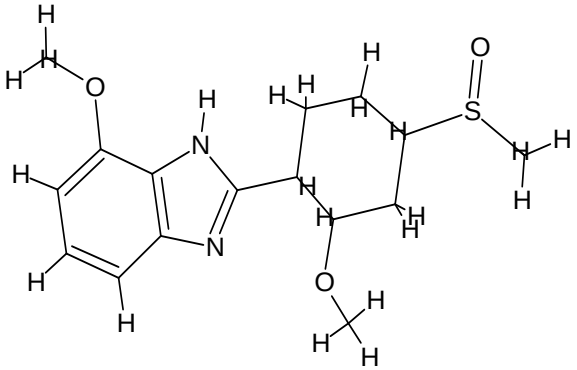


Compound 11

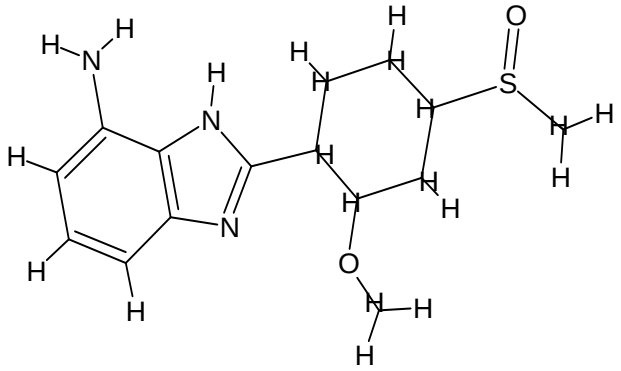


Compound 12

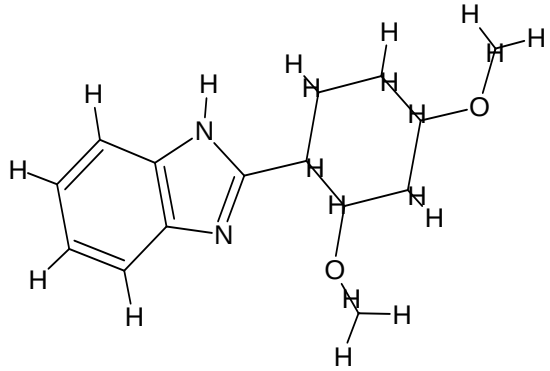




Compound 13



Compound 14



Compound 15

Compound 16

Table-3: Calculated Topological Indices of Sulmazole Analogues

S.No.	log (I/IC$_{50}$)	W	P	M	J	F	G	T	M1	M2
1	4.70	806	31	3512	1.617	6.468	107.8	44	106	126
2	3.00	1073	35	4582	1.601	6.404	119.2129	48	116	138
3	3.63	1073	35	4582	1.601	6.404	119.2129	48	116	138
4	4.10	1045	36	4462	1.642	6.568	122.2658	48	116	139
5	4.64	809	31	3512	1.614	6.456	107.6	44	106	126
6	3.91	929	33	4004	1.615	6.46	113.9544	46	112	133
7	3.76	929	33	4004	1.615	6.46	113.9544	46	112	133
8	3.00	929	33	4004	1.615	6.46	113.9544	46	112	133
9	4.80	929	33	4004	1.615	6.46	113.9544	46	112	133
10	4.01	697	29	3063	1.616	6.464	101.4567	42	100	119
11	3.61	806	31	3512	1.617	6.468	107.8	44	106	126
12	3.73	1045	36	4462	1.642	6.568	122.2658	48	116	139
13	3.75	1073	35	4582	1.601	6.404	119.2129	48	116	138
14	5.08	1045	36	4462	1.642	6.568	122.2658	48	116	139
15	3.72	915	34	3944	1.639	6.556	115.6478	46	112	134
16	4.01	697	29	3063	1.616	6.464	101.4567	42	100	119

Where W= Weiner index, P= Platt's Index, M or SMTI= Schultz Molecular Topological Index,J, F & G= Balaban Indices, T= Topological Index, M_1 & M_2= Zagreb Group Indices

Table-4: Calculated Connectivity Indices of the Compounds Under Study

S.No.	log (I/IC$_{50}$)	^0X	^1X	^2X	^3X	$^0X^v$	$^1X^v$	$^2X^v$	$^3X^v$
1	4.70	14.113	9.686	8.685	7.432	12.521	8.465	6.975	5.51
2	3.00	15.69	10.618	9.488	8.191	13.852	8.998	7.325	5.782
3	3.63	15.69	10.618	9.488	8.191	13.852	8.998	7.347	5.8
4	4.10	15.69	10.635	9.404	8.22	13.852	8.994	7.312	5.826
5	4.64	14.113	9.707	8.589	7.47	11.796	6.877	5.313	4.108
6	3.91	14.983	10.08	9.319	7.768	13.021	8.665	7.24	5.611
7	3.76	14.983	10.08	9.319	7.768	13.578	8.943	7.561	5.763
8	3.00	14.983	10.08	9.319	7.768	13.444	8.876	7.484	5.726
9	4.80	14.983	10.08	9.319	7.768	12.891	8.599	7.165	5.575
10	4.01	13.242	9.313	7.956	7.135	11.296	6.677	5.048	4.007
11	3.61	14.113	9.686	8.685	7.432	12.651	8.605	7.117	5.642
12	3.73	15.69	10.635	9.404	8.22	13.982	9.134	7.457	5.928
13	3.75	15.69	10.618	9.488	8.191	13.982	9.128	7.483	5.941
14	5.08	15.69	10.635	9.404	8.22	13.982	9.134	7.455	5.952
15	3.72	14.983	10.097	9.213	7.906	13.151	8.811	7.342	5.803
16	4.01	13.242	9.313	7.956	7.135	11.426	6.817	5.191	4.139

Where ^0X= Zero-order Randic's connectivity index, ^1X= First-order Randic's connectivity index, ^2X= Second-order Randic's connectivity index, ^3X= Third-order Randic's connectivity Index,
$^0X^v$= Zero-order Kier & Hall's valency connectivity index, $^1X^v$= First-order Kier & Hall's valency connectivity index, $^2X^v$= Second-order Kier & Hall's valency connectivity index, $^3X^v$= Third-order Kier & Hall's valency connectivity index.

Table-5: Selection Results Section

Model Size	R-Squared	R-Squared Change	Coded Variables	Decoded Variables
1	0.113561	0.113561	D	J
2	0.278722	0.165161	EG	G, $^{0}\chi^{v}$
3	0.369405	0.090682	BEG	P,G, $^{0}\chi^{v}$
4	0.533518	0.164113	BEHJ	P,G, $^{1}\chi^{v}$, $^{3}\chi^{v}$
5	0.552641	0.019124	BEHIJ	P,G, $^{1}\chi^{v}$, $^{2}\chi^{v}$, $^{3}\chi^{v}$
6	0.562107	0.009465	BEGHIJ	P,G, $^{0}\chi^{v}$, $^{1}\chi^{v}$, $^{2}\chi^{v}$, $^{3}\chi^{v}$
7	0.582521	0.020414	ADEFGHJ	W,J,G,T, $^{0}\chi^{v}$, $^{1}\chi^{v}$, $^{3}\chi^{v}$
8	0.619245	0.036724	ACDEFGHJ	W,M,J,G,T, $^{0}\chi^{v}$, $^{1}\chi^{v}$, $^{3}\chi^{v}$
9	0.622339	0.003094	ACDEFGHIJ	W,M,J,G,T, $^{0}\chi^{v}$, $^{1}\chi^{v}$, $^{2}\chi^{v}$, $^{3}\chi^{v}$
10	0.624523	0.002184	ABCDEFGHIJ	W,P,M,J,G,T, $^{0}\chi^{v}$, $^{1}\chi^{v}$, $^{2}\chi^{v}$, $^{3}\chi^{v}$

Table-6: Residual Report based on Bi-variate Model

Row	Actual log(1/IC_{50})	Predicted log(1/IC_{50})	Residual
1	4.700	3.790	0.910
2	3.000	3.802	-0.802
3	3.630	3.802	-0.172
4	4.100	4.207	-0.107
5	4.640	4.582	0.058
6	3.910	4.042	-0.132
7	3.760	3.414	0.346
8	3.000	3.565	-0.565
9	4.800	4.189	0.611
10	4.010	4.332	-0.322
11	3.610	3.644	-0.034
12	3.730	4.060	-0.330
13	3.750	3.655	0.095
14	5.080	4.060	1.020
15	3.720	4.120	-0.400
16	4.010	4.185	-0.175

Table-7: Selection Results Section After Deleting Compound no. 1, 2, 8, 9 and 14

Model Size	Coded Variables	Decoded Variables	R-Squared	R-Squared Change
1	J	$^{3}X^{v}$	0.475634	0.475634
2	EJ	G, $^{3}X^{v}$	0.704930	0.229296
3	EHI	G, $^{1}X^{v}$, $^{2}X^{v}$	0.830926	0.125996
4	BEHI	P,G, $^{1}X^{v}$, $^{2}X^{v}$	0.869379	0.038453
5	BEHIJ	P,G, $^{1}X^{v}$, $^{2}X^{v}$, $^{3}X^{v}$	0.895950	0.026571
6	BEGHIJ	P,G, $^{0}X^{v}$, $^{1}X^{v}$, $^{2}X^{v}$, $^{3}X^{v}$	0.938962	0.043012
7	BDEFGIJ	P,J,G,T, $^{0}X^{v}$, $^{2}X^{v}$, $^{3}X^{v}$	0.976048	0.037086
8	BCDEFGIJ	P,M,J,G,T, $^{0}X^{v}$, $^{2}X^{v}$, $^{3}X^{v}$	0.991875	0.015826
9	ABDEFGHIJ	W,P,J,G,T, $^{0}X^{v}$, $^{1}X^{v}$, $^{2}X^{v}$, $^{3}X^{v}$	0.999494	0.007619
10	ABCDEFGHIJ	W,P,M,J,G,T, $^{0}X^{v}$, $^{1}X^{v}$, $^{2}X^{v}$, $^{3}X^{v}$	1.000000	0.000506

Table-8: Best Estimated Models

Model Size	Selected Variables	R^2	R^2A	CV	F-ratio	Model Log (1/IC_{50})
1	$^{3}X^{v}$	0.4756	0.4174	0.0578	8.164	5.2479 - 0.2537686 * $^{3}X^{v}$
2	G, $^{3}X^{v}$	0.7049	0.6312	0.0460	9.556	2.6891 - 0.5549 * $^{3}X^{v}$ + 3.6782 * G
3	G, $^{1}X^{v}$, $^{2}X^{v}$	0.8309	0.7585	0.0372	11.467	3.7888 - 1.9636 * $^{1}X^{v}$ + 1.2918 * $^{2}X^{v}$ + 6.8308 * G
4	P,G, $^{1}X^{v}$, $^{2}X^{v}$	0.8694	0.7823	0.0354	9.984	0.4928 - 1.7336 * $^{1}X^{v}$ + 1.0524 * $^{2}X^{v}$ - 0.5376 * P + 0.2512 * G
5	P,G, $^{1}X^{v}$, $^{2}X^{v}$, $^{3}X^{v}$	0.8959	0.7919	0.0346	8.611	-0.8860 - 2.9777 * $^{1}X^{v}$ + 0.9292 * $^{2}X^{v}$ + 1.5590 * $^{3}X^{v}$ - 0.9522 * P + 0.4095 * G
6	P,G, $^{0}X^{v}$, $^{1}X^{v}$, $^{2}X^{v}$, $^{3}X^{v}$	0.9390	0.8474	0.0296	10.255	-2.4180 + 1.4027 * $^{0}X^{v}$ -9.1765 * $^{1}X^{v}$ + 1.4038 * $^{2}X^{v}$ + 7.1442 * $^{3}X^{v}$ – 1.7449 * P + 0.6581 * G

Table-9: Predicted Values with Confidence Limits of Means on the Basis of Best Model

Row	Actual C3	Predicted C3	Standard Error of Predicted
1	3.630	3.856	0.071
2	4.100	3.954	0.098
3	4.640	4.367	0.120
4	3.910	3.767	0.062
5	3.760	3.683	0.075
6	4.010	4.198	0.108
7	3.610	3.524	0.131
8	3.730	3.897	0.092
9	3.750	3.778	0.071
10	3.720	3.723	0.067
11	4.010	4.124	0.104

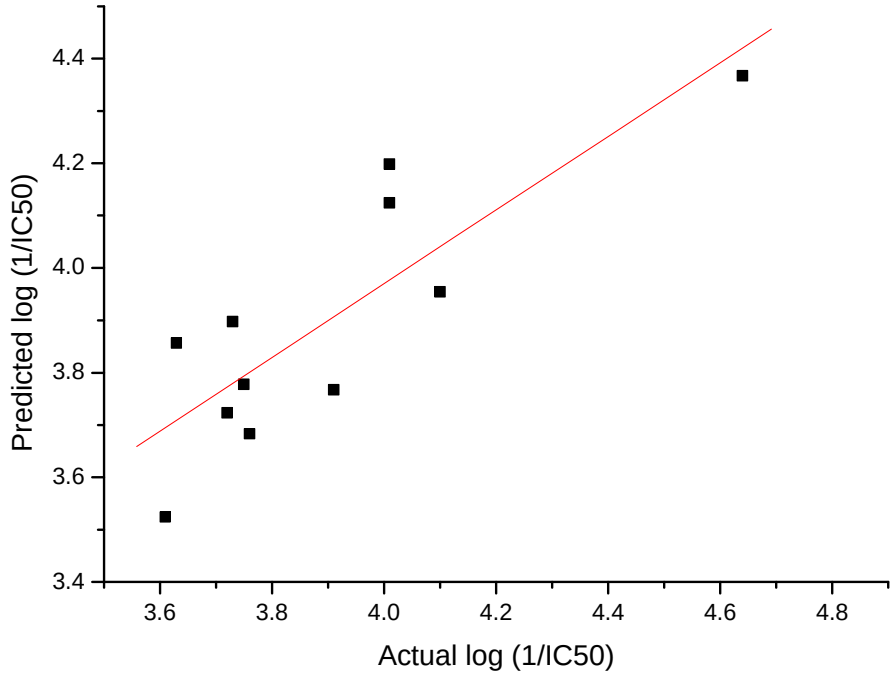


Figure 2 Plot of actual vs predicted log(1/IC50) values based on best model

Y = 1.15226 + 0.70437 * X
R = 0.83936, SD = 0.142, N = 11, P = 0.00123

Table-10: Least Squares Multicollinearity Section

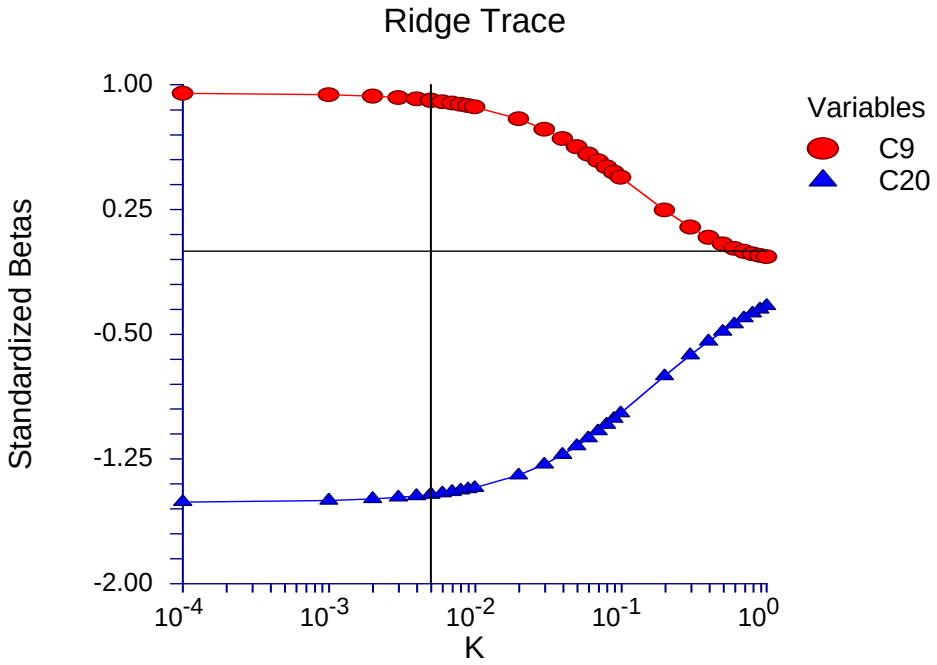
Independent Variable	Variance Inflation	R-Squared Vs Other X's	Tolerance
C9	3.9244	0.7452	0.2548
C20	3.9244	0.7452	0.2548

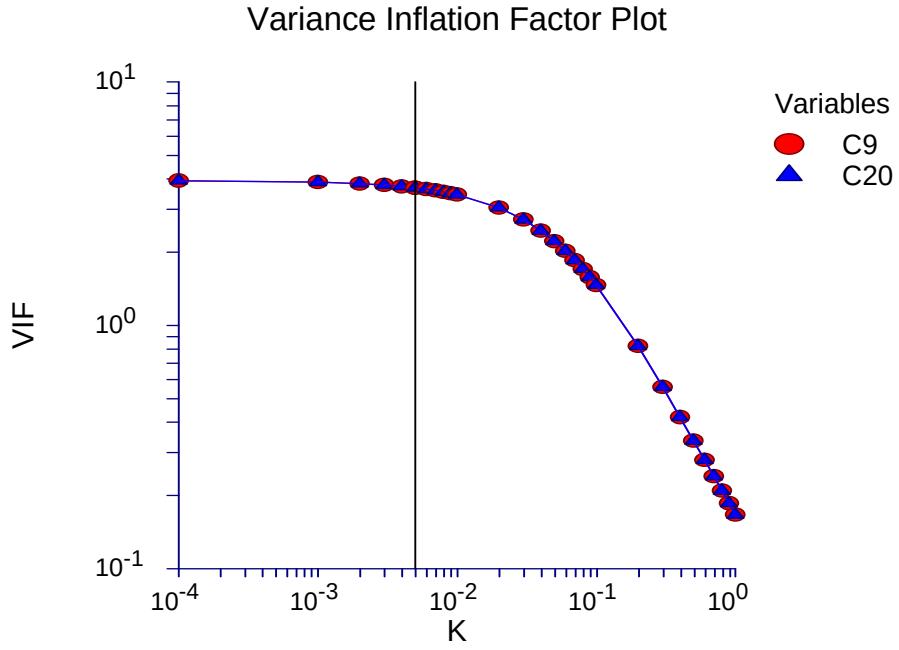
Since all VIF's are less than 10, multicollinearity is NOT a problem.

Table-11: Eigen Values of Correlations

No.	Eigen value	Incremental Percent	Cumulative Percent	Condition Number
1	1.863241	93.16	93.16	1.00
2	0.136759	6.84	100.00	13.62

All Condition Numbers less than 100. Multicollinearity is NOT a problem.





CONCLUSION

The aforementioned results and discussion lead us to conclude that in modeling of Cardiotonic Activity of Sulmazole analogues, Balaban Index G and Kier & Hall's third-order Valency Connectivity Index $^{3}\chi^{v}$ play a dominating role and yield excellent bi-parametric model.

Acknowledgement
Authors are thankful to Late Dr. P.V. Khadikar for his guidance and encouragement. Authors are also thankful to Vice-chancellor and Head of School of Chemical Sciences, DAVV, Indore, for providing facilities to carry out research work. We are also thankful to the referees for providing useful and fruitful suggestions.

REFERENCES

[1] S.L. Rabinowitz, R.M. Myerson, Topics in Medicinal Chemistry, Interscience Pub., John Wiley & Sons, NY, **1967**.
[2] M.V. Diudea, M.S. Floresar, P.V. Khadikar, Molecular Topology and its Applications, EFICON, Bucharest, **2006**.
[3] W.C. Bowman, M.J. Rand, Text Book of Pharmacology, 2nd Ed., Blackwell Science Publication, London, **1980**.
[4] M.L. Burger, Medicinal Chemist, 4th ed., John Wiley & Sons, NY, **1981**.
[5] D.G. Allen, J.R. Blinks, *Nature*, **1978**, 273, 509.
[6] Phosphodiesterase-inhibiting properties of newer inotropic agents. H. Scholz, W. Meyer, *Circulation*, **1986** [Pubmed].
[7] T. Pareek, A.V. Bajaj, D. Mandloi, *Int. J. Sci. Engg. Res.*, **2014**, 5, 159.
[8] T. Pareek, A.V. Bajaj, D. Mandloi, *Int. J. Lat. Tech. Engg. Man. App. Sci.* **2014**, 3 (4A), 32
[9] V. K. Agrawal, R. Sharma, P.V. Khadikar, *Bioorg. Med. Chem.*, **2002,** 10, 2993.
[10] A. Saxena, P.V. Khadikar, *Acta Pharm.*, **1999** 49, 171.
[11] S. Karmarkar, P. V. Khadikar, M. Mandloi, S. Joshi, V. K. Agrawal, *Indian J. Chem.*, **2001,** 40A, 12.
[12] D. Mandloi, S. Joshi, P.V. Khadikar, N. Khosla, *Bioorg. Med. Chem. Lett.*, **2005**, 15, 405.
[13] V. K. Agrawal, J. Singh, P. V. Khadikar, C. T. Supuran, *Bioorg. Med. Chem. Lett.*, **2006**, 16, 2044.
[14] S. Karmarkar, S. Joshi, A. Das,P. V. Khadikar, *J. Serb. Chem. Soc.*, **1997**, 62, 227.
[15] J. Singh, M. Lakhwani, P. V. Khadikar, V. K. Agrawal, A. T. Balaban, W. B. Clare, *Romania Rev*., **2006**, 14, 103.
[16] R. Todeschini, V. Consonni, *Handbook of Molecular Descriptors*, Wiley- VCH: Weinheim (GER), **2000**.
[17] M. Karelson, *Molecular descriptors in QSAR/QSPR*, Wiley-Interscience, New York, **2000**.
[18] H. Wiener, *J. Am. Chem. Soc.*, **1947**, 69, 17.
[19] L. Pogliani, *Amino Acids*, **1994**, 6, 141.
[20] L. Pogliani, *J. Phys. Chem.*, **1996**, 100, 18065.
[21] L. Pogliani, *Chem. Rev.*, **2000** 100, 3827.
[22] I. Gutman, *J. Chem. Inf. Comput. Sci.*, **1994**, 34, 1087.
[23] S. Klavzar, *J. Chem. Inf. Comput. Sci.*, **1996**, 36, 1001.
[24] A. T. Balaban, *Chem. Phys. Lett.*, **1982**, 89, 399.
[25] I. Gutman, Ruscic, N. Trinajstic, C. F. Wilkox, Jr., *J. Chem. Phys.*, **1975**, 62, 3339.
[26] M. Randic, *Acta Chem. Slov.*, **1998**, 45, 239.
[27] M. Randic, *J. Chem. Inf. Comput. Sci.*, **1997**, 37, 672.
[28] M. Randic, *J. Am. Chem. Soc.*, **1975**, 97, 6609.
[29] M. Randic, *J. Mol. Graph Model.*, **2001**, 20, 19.
[30] L. B. Kier, L. H. Hall, W. J. Murray, M. Randic, *J. Pharm. Sci.*, **1975**, 64, 1971.
[31] L. B. Kier, L. H. Hall, *J. Pharm. Sci.*, **1981**, 70, 583.
[32] J. Stevens, Applied Multivariate Statistics for the Social Sciences, Lawrence Erlbaum Associates Publishers, Hillsdale, New Jersey, **1986**.
[33] W. W. Cooley, P. R. Lohnes, Multivariate Data Analysis, Wiley, New York, **1971**.
[34] M. El-Dereny, N.I. Rashwan, *Int. J. Contemp. Math. Sciences*, **2011**, 12, 585.
[35] M. Sana, C. Eyup, *Int. J. Contemp. Math. Sciences*, **2008**, 3, 527.
[36] J. Singh, B. Shaik, S. Singh, V.K. Agrawal, P.V. Khadikar, O. Deeb, C.T. Supuran, *Chem. Biol. Drug Des.*, **2008**, 71, 244.
[37] Acd-Lab software for calculating the referred physicochemical parameters. Chemsketch. http://www.acdlabs.com/acdlabs-rss-feed.xml
[38] DRAGON Software for calculation of topological indices. http://www.disatunimib.it
[39] J. Hintze, NCSS 8 (**2012**) NCSS, LLC. Kaysville, Utah, USA. www.ncss.com