QSAR Analysis of N-alkyl imidazole Analogues as Antibacterial Agents

NAZNEEN KHAN[1], LOVE K. SONI[1], A. K. GUPTA[1], S. R. WAKODE[2], R. D. WAGH[3] AND S. G. KASKHEDIKAR[1]*

[1]Molecular Modeling Study Group, CADD Laboratory, Department of Pharmacy, Shri G. S. Institute of Technology and Science, 23 Park Road, Indore-452 003, [2]Delhi Institute of Pharmaceutical Sciences & Research, Pushp Vihar, New Delhi-110 017, [3]Institute of Pharmacy, Nagaon, Dhule-424 005, India.

A quantitative structure activity relationship study on a series of N-alkyl imidazole analogues was made using combination of various thermodynamic electronic and spatial descriptors. Several statistical expressions were developed using stepwise multiple liner regression analysis. The best quantitative structure activity relationship models were further validated by leave-one-out method of cross-validation. The study revealed that the electronic property, i.e., dipole moment contributed positively, and spatial descriptor (principal moment of inertia at Y axis) contributed negatively. The study suggested that substitution of group at R1 position on imidazole ring with hydrophobic nature and low bulkiness are favourable for the antibacterial activity in the concerned microbes. The quantitative structure activity relationship study provides important structural insights in designing of potent antibacterial agents.

The development of new and different antimicrobial agents has been a very important step[1]. Much of the research programme efforts are directed toward the design of new and available drugs, because of the unsatisfactory status of present drugs' side effects, and the acquisition of resistance by the infecting organisms to the present drugs[2-3]. The resistance of common pathogens to standard antibiotic therapy is rapidly becoming a major health problem throughout the world. The incidence of multidrug-resistant Gram-positive bacteria is increasing, and infections caused by *Staphylococcus aureus* (multidrug-resistant *S. aureus* and coagulase-negative *staphylococcus*), *enterococci*, and *pneumococci* are particularly problematic[4]. There is a real perceived need for the discovery of new compounds endowed with antibacterial property, possibly acting through mechanisms of action that are distinct from those of the well-known classes of antibacterial agents to which many clinically relevant pathogens are now resistant[5].

The *N*-alkyl derivatives of imidazole were identified as potential candidates for the antibacterial agents. *N*-substituted imidazoles exhibit a variety of valuable pharmacological properties such as antiparasitic, antifungal, and antibacterial activity. It has been reported that *N*-alkylimidazoles with the most simple structure posses inhibitory effect on microsomal oxidation and cytotoxic activity.

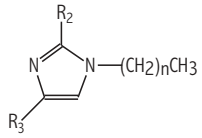
In the present paper, we describe the QSAR studies to investigate the relationship between the various physicochemical parameters and antibacterial activity of *N*-alkyl derivatives of imidazole that may be helpful in development of potent antibacterial agents.

MATERIALS AND METHODS

The antibacterial data of *N*-alkyl imidazole derivatives with activity on *S. aureus*, *E. coli*, *P. aeruginosa* were taken from the reported work of Khabnadideh *et al.*[6] (Table 1). The biological activity data MIC (minimum inhibitory concentration in mg/ml) were converted to negative logarithmic dose in moles (pMIC) for QSAR analysis. The correlations were sought between inhibitory activity and various substituent constants at position R_1 of molecule (fig. 1) and indicator variable for presence of methyl group at R_2, and presence of nitro group in the ring system at R_3 position. The values of substituent constants like hydrophobic (π), steric (Molar refractivity

***For correspondence**
E-mail: sgkaskhedikar@rediffmail.com

TABLE 1: ANTIBACTERIAL ACTIVITY OF *N*-ALKYL IMIDAZOLE ANALOGUES ON *S. AUREUS, P. AERUGINOSA, AND E. COLI*



R_2	R_3	n	Minimum inhibitory concentration (MIC) in µg/ml		
			S. aureus	*P. aeruginosa*	*E. coli*
H	H	1	480	960	480
H	H	2	44	550	110
H	H	3	87	248	99
H	H	4	69	138	97
H	H	5	46	122	91
H	H	6	33	83	50
H	H	7	27	54	45
H	H	8	10	39	19
H	H	9	208	416	208
H	H	10	1,110	2,220	1,110
H	H	11	14,160	16,520	14,160
H	H	12	17,500	20,000	18,750
H	H	13	19,800	23,760	21,120
CH3	H	5	25	-	25
CH3	H	11	17,500	-	17,500
CH3	NO2	3	17	-	18
CH3	NO2	7	24	-	31

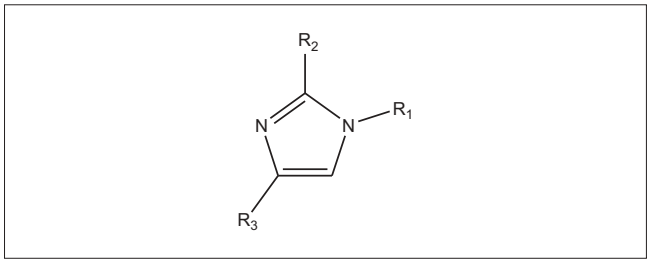


Fig. 1: Structure used for 2D-QSAR analysis.

or MR), hydrogen acceptor (HA), hydrogen donor (HD) and electronic (field effect or *F*, resonance effect or *R*, and Hammett's constant or σ) were taken from the literature[7]. The series was also subjected to molecular modelling via 3D-QSAR studies using CS Chem-Office 6.0[8] running on a P-III processor. Structures of all the compounds were sketched using builder module of the programme. These structures were then subjected to energy minimization using force field molecular mechanics-2 (MM2) until the root mean square (RMS) gradient value became smaller than 0.1 kcal/mol. Å. Minimized molecules were subjected to re-optimization via Austin model-1 (AM1)[9] method until the RMS gradient attained a value smaller than 0.0001 kcal/mol. Å using MOPAC. The geometry optimization of the lowest energy structure was carried out using Eigenvector following routine. The descriptor values for all the molecules were calculated using "compute properties" module of programme.

Calculated thermodynamic descriptors included critical temperature (T_c), ideal gas thermal capacity (C_p), critical pressure (P_c), boiling point (BP), Henry's law constant (H), bend energy (E_b), heat of formation (H_f), total energy (TE), and partition coefficient (PC).

Steric descriptors derived were connolly accessible area (CAA), connolly molecular area (CMA), connolly solvent excluded volume (CSEV), exact mass (EM), molecular weight (MW), principal moment of inertia-X component (PMI-X), principal moment of inertia-Y component (PMI-Y), principal moment of inertia-Z component (PMI-Z), molar refractivity (MR), and Ovality (OVAL).

Electronic descriptors such as dipole (DPL), electronic energy (ElcE), highest occupied molecular orbital energy (HOMO), lowest unoccupied molecular orbital energy (LUMO), repulsion energy (NRE), VDW-1,4-energy (E14), Non-1, 4-VDW energy (E_v), and total energy (E) were calculated.

Stepwise multiple linear regression analysis method was used to perform QSAR analysis employing in-house VALSTAT[10] programme. The ±data within the parentheses are the error of regression coefficients associated with corresponding regression coefficients in regression equation. The best model was selected on the basis of various statistical parameters such as correlation coefficient (r), standard error of estimation (SE), sequential Fischer test (F). Quality of the each model was estimated from the cross-validated squared correlation coefficient (q^2)[11], calculated root mean square error (S_{DEP}), chance statistics evaluated as the ratio of the equivalent regression equations to the total number of randomized sets; a chance value of 0.01 corresponds to 1% chance of fortuitous correlation and boot-strapping square correlation coefficient (r^2_{bs}), which confirm the robustness and applicability of QSAR equation.

RESULTS AND DISCUSSION

When data set was subjected to stepwise multiple linear regression analysis, in order to develop 2D-QSAR between antibacterial activity in various microbes as dependent variables and substituent constants as independent variables, several equations were obtained. The statistically significant equation (Eqn. 1) with coefficient of correlation (r=0.804) was considered as

model for antibacterial activity in *S. aureus*. The model showed overall internal statistical significance level better than 99.9% as it exceeded the tabulated $F_{(2,14a0.001)}$=13.9 (Table 2 and fig. 2). pMIC=35.271(±11.593)π-4.088(±1.330)MR+9.079, n=17, r=0.804, r^2=0.646, SE=0.718, F=12.790 (Eqn. 1)

In case of antibacterial activity on *E. coli*, good correlation was shown by the model (Eqn. 2), which explains for 67.2% variance in inhibitory activity with significant F-value ($F_{(2,14)}$=14.357 against $F_{(2,14a0.001)}$=13.9 (Table 2 and fig. 3). pMIC=35.544(±10.535)π-4.117(±1.208)MR+8.909, n=17, r=0.820, r^2=0.672, SE=0.652, F=14.357 (Eqn. 2)

While for another microbe (*P. aeuginosa*), correlation (r=0.829) was identified in Eqn.3. The equation showed overall internal statistical significance level better than 98.0% as it exceeded the tabulated $F_{(2,10a0.02)}$=7.56 (Table 2 and fig. 4). pMIC=34.556(±10.657)π-3.993(±1.224) MR+8.111, n=13, r=0.795, r^2=0.632, SE=0.612, F=8.593 (Eqn. 3)

2D-QSAR analysis suggested that for all the three microbes, substitution at R_1 position very much dominates the activity as compared to the indicator variable at R_2 and R_3 position (fig. 1). At R_1 position, π contributed positively, which is responsible for hydrophobicity of the molecules; but MR contributed negatively, which suggests that less bulky substitutions form the activity.

The series was also subjected to molecular modelling using 3D-QSAR; all the descriptor values for the molecules, calculated from the programme, were considered as independent variables, and inhibitory concentration data (pMIC) in microbes like *S. aureus, E.*

TABLE 2: OBSERVED AND CALCULATED pMIC AND RESIDUAL VALUES OF *S. AUREUS, P. AERUGINOSA*, AND *E. COLI* BY *N*-ALKYL IMIDAZOLE ANALOGUES USING 2D-QSAR MODELS

S. aureus (pMIC)			*P. aeruginosa* (pMIC)			*E. coli* (pMIC)		
Obs.	Cal.	Res.	Obs.	Cal.	Res.	Obs.	Cal.	Res.
2.302	2.947	-0.645	2.001	2.230	-0.229	2.302	2.761	-0.459
3.399	2.589	0.809	2.302	1.937	0.364	3.001	2.414	0.586
3.155	4.036	-0.882	2.700	3.412	-0.713	3.098	3.886	-0.788
3.302	4.072	-0.771	3.001	3.505	-0.505	3.154	3.937	-0.783
3.520	3.797	-0.277	3.096	3.293	-0.196	3.223	3.673	-0.449
3.702	3.480	0.222	3.302	3.040	0.262	3.522	3.368	0.154
3.825	3.163	0.661	3.524	2.787	0.736	3.603	3.063	0.540
4.289	2.847	1.442	3.697	2.535	1.163	4.010	2.758	1.252
3.001	2.530	0.471	2.700	2.282	0.418	3.001	2.452	0.548
2.302	2.213	0.088	2.001	2.029	-0.028	2.302	2.147	0.154
1.223	1.897	-0.674	1.156	1.776	-0.621	1.223	1.842	-0.620
1.156	1.580	-0.424	1.098	1.524	-0.426	1.126	1.537	-0.412
1.126	1.263	-0.138	1.046	1.271	-0.225	1.098	1.232	-0.135
3.823	3.797	0.026	-	-	-	3.823	3.673	0.150
1.156	1.897	-0.741	-	-	-	1.156	1.842	-0.687
4.033	4.036	-0.004	-	-	-	4.008	3.886	0.121
3.999	3.163	0.835	-	-	-	3.888	3.063	0.825

Obs.: Observed pMIC, Cal.: Calculated pMIC, Res.: Residual

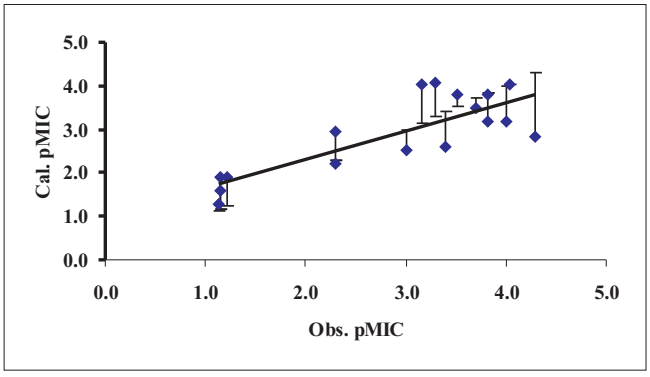


Fig. 2: Plot between observed vs. calculated pMIC values of *S. aureus* for Eqn. 1.
Obs. pMIC: Observed pMIC, Cal. pMIC: Calculated pMIC, y = 0.646x+1.026, r^2 = 0.646.

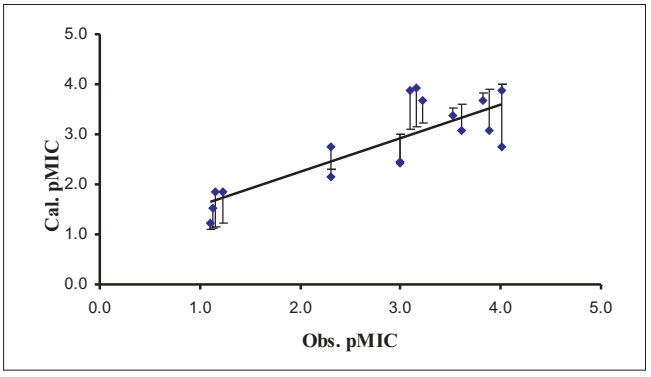


Fig. 3: Plot between observed vs. calculated pMIC values of *E. coli* for Eqn. 2.
Obs. pMIC: Observed pMIC, Cal. pMIC: Calculated pMIC, y = 0.672x + 0.917, r^2 = 0.672.

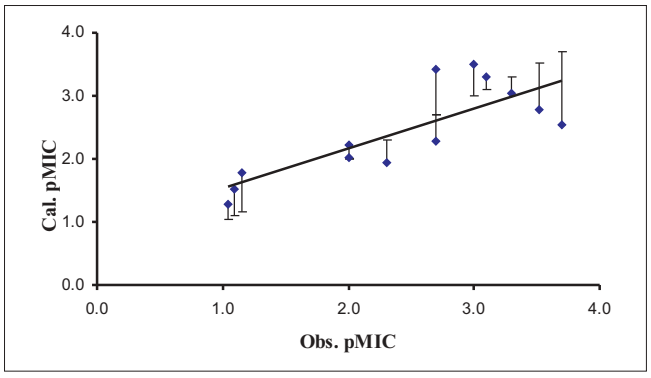


Fig. 4: Plot between observed vs. calculated pMIC values of *P. aeruginosa* for Eqn. 3.
Obs. pMIC: Observed pMIC, Cal. pMIC: Calculated pMIC, y = 0.632x+0.895, r^2 = 0.632.

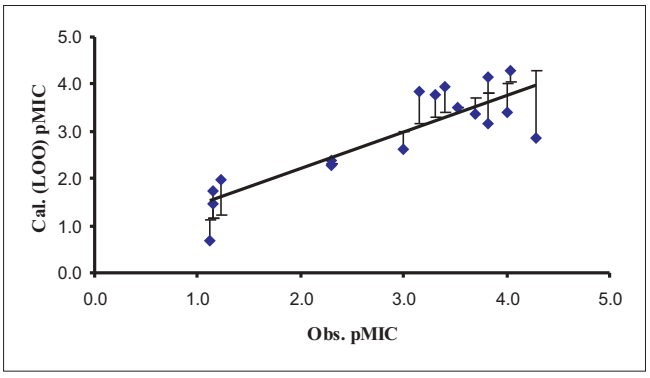


Fig. 5: Plot between observed vs. calculated by leave-one-out cross-validation pMIC values of *S. aureus* for Eqn. 4.
Obs. pMIC: Observed pMIC, Cal. (LOO): calculated by leave-one-out cross-validation pMIC, y = 0.784x+0.636, r^2 = 0.734.

coli, and *P. aeruginosa* were taken as dependent variables. Regression gave various multivariant equations with high correlation coefficients, but the equations having inter-correlation among the physicochemical properties (ICAP) less than 0.3 were considered. The regression analysis study of antibacterial data and physicochemical descriptors for *S. aureus* gave a significant equation (Eqn.4) with better correlation coefficient (r=0.894), which accounted for more than 79.9% of the variance in activity. The data showed overall internal significance level better than 99.9% as $F_{(2,14a0.001)}$=27.878, which exceeded the tabulated $F_{(2,14a0.001)}$=13.9. The equation was further subjected to cross-validation method to confirm the internal consistency; the cross-validated squared correlation coefficient (q^2=0.730) suggested good predictive ability of the antibacterial activity (Table 3 and fig. 5). The robustness and wide pragmatism of the equation was further supported by r^2_{bs}=0.808 and chance <0.001 (Table 4). At par value of bootstrapping squared correlation coefficient (r^2_{bs}) with conventional squared correlation coefficient (r^2) suggested that the model is a proper representative of analogues. pMIC=0.146(±0.075)DPL-3.462e-004(±5.373e-005)PMIY+3.399, n=17, r=0.894,

TABLE 3: CALCULATED AND CALCULATE BY LEAVE-ONE-OUT CROSS-VALIDATION pMIC VALUES OF *S. AUREUS, P. AERUGINOSA, AND E. COLI* BY *N*-ALKYL IMIDAZOLE ANALOGUES USING 3D-QSAR MODELS

S. aureus (pMIC)				*P. aeruginosa* (pMIC)				*E. coli* (pMIC)			
Cal.	†C_{res}	Cal. (LOO)	‡P_{res}	Cal.	†C_{res}	Cal. (LOO)	‡P_{res}	Cal.	†C_{res}	Cal. (LOO)	‡P_{res}
2.366	-0.065	2.372	-0.070	2.125	-0.124	2.137	-0.137	2.281	0.020	2.280	0.022
3.839	-0.440	3.933	-0.535	3.260	-0.958	3.474	-1.172	3.616	-0.615	3.748	-0.748
3.747	-0.592	3.858	-0.704	3.186	-0.487	3.282	-0.582	3.534	-0.435	3.616	-0.517
3.707	-0.405	3.784	-0.482	3.167	-0.167	3.199	-0.198	3.493	-0.340	3.558	-0.404
3.507	0.013	3.505	0.015	3.026	0.070	3.015	0.081	3.307	-0.084	3.321	-0.097
3.419	0.284	3.381	0.321	2.945	0.356	2.898	0.404	3.232	0.290	3.193	0.328
3.211	0.614	3.146	0.678	2.785	0.738	2.707	0.817	3.044	0.559	2.985	0.618
2.962	1.326	2.851	1.437	2.591	1.107	2.494	1.204	2.819	1.190	2.719	1.290
2.650	0.350	2.623	0.377	2.354	0.346	2.325	0.375	2.535	0.466	2.499	0.502
2.294	0.008	2.293	0.008	2.079	-0.078	2.087	-0.086	2.212	0.090	2.204	0.098
1.874	-0.651	1.966	-0.743	1.755	-0.599	1.859	-0.703	1.831	-0.608	1.917	-0.694
1.396	-0.241	1.458	-0.303	1.387	-0.289	1.481	-0.383	1.398	-0.272	1.468	-0.343
0.844	0.282	0.695	0.430	0.961	0.085	0.902	0.144	0.898	0.200	0.792	0.306
4.048	-0.225	4.157	-0.334	-	-	-	-	3.993	-0.170	4.075	-0.252
1.647	-0.491	1.738	-0.582	-	-	-	-	1.621	-0.465	1.707	-0.551
4.201	-0.168	4.292	-0.259	-	-	-	-	4.138	-0.130	4.208	-0.200
3.596	0.403	3.393	0.605	-	-	-	-	3.583	0.305	3.429	0.458

Cal.: Calculated pMIC, Cal. (LOO): Calculated by leave-one-out cross-validation pMIC, †Calculated residual, ‡calculated by leave-one-out cross-validation residual

TABLE 4: QSAR STATISTICS OF SIGNIFICANT EQUATIONS

Eqn. No.	N	r^2	SE	F	ICAP[a]	R^2_{bs}	s_{bs}	Chance	q^2	S_{PRESS}	S_{DEP}
4	17	0.799	0.541	27.878	0.263	0.808	0.132	<0.001	0.730	0.627	0.569
5	17	0.804	0.505	28.661	0.263	0.828	0.113	<0.001	0.740	0.581	0.527
6	13	0.639	0.578	19.449	-	0.666	0.196	<0.003	0.521	0.665	0.611

[a]The maximum limit of inter-correlation among the descriptors used in generation of equations

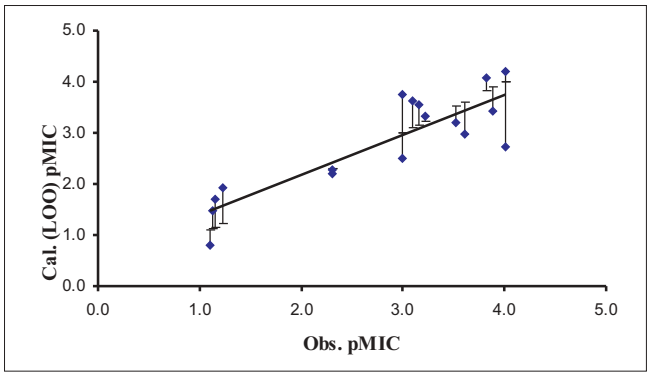


Fig. 6: Plot between observed vs. calculated by leave-one-out cross-validation pMIC values of *E. coli* for Eqn. 5. Obs. pMIC: Observed pMIC, Cal. (LOO): calculated by leave-one-out cross-validation pMIC, y = 0.786x+0.611, r^2 = 0.743.

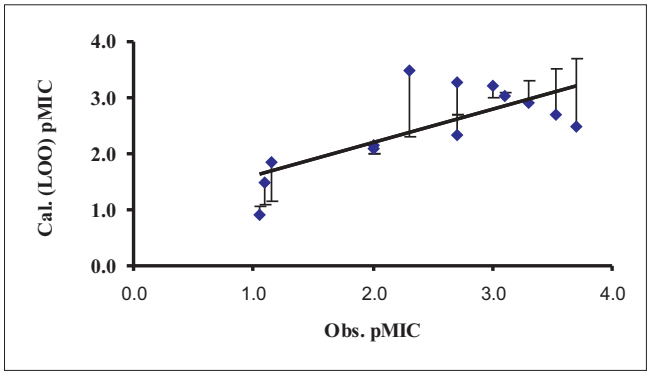


Fig. 7: Plot between observed vs. calculated by leave-one-out cross-validation pMIC values of *P. aeruginosa* Eqn. 6. Obs. pMIC: Observed pMIC, Cal. (LOO): calculated by leave-one-out cross-validation pMIC, y = 0.596x+1.001, r^2 = 0.530.

r^2=0.799, SE=0.541,F=27.878 (Eqn. 4)

For *E. coli*, various statistical significant equations were obtained, but equation 5 was considered as the best model, which accounted for more than 80.4% of the variance in activity. The data showed overall internal significance level better than 99.9% as $F_{(2,14a0.001)}$=28.661 against tabulated $F_{(2,14a0.001)}$=13.9. The equation was further subjected to cross-validation method to confirm the internal consistency; the cross-validated squared correlation coefficient (q^2=0.740) suggested good predictive ability of the antibacterial activity (Table 3 and fig. 6). The wide expediency of the equation was further confirmed by bootstrapping squared correlation coefficient (r^2_{bs}=0.828) and low value of opportunistic (chance <0.001) correlation. pMIC=0.173(±0.070)DPL-3.139e-004(±5.016e-005)PMIY+3.051, n=17, r=0.896, r^2=0.804, SE=0.505, F=28.661 (Eqn. 5)

In case of *P. aeruginosa*, monovariant (Eqn. 6) was considered as model on the basis of fitness of the equations, which explained 63.9% of the variance in activity. The model having significance level better than 99.9% as $F_{(2,10\alpha0.001)}$=19.449, which exceeds the tabulated $F_{(2,10\alpha0.001)}$=17.9. The equation 4 also has better internal predictivity of the activity as compared to equation 3 with high q^2 value (Table 3 and fig. 7). pMIC = -2.670e-004(±6.054e-005)PMIX+3.386, n=13, r=0.799, r^2=0.639, SE=0.578, F=19.449 (Eqn. 6)

The 2D analysis suggested that the substitution at R_1 position with various alkyl groups affect the antibacterial activity of imidazole analogues as compared to substitutions at R_2 and R_3 position. The QSAR studies revealed that spatial parameter PMIY plays a significant role in explaining antibacterial activity of *N*-alkyl imidazoles. It contributed negatively to the expression, which suggested that less bulky groups around Y-axis in the molecules are favourable for the activity. DPL contributed positively to the activity up to a small extent as compared to the PMIY, suggesting that the moiety, which increases the charge distribution over the molecules, is favourable for the activity. It could be concluded that new molecules should be designed by considering the shape and size of the molecule, possessing at least one electron-withdrawing group on imidazole ring system and optimising the hydrophobicity and bulkiness at R_1 position.

ACKNOWLEDGEMENTS

Authors thank the Director, Shri G. S. Institute of Technology and Science, Indore, for providing facilities to complete this work. The author AKG is thankful to CSIR, New Delhi, for providing SRF.

REFERENCES

1. Agrafiotis, D.K., **J. Comp. Aided Mol. Des.**, 2002, 16, 335.
2. Yalcin, I., Oren, I., Sener, E., Akin, A. and Ucarturk, N., **Eur. J. Med. Chem.**, 1992, 27, 401.
3. Ertepinar, H., Gok, Y., Geban, O. and Ozden, S., **Eur. J. Med. Chem.**, 1995, 30, 171.
4. Neu, H.C., **Science**, 1992, 257, 1064.
5. Antolini, M., Bozzoli, A., Ghiron, C. and Gordon, K., **Bioorg. Med. Chem. Lett.**, 1999, 9, 1023.
6. Khabnadideh, S., Rezaei, Z. Khalafi, N. Bahrinajafi, R. Mohamadi, R. and Farrokhroz, A., **Bioorg. Med. Chem. Lett.**, 2003, 13, 2863-2865.
7. Hansch, C. and Leo, A., In; Substituent Constants for Correlation Analysis in Chemistry and Biology, John Wiley & Sons, New York, 1979, 48.
8. CS Chem Office, Version 6.0, Cambridge Soft Corporation, Software Publishers Association, 1730 M Street, NW, Suite 700, Washington, D.C. 20036.

9. Keir, L.B., In; Molecular Orbital Theory in Drug Research, Academic press, New York, 1971, 62.
10. Gupta, A.K., Arockia Babu M. and Kaskhedikar, S.G., **Indian. J. Pharm. Sci.**, 2004, 66, 396.
11. Kubinyi, H., In; QSAR: Hansch Analysis and Related Approaches, Weinheim, New York, 1993, 91

Accepted 27 May 2006
Revised 20 September 2005
Received 11 June 2005
Indian J. Pharm. Sci., 2006, 68 (3): 341-346