

QSAR study of substituted 2-pyridinyl guanidines as selective urokinase-type plasminogen activator (uPA) inhibitors

C. KARTHIKEYAN, N.S. HARI NARAYANA MOORTHY[†], & PIYUSH TRIVEDI

School of Pharmaceutical Sciences, Rajiv Gandhi Technical University, Airport Bypass Road, Gandhi Nagar, Bhopal 462036, MP, India

(Received 9 June 2007; accepted 19 October 2007)

Abstract

A quantitative structure–activity relationship analysis was conducted on two different series of pyridinylguanidines acting as inhibitors of urokinase-type plasminogen activator using QuaSAR descriptors of molecular modeling software MOE. Multiple linear regression analysis following a stepwise scheme was employed to generate QSARs that relate molecular descriptors to uPA inhibitory activity data of the title compounds. Among the several QSARs generated by MLR analysis, the best models were selected on the basis of their statistical significance and predictive potential. The interpretation of the selected QSAR models suggest that uPA inhibitory activity of compounds in series 1 is influenced by their molecular shape, molecular flexibility and halogen atoms in the molecule whereas the uPA inhibitory potency of compounds in series 2 is dependent on molecular lipophilicity, number of double bonds and spatial orientation of bulky substituents in the molecule.

Keywords: MOE, QSAR, pyridinyl guanidines, uPA, inhibitors

Introduction

Metastasis is the cause of most cancer-related deaths. The proteolytic degradation of the extracellular matrix is recognized as a mechanism that plays an important role in the metastatic process. Proteolytic enzymes are required to mediate tumor cell invasion into adjacent tissues and initiate the metastatic process [1]. Urokinase plasminogen activator (urokinase, uPA) a trypsin-like serine protease has been implicated in many cellular processes such as tumorigenesis, cell proliferation and migration, cell adhesion, angiogenesis and metastasis [2–4]. The primary role of uPA is to convert plasminogen into its active form plasmin [5]. Plasmin, a serine protease, is involved in the degradation of several extracellular matrix components, the key stage in the process of angiogenesis and thus contributes to the tumor cell survival and metastatic properties of cancer cells [6]. Clinical studies have demonstrated that

increased expression of uPA in tumour tissues is highly correlated with tumor cell migration, invasion, proliferation, progression and metastasis [7]. Additionally, it has also been reported that uPA activity is increased in metastatic tumors compared with primary tumors in experimental animals [8]. The pivotal role of uPA in tumor angiogenesis and metastases makes it a potentially attractive target for preventing metastasis of tumors.

Several small molecules have been identified as potent inhibitors of uPA such as benzamidines, phenylguanidines, acylguanidines and bisbenzamidines [6]. However, these compounds also inhibit other related serine proteases such as tissue plasminogen activator, plasmin and thrombin, which play a significant role in the fibrinolytic cascade [6]. In order to circumvent these unwanted adverse effects arising out of inhibition of homeostatic enzymes, efforts were

Correspondence: P. Trivedi, School of Pharmaceutical Sciences, Rajiv Gandhi Technological University, Bhopal, India. Fax: 91 0755 2742006. E-mail: drtrivedislab@yahoo.co.in

[†]nshnm06@yahoo.co.in

devoted to develop inhibitors with high degree of selectivity towards uPA. Recently, Barber et al. [9,10] reported two series of pyridinylguanidines with selective uPA inhibitory activity. In view of further progression in the development of such inhibitors, the present study attempts to explore QSAR of both the series of uPA inhibitory pyridinylguanidines reported by Barber et al. QSAR is a powerful lead-compound optimization technique, which quantitatively relates variations in biological activity to changes in molecular properties (descriptors). In other words, it attempts to link activity data with descriptors chosen via identification of the “rules” that can be further used to guide chemical synthesis when new chemical entities are developed. In the past, many quantitative structure activity relationship studies were performed on potential uPA inhibitors to explore the structural features determining the binding affinity and selectivity for the enzyme. Yang et al. [11] investigated the features contributing to selective inhibition of urokinase using Hansch analysis of a series of phenyl guanidines exhibiting inhibitory activity against six serine proteinases. Bhongade et al. [12,13] reported comparative molecular field analysis and comparative molecular similarity indices analysis of uPA inhibitory indole/benzimidazole-5-carboxamides. Related to the foregoing and in continuation of our efforts focused on the design and development of novel and potent protease inhibitors targeting cancer, the present work strives to examine the applicability of QSAR approach to the series of pyridinyl guanidines reported by Barber et al. [9,10] and gain some fundamental understanding on factors influencing uPA binding affinity of these molecules.

Materials and methods

The dataset investigated for QSAR study includes two series of pyridinyl guanidines; series 1 and series 2. Series 1 comprises of 16 compounds whereas series 2 comprises of 19 compounds. The structure of molecules in Series 1 and Series 2 are presented in Table I and Table II respectively. The uPA inhibitory activity data of compounds in both the series have been reported as IC_{50} values, where IC_{50} refers to the experimentally determined micro-molar concentration of the compounds in the selected series required to inhibit the enzyme by 50%. The biological activity values [IC_{50} (μM)] were transformed into their molar scale and subsequently converted to negative logarithmic scale and then used as the response variable for the QSAR analysis.

The softwares employed for the present study are Molecular Operating Environment [14] (MOE 2002.03), statistical software SYSTAT [15] (Version 10.2) and inhouse validation program VALSTAT [16]. All the computations were carried out on Compaq PIV workstation. The structures of the

compounds in the selected series were sketched using builder module of MOE software and sketched structures were subsequently energy minimized up to root mean square gradient of 0.01 kcal/molÅ using MMFF94 force field. Conformational search of each energy-minimized structure was performed employing stochastic search routine. All the conformers generated for each structure were carefully scrutinized in conformational geometry panel and the lowest energy conformer of each structure selected for re-optimization by applying AM1 semi-empirical method as implemented in semi-empirical molecular orbital software program MOPAC (version 7) of MOE.

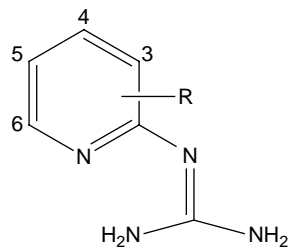
Molecular descriptors programmed into QuaSAR module of MOE were calculated for the geometry-optimized structures of the compounds in the series. The MOE package calculates nearly about 180 descriptors named as QuaSAR descriptors [17]. The QuaSAR descriptors can be broadly classified into 2D and 3D descriptors. The two dimensional descriptors include traditional physicochemical properties, (atom counts and bond counts, mr, logP and vdw_area etc), connectivity-based topological descriptors (Kier and Hall connectivity and Kappa Shape indices; adjacency and distance matrix descriptors), pharmacophore feature descriptors (e.g. Donor, Acceptor, Polar, Positive, Negative, Hydrophobic.), partial charge descriptors based on Partial Equalization of Orbital Electronegativities method and quantum chemical descriptors such as HOMO and LUMO energies, dipole, etc. The 3D descriptors in the module include potential energy descriptors, surface area descriptors, volume descriptors, shape descriptors, and conformation-dependent charge descriptors.

The calculated descriptors were initially screened for invariant nature (constant and near constant values) to reduce the number of descriptors for statistical analysis. Further reduction in the descriptor pool was accomplished by eliminating the descriptors with no significant correlation with the dependent variable ($R^2 < 0.1$) using QuaSAR-contingency module of MOE. QuaSAR-Contingency is a statistical application designed to assist in the selection of descriptors for QSAR.

In the first step of statistical analysis, pairwise correlation analysis of the reduced set of descriptors was established by the calculation of correlation matrix. If given pair of descriptors have a Pearson correlation greater than 0.7, then one member of the pair is randomly omitted for the QSAR study. As a resultant, the descriptor pool is further reduced to 45 descriptors for series 1 and 52 descriptors for series 2. This reduced descriptor pool was used in the generation of QSAR models.

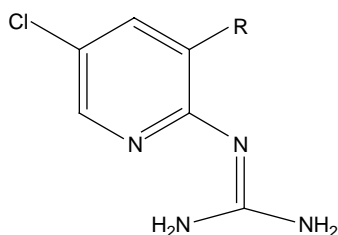
The QSAR models were found through forward stepwise regression procedure employing statistical program SYSTAT 10.2 version. The best QSAR models were selected on the basis of the highest

Table I. Structure, biological activity and physicochemical descriptors of series 1



S. No	R	BA (μM)	$-\log\text{BA}$	petitjeanSC	b_IrotR	KierFlex
1	H	29.5	4.5302	1	0.1667	1.4283
2	3-CH ₃	37.7	4.4237	1	0.1905	1.6467
3	4-CH ₃	83.8	4.0768	1	0.1905	1.6467
4	5-CH ₃	32.3	4.4908	0.75	0.1905	1.6467
5	6-CH ₃	173.0	3.7619	1	0.1905	1.6467
6	4-C ₆ H ₅	7.10	5.1487	0.8	0.1379	2.0737
7	4-OCH ₃	53.3	4.2732	0.75	0.2273	2.1281
8	3-OH	147.0	3.8327	1	0.2105	1.6467
9	3-OCH ₃	62.0	4.2076	1	0.2273	2.1281
10	5-Cl	10.0	5.0000	0.75	0.1667	2.2402
11	5-Br	3.13	5.5044	0.75	0.1667	2.7614
12	3,5-diCl	5.47	5.2620	0.75	0.1667	3.2121
13	3-Br, 5-Cl	4.83	5.3160	0.75	0.1667	3.8483
14	3-Cl, 5-Br	2.90	5.5376	0.75	0.1667	3.8483
15	3-Cl, 5-CF ₃	40.7	4.3904	1	0.1905	2.9205
16	3,5-diCl-4-Me	8.70	5.0605	0.75	0.1905	3.4314

Table II. Structure, biological activity and physicochemical descriptors of series 2



S.No	R	BA (μM)	$-\log\text{BA}$	chilv_C	b_double	E_sol	PMI - Z	logP (o/w)
1	(E)-CH=CHCO ₂ H	1.5	5.8239	2.0267	3	-15.3085	1328.947	1.04
2	(E)-CH=CHCONHMe	1.28	5.8928	2.0267	3	-15.1895	1889.878	0.662
3	E)-CH=CHCO(1-morpholino)	2.1	5.6778	3.0267	3	-15.4765	361.3822	0.425
4	(E)-CH=CHCONHCH ₂ Ph	0.77	6.1135	4.2909	3	-12.4917	1392.558	2.45
5	(E)-CH=CHCON(Me)CH ₂ Ph	2	5.6990	4.2909	3	-6.5011	2556.398	2.647
6	CH ₂ CH ₂ CO ₂ H	9.31	5.0311	2.3231	2	-14.8113	638.4941	0.572
7	OCH ₂ CO ₂ H	90.2	4.0448	1.4696	2	-13.5755	609.7509	0.18
8	OCH ₂ CONHCH ₂ Ph	16.1	4.7932	3.7338	2	-9.0434	353.2897	1.59
9	OCH ₂ CH ₂ OCH ₃	16.4	4.7852	1.6160	1	-11.0577	1442.224	0.345
10	C ₆ H ₅	166.3	3.7791	3.2767	1	-9.12613	939.6092	2.679
11	OC ₆ H ₅	2.13	5.6716	3.0267	1	-12.9468	484.1166	2.329
12	OCH ₂ C ₆ H ₅	0.92	6.0362	3.3803	1	-13.5082	2463.676	2.463
13	C≡CC ₆ H ₅	0.77	6.1135	3.7767	1	-14.6776	1390.561	3.425
14	(E)-CH=CHC ₆ H ₅	0.55	6.2596	3.9374	2	-13.2539	1335.126	3.323
15	(E)-CH=CHC ₆ H ₁₁	0.73	6.1367	4.8879	2	-10.8807	118.8222	3.338
16	(E)-CH=CH(2-pyridinyl)	1.6	5.7959	3.3154	2	-13.3939	1285.165	2.052
17	(E)-CH=CH(4-C ₆ H ₄ OMe)	0.69	6.1611	3.8481	2	-12.8072	1330.748	3.279
18	E)-CH=CH[(3,4-OCH ₂ O)C ₆ H ₄]	0.49	6.3098	3.7647	2	-14.2263	2392.332	3.028
19	(E)-CH=CH(3-C ₆ H ₄ CO ₂ H)	0.17	6.7695	4.0981	3	-14.1573	2820.975	3.037

correlation coefficient (R^2), the lowest standard error SEE, F statistics and the statistical relevance of the incorporated descriptors. Another important criterion for the model selection is low variable collinearity between the descriptors in the same model. To confirm the absence of multicollinearity in the selected correlations, variance inflation factor (VIF) values [18] were calculated for each parameter in the regression. VIF value was calculated from $1/(1 - R^2)$, where R^2 is the multiple correlation coefficient of one parameter's effect regressed on the remaining parameters. Although it is stated that VIF values less than 10 are statistically satisfactory, we chose a more stringent criterion ($VIF < 5$) for selection of QSAR models. The Z-score method was adopted for the detection of outliers. Z-Score can be defined as absolute difference between the value of the model and the activity field, divided by the square root of the mean square error of the data set. Any compound which shows a value of Z-score higher than 2.5, during generation of a particular QSAR model is considered as outlier.

Finally, the stability and predictive ability of every potential model was tested by a cross validation method following a leave-one-out scheme using inhouse program VALSTAT. The validation parameters calculated are squared cross-correlation coefficient (q^2), standard deviation of sum of square of difference between predicted and observed values (S_{PRESS}) and standard deviation of error of prediction (S_{DEP}). The validation parameter q^2 describes the stability of a regression model obtained by focusing on the sensitivity of the model to the elimination of any single data point and capacity to estimate the activity of compounds outside the training set. For a reliable QSAR model, the calculated squared cross-correlation coefficient q^2 should be greater than 0.6.

Results and discussion

QSAR model for series 1

The best QSAR model generated for uPA inhibitory activity of compounds in Series 1 is given below

$$\begin{aligned} -\log BA = & -1.712(\pm 1.248)\text{PetitjeanSC} \\ & - 10.814(\pm 5.515)\text{b_1rotR} \\ & + 0.281(\pm 0.183)\text{KierFlex} \\ & + 7.472(\pm 1.557) \end{aligned} \quad (1)$$

$N = 16$, $R = 0.94$, $R^2 = 0.89$, $SEE = 0.21$, $F = 33.28$, $P > 0.000$, $Q^2 = 0.83$, $S_{PRESS} = 0.27$, $S_{DEP} = 0.24$.

In QSAR model given above, N is the number of data points and the figures in the parentheses are 95% confidence limits.

Model 1 produces good description for uPA inhibitory activity of pyridinyl guanidines, as it accounts for more than 80% of the total variance in the biological activity. The F statistic of the model is significant at 99% level as the calculated F-test value exceed the critical F value ($F_{(3,12)} = 5.95$) by a large margin. Since the p value less than 0.01, there exist a statistically significant relationship between the variables in the model at 99% confidence level. Furthermore, the intercorrelation coefficient between the descriptors in the model is less than 0.7 as indicated by correlation matrix (Table III) and the variance inflation factors of all the descriptors in the model as shown in Table V are less than 5, which indicate the absence of multicollinearity in the models.

The molecular descriptors in the model 1 include petitjeanSC, KierFlex and b_1rotR. Among the three descriptors found in the model, petitjeanSC and KierFlex are topological descriptors and b_1rotR is a constitutional descriptor, which indicates that variation in the uPA inhibitory activity of pyridinyl guanidines can be described in terms of structural components of the molecule. The topological descriptor petitjeanSC refers to petitjean shape coefficient, which encodes information regarding the shape of the molecule. The regression coefficient of the descriptor petitjeanSC bears a negative sign, which suggest that shape of the molecule is an important determinant for uPA inhibition exhibited by pyridinyl guanidines. The topological descriptor KierFlex [19] encodes the structural properties that restrict a molecule from being "infinitely flexible", the model for which is an endless chain of $C(sp^3)$ atoms. The value of KierFlex decreases with the presence structural features considered to preventing a molecule from attaining infinite flexibility are fewer atoms, cyclicity, branching, conjugation and presence of atoms with a covalent radii smaller than $C(sp^3)$. However, an observation of the substituents in the pyridinyl ring indicates none of the substituents except for phenyl ring seem to exhibit cyclicity, branching or conjugation. So, it appears that descriptor might actually characterize the covalent radii of the atoms in the substituents. Hence, the positive coefficient of the descriptor in the model may be interpreted that the presence of atom with larger covalent radii than carbon such as chlorine, bromine will increase the uPA binding affinity of pyridinyl guanidines. The constitutional descriptor b_1rotR [12] represents fraction of

Table III. Correlation matrix for series 1.

	$-\log BA$	petitjeanSC	b_1rotR	KierFlex
$-\log BA$	1.00	-0.77	-0.70	0.75
PetitjeanSC	-0.77	1.00	0.38	-0.59
b_1rotR	-0.70	0.38	1.00	-0.30
KierFlex	0.75	-0.59	-0.30	1.00

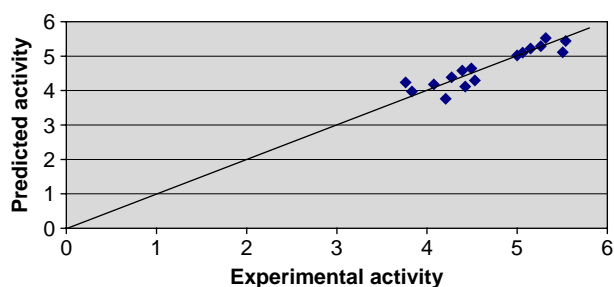


Figure 1. Experimental vs predicted activity of model 1.

rotatable single bonds in the molecule. The negative coefficient of the descriptor in model suggests that increase in the fraction of rotatable single bonds will decrease the uPA inhibitory potency of pyridinyl guanidines. Furthermore, it is quite evident from Table I that fractional increase in number of rotatable single bonds occurs with methoxy and hydroxyl substitution in the pyridinyl ring hence it may be interpreted that presence of aforementioned substituents is not conducive for uPA inhibitory activity.

The predictive potential of model 1 as judged by leave-one-out procedure is fairly high ($Q^2 > 0.8$) suggesting that the models will be useful for meaningful predictions. Further support in this regard is obtained from the low values of the cross-validation parameters S_{press} and S_{dep} . Predicted activity values for the compounds in the training set alongwith their corresponding experimental activity values are recorded in Table VI. The predicted $-\log IC_{50}$ values of the compounds in the test set are in good agreement with the corresponding experimental values as shown in Figure 1.

QSAR models for series 2

The best QSAR models generated for series 2 are given below alongwith regression statistics as well as crossvalidation statistics.

$$\begin{aligned}
 -\log BA = & 0.544(\pm 0.164)\text{chi1v_C} \\
 & - 0.156(\pm 0.069)\text{E_SOL} \\
 & + 0.0003(\pm 0.0001)\text{PMI} - Z \\
 & + 1.505(\pm 1.220) \quad (2)
 \end{aligned}$$

$N = 18, R = 0.92, R^2 = 0.85, SEE = 0.28, F = 26.86, P > 0.000, Q^2 = 0.74, S_{PRESS} = 0.37, S_{DEP} = 0.33.$

$$\begin{aligned}
 -\log BA = & 0.419(\pm 0.099)\log P(o/w) \\
 & - 0.110(\pm 0.045)\text{E_Sol} \\
 & + 0.272(\pm 0.139)\text{b_double} \\
 & + 2.935(\pm 0.748) \quad (3)
 \end{aligned}$$

$N = 17, R = 0.94, R^2 = 0.89, SEE = 0.19, F = 35.22, P > 0.000, Q^2 = 0.78, S_{PRESS} = 0.27, S_{DEP} = 0.24$

Both the models given above manifest a satisfying correlation, as they account for more than 80% of the observed variance in the biological activity. Accuracy in the analysis is shown by low values of standard error of estimate. The F statistics are significant at 99% level and the p value less than 0.000 indicates that the models are not a mere result of chance. The correlation matrix and VIF data recorded in Tables IV and V respectively depicts the orthogonal nature of the descriptors in the selected correlations.

In addition to statistical significance, the generated models also exhibit good predictive capacity as shown by cross-validation statistics given alongwith the models. Both the models show predicted variance greater than 70% since their Q^2 values are greater than 0.7. It is also noteworthy that the PRESS statistics (S_{PRESS}) of the both the models are comparable to their respective standard error of estimate (SEE), suggesting that the models will be useful for prediction. Predicted $-\log IC_{50}$ values of the compounds in the training set together with the experimental $-\log IC_{50}$ values are recorded in Table VI. Plot of experimental versus predicted $-\log IC_{50}$ values of the compounds in the training set for model 2 and 3 is illustrated in Figures 2 and 3 respectively.

Model 2 shows linear positive dependence of uPA inhibitory activity on topological descriptor chi1v_C , steric descriptor $\text{PMI} - Z$ and linear negative dependence of activity on potential energy descriptor E_Sol . The topological descriptor, chi1v_C [12] encodes information regarding degree of branching, cyclization, heteroatom content, and heteroatom position in

Table IV. Correlation matrix for series 2.

	$-\log BA$	chi1v_C	B_double	E_sol	$\text{PMI} - Z$	$\log P(o/w)$
$-\log BA$	1.00	0.55	0.36	-0.37	0.48	0.54
chi1v_C	0.55	1.00	0.10	0.37	0.17	0.86
b_double	0.36	0.10	1.00	-0.15	0.20	-0.19
E_sol	-0.37	0.37	-0.15	1.00	-0.02	0.24
$\text{PMI} - Z$	0.48	0.17	0.20	-0.02	1.00	0.31
$\log P(o/w)$	0.54	0.86	-0.19	0.24	0.31	1.00

Table V. Variance inflation factor (VIF) value of descriptors in QSAR models.

Model No	Parameters	VIF
Model 1	PetitjeanSC	1.66
	b1rotR	1.17
	KierFlex	1.56
Model 2	chi1v_C	1.22
	E_sol	1.18
	PMI – Z	1.03
Model 3	b_double	1.03
	E_sol	1.04
	logP(o/w)	1.07

the molecule and is calculated as follows:

$$\text{chi1v_C} = \sum (\delta_i^v \delta_j^v)^{-1/2} \quad (4)$$

Where δ_i^v and δ_j^v are the vertex connectivity indices of carbon atoms i and j , respectively, and the summation extends to all bonded pairs of non hydrogen carbon atoms in the group or molecule. For second and third rows of atoms, Kier [19] gave a unified definition of δ_i^v , as expressed by Equation 5. In this equation, Z_i^v is the number of valence electrons of atom i , h_i is the number of hydrogen atoms attached to it, and Z_i is its atomic number.

$$\delta_i^v = (Z_i^v - h_i) / (Z_i - Z_i^v - 1) \quad (5)$$

The value of the δ_i^v increase with branching and with increase in the number of heteroatoms in the molecule. Thus, the positive coefficient of the descriptor in model 3 implies that increase in molecular branching and presence of heteroatoms in

the molecule will decrease the uPA inhibitory potency of pyridinyl guanidines.

The steric descriptor PMI – Z refers to the principal moment of inertia along Z-axis of the molecule. The positive coefficient of the descriptor suggests that orientation of molecular distribution along Z-axis of the molecule is conducive for the uPA inhibitory potency of pyridinyl guanidines. Moreover, the appearance of the descriptor also highlights the likelihood of shape specific steric interactions between molecules and the receptor. The potential energy descriptor E_Sol bears a negative coefficient in model 3, which indicates that increase in the solvation energy of the molecule will decrease the uPA binding affinity of pyridinyl guanidines.

Model 3 comprises of the following descriptors; lipophilicity descriptor logP(o/w), potential energy descriptor E_Sol and constitutional descriptor b_double. The descriptor logP(o/w) refers to log of the octanol/water partition coefficient of the molecule and is considered as a measure of lipophilicity of a molecule. The positive coefficient of the descriptor in model 3 suggests that increase in the overall lipophilicity of the molecule will in turn increase the uPA inhibitory activity of pyridinyl guanidines. The negative coefficient associated with potential energy descriptor E_Sol reaffirms the conclusions drawn from model 2. The third descriptor b_double in model 3 represents number of double bonds in the molecule. From the positive coefficient of this term it appears that increase in the number of double bonds in the molecule have a beneficial effect on the uPA inhibitory activity shown by pyridinyl guanidines. As it can be seen from table the descriptor values are larger when an intermittent double bond is present in the side

Table VI. Predicted activity values of selected models.

S. No	Series 1		Series 2		
	–logBA	Predicted activity of Model 1	–logBA	Predicted activity of Model 2	Predicted activity of Model 3
1	4.5302	4.2920	5.8239	5.3329	5.8936
2	4.4237	4.1166	5.8928	5.4878	5.6459
3	4.0767	4.1781	5.6778	5.6856	5.6233
4	4.4908	4.6426	6.1135	6.2420	6.1681
5	3.7619	4.2339	5.6989	5.5741	5.4028
6	5.1487	5.2213	5.0310	5.3285	5.4378
7	4.2733	4.3869	4.0448	4.8388	NA*
8	3.8327	3.9713	4.7932	5.1826	5.2746
9	4.2076	3.7601	4.7852	4.4258	4.3739
10	5.0000	5.0184	3.7791	NA*	NA*
11	5.5044	5.1105	5.6716	5.2796	5.6009
12	5.2620	5.2925	6.0362	6.2613	5.6597
13	5.3160	5.5237	6.1135	6.3142	6.3304
14	5.5376	5.4399	6.2596	6.1225	6.3503
15	4.3904	4.5799	6.1367	5.7474	6.0719
16	5.0605	5.1003	5.7959	5.8036	5.8212
17			6.1611	6.0032	6.2858
18			6.3098	6.5722	6.3229
19			6.7695	6.8517	6.5341

*Outlier compounds

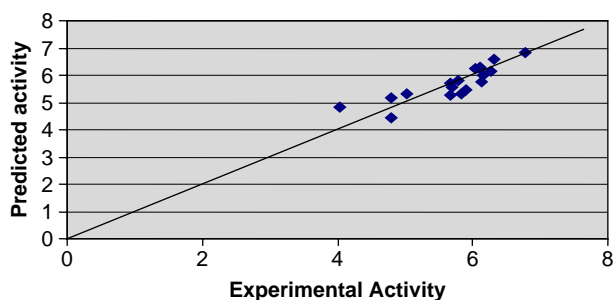


Figure 2. Experimental vs predicted activity of model 2.

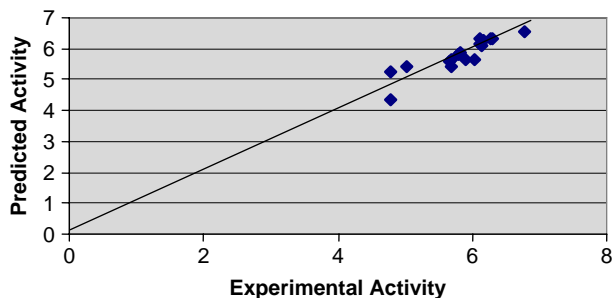


Figure 3. Experimental vs predicted activity of model 3.

chain at carbon-3 of pyridine ring. The positive effect of intermittent double bond may be attributed to the fact that incorporation of the double bond makes the molecule assume a trans configuration and orients the substituents attached to the double bond to the opposite direction of the pyridinyl ring which in turn might help in the interaction of the molecule with the enzyme.

Compound number **9** was found to be a statistical outlier (Z score > -2.72551) during the development of model 2 while compounds **9** and **6** were found to be successive outliers in model 3. The reason for outlying behavior of compound **9** may be attributed to the steric hindrance owing to presence of a phenyl substitution vicinal to guanidino moiety in the pyridine ring. Furthermore, the phenyl ring is directly linked to the pyridinyl ring form a biphenyl like structure, which may not permit proper orientation of the guanidino group in the active site of the enzyme. However, the reason for outlying behavior of compound **6** is not immediately apparent and merits further studies.

Conclusion

Finally to summarize, two series of pyridinyl guanidines reported by Barber et al. have been analyzed applying a QSAR approach. For the 16 compounds evaluated for uPA inhibitory activity in series 1, a statistically satisfactory QSAR model is developed with topological descriptors Petitjean shape coefficient, Kierflex and constitutional descriptor b_{1rotR} . Structural information encoded by the

forementioned descriptors suggests that uPA binding affinity of the compounds studied is dependent upon molecular shape, presence of halogen atoms and molecular flexibility. Furthermore, the predictive ability of the model was established by cross-validation procedure following a leave-one-out scheme.

Two QSAR models of good statistical quality and predictive capacity were generated for describing uPA inhibitory activity of compounds in series 2. The interpretation of the models suggest that uPA inhibitory activity of pyridinyl guanidines in series 2 increases with lipophilicity of molecules while decreases with molecular branching and increase in the heteroatom content of the molecule. The models also highlight the possibility of shape specific steric interactions between the molecules and the enzyme.

Acknowledgements

One of the authors, C. Karthikeyan wishes to thank CSIR, India for providing a senior research fellowship. The authors wish to thank Tata Elxsi for providing MOE software for the study. Grateful acknowledgements to Prof. P. B. Sharma, Vice Chancellor, Rajiv Gandhi Technical University, Bhopal and Director, SGSITS, Indore for the experimental facilities provided. The authors also wish to thank Prof. L.B. Kier and Prof. M. Petitjean for providing reprints of their research works. This research work is funded by grant from All India Council of Technical Education, New Delhi, India.

References

- [1] Dano K, Andreasen PA, Grondahl-Hansen J, Kristensen P, Nielsen LS, Skriver L. Plasminogen activators, tissue degradation and cancer. *Adv Cancer Res* 1985;44:139–266.
- [2] Rabbani SA, Xing RHM. Role of urokinase and its receptor in invasion and metastasis of hormone-dependent malignancies. *Int J Oncol* 1998;12:911–920.
- [3] Irigoyen JP, Munoz-Canoves P, Montero L, Koziczak M, Nagamine Y. The plasminogen activator system: Biology and regulation. *Cell Mol Life Sci* 1999;56:104–132.
- [4] Sidenius N, Blasi F. The urokinase plasminogen activator system in cancer. *Cancer Metast Rev* 2003;22:205–222.
- [5] Ellis V, Pyke C, Eriksen J, Solberg H, Dano K. The urokinase receptor: Involvement in cell surface proteolysis and cancer invasion. *Ann N Y Acad Sci* 1992;667:13–31.
- [6] Chapman HA. Plasminogen activators, integrins, and the coordinant regulation of cell adhesion and migration. *Curr Opin Cell Biol* 1997;9:714–724.
- [7] Giovanni A, David PF. Protease inhibitors in the clinic. *Med Chem* 2005;1:71–104.
- [8] Abderrahim A, Stephanie K, Gilles T, Louis-Georges S, Pnina B, David G, Shafaat AR. Urokinase overproduction results in increased skeletal metastasis by prostate cancer cells *in vivo*. *Cancer Res* 1994;54(9):2372–2377.
- [9] Christopher GB, Roger PD, Valerie AH. Selective urokinase-type plasminogen activator (uPA) inhibitors. Part 1: 2-Pyridinylguanidines. *Bioorg Med Chem Lett* 2002;12: 181–184.

- [10] Christopher GB, Roger PD. Selective urokinase-type plasminogen activator (uPA) inhibitors. Part 2: (3-Substituted-5-halo-2-pyridinyl) guanidines. *Bioorg Med Chem Lett* 2002; 12:185–187.
- [11] Heechung Y, Jack H, Ki HK, Greer J. Selective inhibition of urokinase by substituted phenylguanidines: Quantitative structure–activity relationship analyses. *J Med Chem* 1990;33(11):2956–2961.
- [12] Bhoomendra AB, Veerappa VG, Andanappa KG. 3D-QSAR CoMFA studies on trypsin-like serine protease inhibitors: A comparative selectivity analysis. *Bioorg Med Chem* 2005;13:2773–2782.
- [13] Bhongade BA, Gadad AK. 3D-QSAR CoMFA/CoMSIA studies on urokinase plasminogen activator (uPA) inhibitors: A strategic design in novel anticancer agents. *Bioorg Med Chem* 2004;12:2797–2805.
- [14] MOE is a molecular modeling package developed by Chemical Computing Group Inc., Canada, 2002.
- [15] SYSTAT 10.2 is a statistical package developed by SYSTAT software Inc, USA, 2003.
- [16] VALSTAT is a statistical program developed by Dr. Arun Kumar Gupta, SGSITS, Indore, 2004.
- [17] Lin A. J Chem Comput Group, http://www.chemcomp.com/Journal_of_CCG/Features/descr.htm
- [18] Cho DH, Lee SK, Kim BT, No KT. Quantitative structure-activity relationship (QSAR) study of new fluorovinylxyacetamides. *Bull Korean Chem Soc* 2001;22:388–394.
- [19] Kier LB, Hall LH. Structure–activity studies on hallucinogenic amphetamines using molecular connectivity. *J Med Chem* 1977;20(12):1631–1636.