Available online at www.sciencedirect.com
ScienceDirect
European Journal of Medicinal Chemistry 41 (2006) 1339–1346
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY
http://france.elsevier.com/direct/ejmech

Short communication

Quantitative structure activity relationship studies of aryl heterocycle-based thrombin inhibitors

Sumit Deswal, Nilanjan Roy*

Pharmacoinformatics division National Institute of Pharmaceutical Education and Research, Sector 67, Phase X, 160062 S.A.S. Nagar, Punjab, India

Revised and accepted 3 July 2006
Available online 01 August 2006

Abstract

A quantitative structure activity relationship (QSAR) analysis has been performed on a data set of 42 aryl heterocycle-based thrombin inhibitors. Several types of descriptors including topological, spatial, thermodynamic, information content and E-state indices were used to derive a quantitative relationship between the anti thrombin activity and structural properties. Genetic algorithm based genetic function approximation method of variable selection was used to generate the model. Best model was developed when number of descriptors in the equation was set to five. Highly statistically significant model was obtained with atom type logP descriptors, logP and Shadow_YZ. The model is not only able to predict the activity of new compounds but also explained the important regions in the molecules in a quantitative manner.

Keywords: QSAR; Thrombin inhibitors; Genetic function approximation; Aryl heterocyclic

1. Introduction

Thrombin is a serine protease involved in the conversion of fibrinogen to fibrin [1]. It is one of the important drug targets for cardiovascular disorders like cardiac arrest, ischemic stroke and pulmonary infraction [2]. Older drugs like heparin and warfarin interfere with activity of many clotting factors. These agents suffer from a number of side effects like increased rate of bleeding and need for parenteral administration. In addition these drugs specially warfarin, is involved in a number of drug–drug interactions, which can be problematic in many cases. Antiplatelet drug aspirin does not have a similar thromboprophylactic efficacy to warfarin, with the degree of reduction in risk of stroke being less pronounced [3]. In view of these facts, the agents with specific thrombin inhibitor activity and improved oral bioavailability may be a good alternative for the older drugs. So there is an urgent need to design such agents.

Quantitative structure activity relationships (QSAR) are one of the most important methods in chemometrics, which give information that is useful for drug design and medicinal chemistry [3,4]. A QSAR equation is a mathematical equation that correlates the biological activity to a wide variety of physical or chemical parameters [5]. It started in 1962 when Corwin Hansch first time developed a QSAR model which correlate biological activity to Hammett constant and hydrophobicity [6]. The parameter π which is the relative hydrophobicity of a substituent, was defined by Iwasa et al. in 1964 [7]. Hansch and Fujita combined Hammett and hydrophobic constants to give linear Hansch equation, which later on resulted in the development of Hansch parabolic equation [8]. After that a number of structural, topological, thermodynamic, spatial and other type of descriptors have been developed which can be used for QSAR model generation [9]. There are many examples available in literature in which QSAR models have been used successfully for the screening of compounds for the biological activity [10–13]. Several efforts have been done in past to develop QSAR model for thrombin inhibitors [14–16]. But most of the models lack the high statistical quality. Here we have developed a QSAR model for the aryl heterocycle-based thrombin inhibitors. The behavior of QSAR model is examined with a variety of statistical parameters and the contribution of various descriptors was analyzed.

* Corresponding author.
E-mail address: nilanjanroy@niper.ac.in (N. Roy).

0223-5234/$ - see front matter © 2006 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2006.07.001

2. Experimental

2.1. Data set

The inhibitory activity of the aryl heterocycle-based thrombin antagonist was taken from literature in terms of K_i values [17]. The K_i values were converted to pK_i to get the linear relationship in the equation using following formula:

$$pK_i = -\log K_i$$

Total set of 42 compounds was divided in training and test set of 34 and eight compounds randomly. The structure and actual and predicted activity for both the training and test set compounds are shown in Tables 1–3. Because of the structural uniqueness compounds **1**–**3** and **42** could not accommodated in any table hence shown in Fig. 1.

2.2. Molecular modeling

The X-ray crystal structure of most active compound (**32**) bound with thrombin was extracted from Protein Data Bank (PDB code 1SL3). As the PDB structures do not contain hydrogen atoms, so the hydrogen atoms were attached to structure and it was further energy minimized using semi-empirical AM1 method of energy minimization included in MOPAC 6.0 [18]. Other compounds were built using this structure as template and these structures were also energy minimized using same method.

2.3. Descriptor calculation

E-state indices [19,20], electronic, information content, spatial, structural, thermodynamic and topological descriptors [21] were calculated using the Cerius^2 4.10 software package. Descriptors included in the model are listed and described in Table 4.

2.4. Regression analysis

The total number of descriptors calculated was more than 150 but some of the descriptors were rejected because they contain a value of zero for all the compounds. Further, the inter correlation of descriptors was taken in to account and highly correlated descriptors were grouped together and descriptor with highest correlation with biological activity was taken from the group. From descriptors thus remained, the selection of variables to obtain the QSAR models were carried out using genetic function approximation (GFA) method. GFA is genetics based method of variable selection, which combines Holland's genetic algorithm (GA) with Friedman's multivariate adaptive regression splines (MARS) [22,23]. The GFA method works in the following way: first of all a particular number of equations (set at 100 by default in the Cerius^2 software) are generated randomly. Then pairs of "parent" equations are chosen randomly from this set of 100 equations and "crossover" operations are performed at random. The number of crossing over was set by default at 5000. The goodness of each progeny equation is assessed by Friedman's lack of fit (LOF) score, which is given by following formula

$$\mathrm{LOF} = \mathrm{LSE}/\{1 - (c + dp)/\mathrm{m}\}^2$$

Where LSE is the least-squares error, c is the number of basis functions in the model, d is smoothing parameter, p is the number of descriptors and m is the number of observations in the training set. The smoothing parameter, which controls the scoring bias between equations of different sizes, was set at default value of 1.0 and the new term was added with a probability of 50%. Only the linear equation terms were used for model building, which is set by default in the software. The best equation out of the 100 equations was taken based on the statistical parameters such as regression coefficient, adjusted regression coefficient, regression coefficient cross validation and F-test values.

3. Results and discussion

We first determined the number of descriptors necessary and sufficient for the QSAR equation. Taking a brute force approach, we increased the number of terms in the QSAR equation one by one and evaluated the effect of addition of new term on the statistical quality of model. As the r^2 correlation coefficient can be easily increased by number of terms in the QSAR equation, so we took the cross validation correlation coefficient, q^2 as the limiting factor for number of descriptors to be used in the model. As shown in Fig. 2 the q^2 value increases till the number of descriptors in the equation reached up to 5. When number of descriptors in the equation was 6, there was a decrease in q^2 value of model. So the number of descriptors was restricted to 5. The models with increasing number of descriptors are shown in Table 5 along with the statistical parameters.

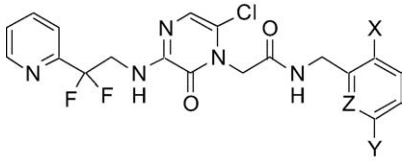
After analyzing equation with five descriptors it was found that compound **18** is an outlier. The reason for **18** being found as outlier is its very low activity. It has highest value of K_i and thus lowest value of pK_i among all the compounds. So it was removed from the training set and a new equation was generated as given under:

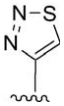
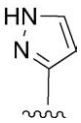
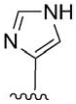
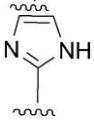
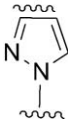
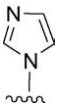
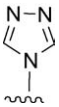
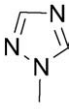
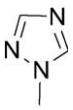
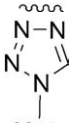
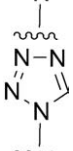
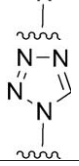
$$pK_i = -5.28169 + 3.32264 * \mathrm{Atype\_N\_69} + 0.04043 * \mathrm{Shadow\_YZ} + 1.86922 * \mathrm{Atype\_Cl\_89} - 0.44212 * \mathrm{LogP} + 0.94893 * \mathrm{Atype\_C\_20} \quad (1)$$

$$N = 33,\ \mathrm{LOF} = 0.114,\ r^2 = 0.959,\ r^2_{adj} = 0.951,\ \mathrm{F-test} = 125.279,\ \mathrm{LSE} = 0.055,\ r = 0.979,\ q^2 = 0.943,\ r^2_{pred}(8) = 0.504,\ r^2_{pred}(7) = 0.947$$

Where N is number of compounds in training set, LOF is lack of fit score, r^2 is squared correlation coefficient, r^2_{adj}. is square of adjusted correlation coefficient, F-test is a variance-related static which compares two models differing by one or

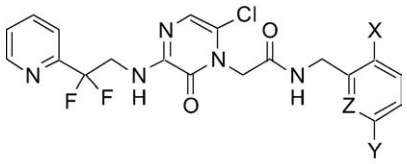
Table 1
Structure, actual and predicted activity of pyrazinone/pyridine analogues used in this study

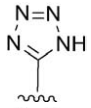
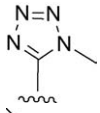
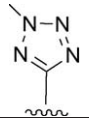


Compound numbers	X	Y	Z	Actual pK_i	Predicted pK_i	Residual
1[b]	–	–	–	–0.477	–0.474	–0.003
2[b]	–	–	–	0.097	0.097	0.000
3[b]	–	–	–	–0.716	–0.830	0.114
4	H	H	CH	–1.079	–1.409	0.330
5[a]	H	Cl	CH	0.356	0.055	0.301
6		H	CH	0.076	0.173	–0.097
7		H	CH	–0.748	–0.656	–0.092
8		H	CH	–0.623	–0.521	–0.102
9		H	CH	–0.857	–0.477	–0.380
10		H	CH	–0.839	–0.677	–0.162
11[a]		H	CH	–0.799	–0.727	–0.072
12		H	CH	–0.544	–0.497	–0.047
13		H	CH	0.347	–0.084	0.431
14		H	N	0.180	0.108	0.072
15[a]		H	CH	1.018	0.618	0.400
16		Cl	CH	2.824	2.363	0.461
17		H	N	0.854	0.969	–0.115

(continued)

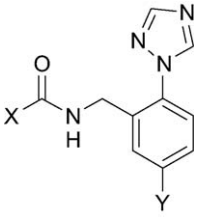
Table 1 *(continued)*

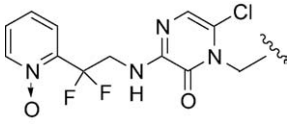
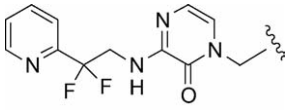
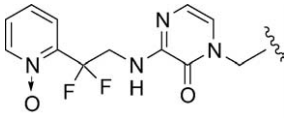
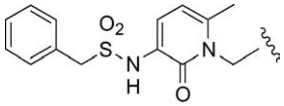
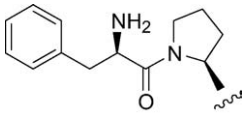
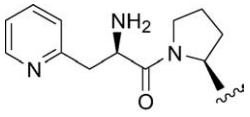


Compound numbers	X	Y	Z	Actual pK_i	Predicted pK_i	Residual
18		H	CH	–2.973	0.642	–0.642
19		H	CH	–0.079	0.409	–0.488
20[a]		H	CH	–0.763	1.107	–1.870

[a] indicates compound taken in test set.
[b] The structure of compounds is given in Fig. 1.

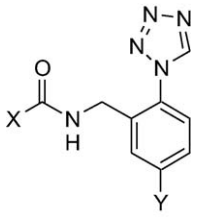
Table 2
Structure, actual and predicted activity of triazole analogues used in this study

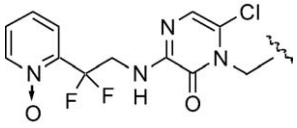
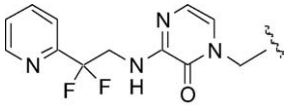
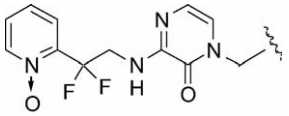
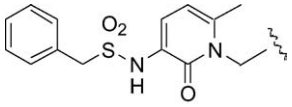
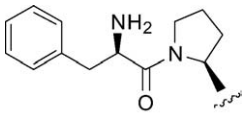
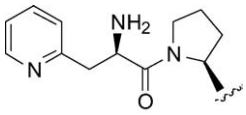


Compound numbers	X	Y	Actual pK_i	Predicted pK_i	Residual
21		H	0.678	0.565	0.113
22		H	–1.204	–1.214	0.010
23		Cl	0.619	0.526	0.094
24[a]		H	–0.857	–0.556	–0.301
25		Cl	1.071	1.199	–0.128
26		H	–0.863	–0.797	–0.066
27		Cl	0.796	0.871	–0.075
28		H	–2.025	–2.226	0.201
29[a]		Cl	–0.255	–0.508	0.253
30		Cl	–0.301	0.097	–0.398

[a] indicates compound taken in test set.

Table 3
Structure, actual and predicted activity of tetrazole analogues used in this study



Compound numbers	X	Y	Actual pK_i	Predicted pK_i	Residual
31[a]		H	1.301	1.277	0.024
32		Cl	2.854	3.044	–0.190
33		H	–0.431	–0.458	0.027
34		Cl	1.481	1.219	0.262
35		H	–0.041	0.147	–0.188
36		Cl	1.886	1.602	0.284
37		H	–0.146	–0.132	–0.014
38[a]		Cl	1.745	1.639	0.106
39		H	–1.146	–1.447	0.301
40		Cl	0.481	0.224	0.257
41		Cl	0.398	0.703	–0.305
42[b]	–	–	–2.398	–2.142	–0.256

[a] indicates compound taken in test set.
[b] The structure of compounds is given in Fig. 1.

more variables to see if the more complex model is more reliable than the less complex one, the model is supposed to be good if the *F*-test is above a threshold value, LSE is least-square error, r is correlation coefficient, q^2 is the square of the correlation coefficient of the cross validation, r^2_{pred} (8) is the predicted correlation coefficient calculated from the predicted activity of all the test set compounds and r^2_{pred} (7) is the predicted correlation coefficient calculated after removing compound **20** as outlier.

Atype_N_69, Atype_Cl_89 and Atype_C_20 are atom-type-based AlogP descriptors. The 120 atom types defined in the calculation of AlogP98 are being used as descriptors. Shadow_YZ is geometric descriptor, which helps to characterize the shape of the molecule. These descriptors are calculated by projecting the molecular surface on three mutual perpendicular planes, XY, YZ and XZ. Shadow_YZ represents the area of molecular shadow in the YZ plane. LogP is the partition coefficient, which represents the lipophilicity of molecule. The negative slope of logP in this equation represents that activity decreases with an increase in lipophilicity of molecule. So substituents, which increase lipophilicity of compound, should be avoided.

Further statistical significance of the relationship between the anti thrombin activity and chemical structure descriptors was demonstrated by randomization procedure. The test was done by repeatedly permuting the activity values of the data set and using the permuted values to generate QSAR models and then comparing the resulting scores with the score of the original QSAR model generated from non-randomized activity values. If the original QSAR model is statistically significant, its score should be significantly better than those from permuted data. The randomized test was performed at 90%, 95%, 98% and 99% confidence interval. The higher the confidence level, the more randomization tests are run. For a 90% confidence level, nine trials are run, 19 trials at 95%, 49 trials at 98% and 99 trials at 99%. The r value of the original model was much higher than any of the trials using permuted data. Hence, model 1 is statistically significant and robust. The results of randomization test at various levels of confidence levels are shown in Table 6.

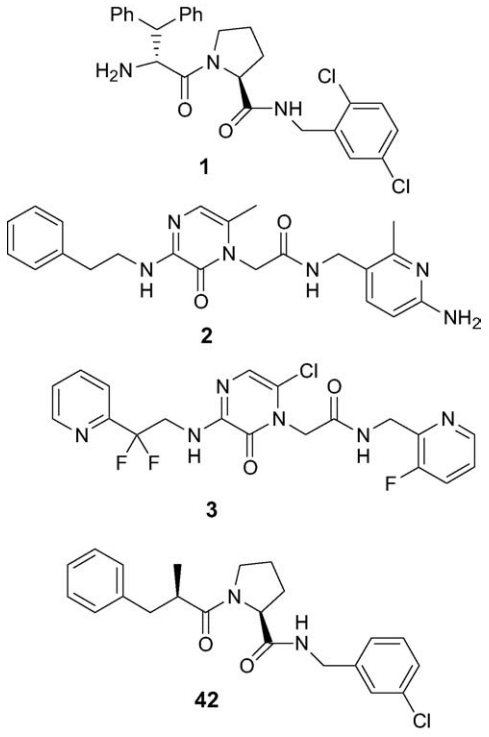


Fig. 1. Structure of compounds **1**–**3** and **42**.

Table 4
Description of the parameters used in the study

Type	Descriptors
E-state indices	Electrotopological-state indices
Electronic	Sum of partial charges, sum of formal charges, Sum of atomic polarizabilities, Dipole moment, Energy of highest occupied orbital, Energy of lowest unoccupied orbital, Superdelocalizability
Information content	Information of atomic composition index, Information indices based on the A-matrix, information indices based on the D matrix, Multigraph information content indices
Spatial	Radius of gyration, Jurs descriptors, Shadow indices, Area, Density, PMI, Vm
Structural	Number of chiral centers, molecular weight, Number of rotatable bonds, Number of hydrogen-bond acceptors, Number of hydrogen-bond donors
Thermodynamic	Log of the partition coefficient, log of the partition coefficient atom-type value, Desolvation free energy of water, Desolvation free energy of octanol, Heat of formation, Molar refractivity
Topological	Wiener index, Zagreb index, Hosoya index, Kier and Hall molecular connectivity index, Balaban indices

Also the inter correlation of the descriptors used in final model was checked and the descriptors were found to be not much correlated. The correlation matrix for the used descriptors is shown in Table 7. To further check the inter correlation of descriptors variance inflation factor (VIF) analysis was performed. VIF value is calculated from $1/1 - r^2$, where r^2 is the multiple correlation coefficient of one descriptor's effect regressed on the remaining molecular descriptors. If VIF value is larger than 10, information of descriptors can be hidden by correlation of descriptors [24,25]. In this model, the

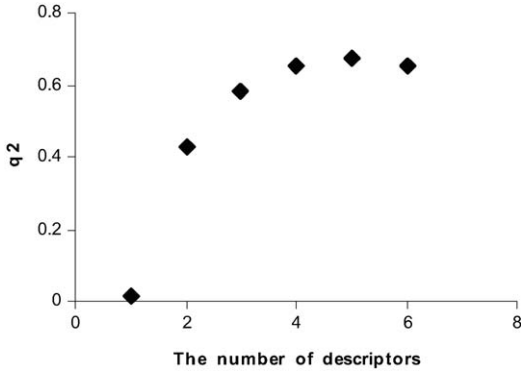


Fig. 2. The cross-validated r^2 (q^2) as a function of number of descriptors.

VIF value of these descriptors are 1.253 (Atype_N_69), 1.189 (Shadow_YZ), 1.567 (Atype_Cl_89), 1.390 (logP) and 1.312 (Atype_C_20). Therefore, from VIF analysis also it is clear that the descriptors used in the final model have very low inter correlation.

The correlation of actual activity to predicted activity is shown in Figs. 3 and 4 for the training and test set compounds, respectively. The predicted activity of test set compounds are very close to their actual activity, which indicate the robustness of model and also indicates that it can be used confidently for predicting the anti thrombin activity of similar compounds.

The QSAR model is also important from the point of suggesting important substitutions in the molecules. Specifically, the atom type descriptors used in the model play important role in this direction and quantitatively suggest which parts of the molecule should not be altered for improving the biological activity. These descriptors developed by Ghose and Crippen [26] use the atomic contribution of individual atom types towards the overall hydrophobicity of molecules. Carbon, hydrogen, oxygen, nitrogen, sulfur and halogens are classified in to 110 atom types. After several revisions the number of atom classifications has increased to 120 [27]. Hydrogen and halogens are classified by the hybridization and oxidation state of the carbon they are bonded to; carbon atoms are classified by their hybridization state and the chemical nature of their neighboring atoms. The complexity of the classification procedure is attested by a total of 44 carbon types alone. The most important descriptor in the model is Atype_Cl_89 as is clear from Table 5. The Atype_Cl_89 corresponds to chlorine atom attached to $C^1{}_{sp2}$ where 1 is oxidation state of carbon and sp^2 is its hybridization state. The descriptor has a positive sign in the model, which indicates this atom has beneficial effect on the activity of compounds. This is also evidenced from the higher activity of compounds with a chlorine substitution on P1 heterocycle. Thus our model predicted the importance of this substitution on biological activity in a quantitative way. Other atype descriptors used in the model are Atype_C_20 and Atype_N_69. Atype_C_20 corresponds to the carbon atom = CX2 where = represent a double bond and X any heteroatom (O, N, S, P, Se, and halogens). As is evidenced from the low activity of compounds **22**–**30** and **33**–**42** compare to their correspond-

Table 5
Statistical assessment of equations with varying number of descriptors

Number of descriptors	Equation	LOF	r^2	r^2_{adj}	F-test	LSE	r	q^2
1	pK_i = -0.459 + 0.982*Atype_Cl_89	1.427	0.184	0.158	7.202	1.264	0.429	0.016
2	pK_i = –1.995 + 1.484*Atype_Cl_89 – 0.368*logP	0.903	0.546	0.517	18.648	0.703	0.739	0.431
3	pK_i = -2.436 + 2.579*Atype_N_69 – 0.447*logP + 1.689 *Atype_Cl_89	0.798	0.651	0.616	18.614	0.541	0.807	0.584
4	pK_i = -2.901 + 0.743*Atype_C_20 – 0.441*logP + 2.006* Atype_Cl_89 + 3.044*Atype_N_69	0.749	0.717	0.678	18.389	0.438	0.847	0.654
5	-4.718 + 0.0323*Shadow_YZ – 0.416*logP + 1.905* Atype_Cl_89 + 3.181*Atype_N_69 + 0.755*Atype_C_20	0.817	0.737	0.690	15.688	0.407	0.858	0.673
6	-4.425 + 0.751786*Atype_C_20 + 1.938* Atype_Cl_89 + 0.029*Shadow_YZ – 0.033*S_aaN – 0.471 *logP +3.182*Atype_N_69	0.958	0.741	0.683	12.864	0.401	0.861	0.655

Table 6
Results of randomization test performed to check the validation of model

Confidence level	90%	95%	98%	99%
Total trials	9	19	48	99
r from non-random	0.979	0.979	0.979	0.979
Random r's > non-random	0	0	0	0
Random r's < non-random	9	19	48	99
Mean value of r from random trial	0.289	0.246	0.310	0.291
Standard deviation of random trials	0.058	0.067	0.109	0.104
Standard deviation from non-random r to mean	11.969	10.966	6.15	6.596

Table 7
Correlation matrix of the descriptors used in equation

	Atype_C_20	Atype_N_69	Atype_Cl_89	Shadow_YZ	logP	Activity
Atype_C_20	1.000					
Atype_N_69	–0.161	1.000				
Atype_Cl_89	–0.435	–0.127	1.000			
Shadow_YZ	–0.081	–0.238	0.225	1.000		
logP	–0.231	0.301	0.330	–0.217	1.000	
Activity	0.120	0.014	0.418	0.438	–0.531	1.000

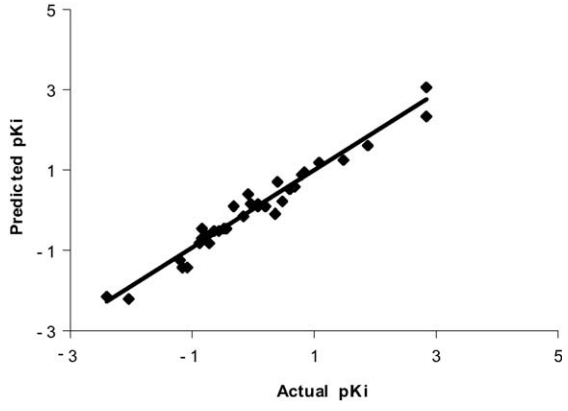


Fig. 3. Correlation of actual activity to the predicted activity of training set compounds by model 1.

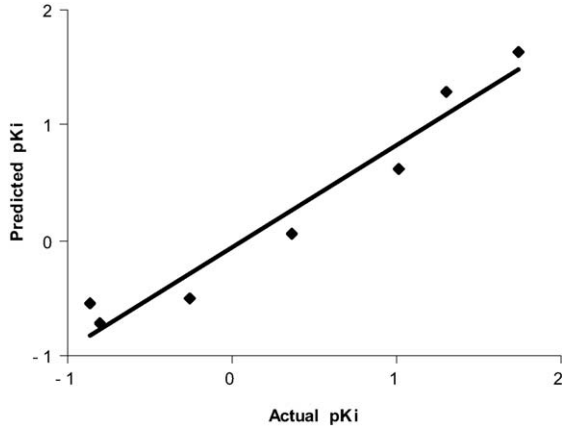


Fig. 4. Correlation of actual activity to the predicted activity for test set compounds by model 1 (compound **20** removed).

ing compounds containing atom corresponding to Atype_C_20 descriptor, the electronic environment of this atom is important for the biological activity of compounds and the developed model is able to predict that in quantitative aspect. Similarly, Atype_N_69 corresponds to the nitrogen atom of type Ar-NH_2 or X-NH_2, where Ar represents the aromatic groups and X any heteroatom (O, N, S, P, Se, and halogens) and this atom type also has a beneficial overall effect on the thrombin inhibitory activity of these compounds. So while designing new inhibitors, these factors should be taken in to account and atom

types corresponding to those predicted important by the model should be included in the new molecules.

4. Conclusion

The QSAR study shows that partition coefficient (logP), partition coefficient-atom type (AlogP98-atom type) and E-state indices are important descriptors responsible for describing the activity of aryl heterocycle-based thrombin inhibitors. The QSAR model is statistically and chemically sound and explains more than 95% of the variance in the experimental activity with excellent predictive power as is evidenced from the predicted activity of test set compounds. The atom type logP descriptors also explain the important atom types in the molecule, which should be taken in to account while designing new inhibitors.

References

[1] S. Anand, Haemostasis 29 (Suppl. S1) (1999) 76–78.
[2] J.M. Maraganore, Thromb. Haemost. 70 (1993) 208–211.
[3] V.J. Marder, Blood Rev. 15 (2001) 143–157.
[4] K. Tuppurainen, Chemosphere 38 (1999) 3015–3030.
[5] C. Hansch, A. Kurup, R. Garg, H. Gao, Chem. Rev. 101 (2001) 619–672.
[6] P.P. Maloney, C. Hansch, T. Fujita, R.M. Muir, Nature 194 (1962) 178–180.
[7] J. Iwasa, T. Fujita, C. Hansch, J. Am. Chem. Soc. 86 (1964) 5175–5180.
[8] C. Hansch, Acc. Chem. Res. 2 (1969) 232–239.
[9] D.J. Livingstone, J. Chem. Inf. Comput. Sci. 40 (2000) 195–209.
[10] L.M. Shi, Y. Fan, T.G. Myers, P.M. O'Connor, K.D. Paull, S.H. Friend, J.N. Weinstein, J. Chem. Inf. Comput. Sci. 38 (1998) 189–199.
[11] S. Oloff, R.B. Mailman, A. Tropsha, J. Med. Chem. 48 (2005) 7322–7332.
[12] A. Meneses-Marcel, Y. Marrero-Ponce, Y. Machado-Tugores, A. Montero-Torres, D.M. Pereira, J.A. Escario, J.J. Nogal-Ruiz, C. Ochoa, V.J. Aran, A.R. Martinez-Fernandez, R.N. Garcia Sanchez, Bioorg. Med. Chem. Lett. 15 (2005) 3838–3843.
[13] L. Santana, E. Uriarte, H. Gonzalez-Diaz, G. Zagotto, R. Soto-Otero, E. Mendez-Alvarez, J. Med. Chem. 49 (2006) 1149–1156.
[14] N. Rameshwar, K. Krishna, B.A. Kumar, T. Parthasarathy, Bioorg. Med. Chem. 14 (2006) 319–325 (Epub 2005 Nov 2004).
[15] D. Labes, V. Hagen, Pharmazie 34 (1979) 554–556.
[16] C.T. Supuran, A. Scozzafava, F. Briganti, B.W. Clare, J. Med. Chem. 43 (2000) 1793–1806.
[17] M.B. Young, J.C. Barrow, K.L. Glass, G.F. Lundell, C.L. Newton, J.M. Pellicore, K.E. Rittle, H.G. Selnick, K.J. Stauffer, J.P. Vacca, P.D. Williams, D. Bohn, F.C. Clayton, J.J. Cook, J.A. Krueger, L.C. Kuo, S.D. Lewis, B.J. Lucas, D.R. McMasters, C. Miller-Stein, B.L. Pietrak, A.A. Wallace, R.B. White, B. Wong, Y. Yan, P.G. Nantermet, J. Med. Chem. 47 (2004) 2995–3008.
[18] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.P. Stewart, J. Am. Chem. Soc. 107 (1985) 3902–3909.
[19] C. de Gregorio, L.B. Kier, L.H. Hall, J. Comput. Aided Mol. Des. 12 (1998) 557–561.
[20] L.B. Kier, L.H. Hall, Pharm. Res. 7 (1990) 801–807.
[21] E. Estrada, E. Uriarte, Curr. Med. Chem. 8 (2001) 1573–1588.
[22] D. Rogers, A.J. Hopfinger, J. Chem. Inf. Comput. Sci. 34 (1994) 854–866.
[23] L.M. Shi, F. Yi, T.G. Myers, P.M. O'Connor, K.D. Paull, S.H. Friend, J.N. Weinstein, J. Chem. Inf. Comput. Sci. 38 (1998) 189–199.
[24] M. Jaiswal, P.V. Khadikar, A. Scozzafava, C.T. Supuran, Bioorg. Med. Chem. Lett. 14 (2004) 3283–3290.
[25] S. Shapiro, B. Guggenheim, Quant. Struct.-, Act. Relat. 17 (1998) 327–337.
[26] A.K. Ghose, G.M. Crippen, J. Comput. Chem. 7 (1986) 565–577.
[27] A.K. Ghose, V.N. Viswanadhan, G.R. Revankar, R.K. Robins, J. Chem. Inf. Comput. Sci. 29 (1989) 163–172.