Available online at www.sciencedirect.com
ScienceDirect
European Journal of Medicinal Chemistry 43 (2008) 293–299
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY
http://www.elsevier.com/locate/ejmech

Original article

Quantitative structure activity relationship studies on thiourea analogues as influenza virus neuraminidase inhibitors

Pramod C. Nair, M. Elizabeth Sobhia*

Centre for Pharmacoinformatics, National Institute of Pharmaceutical Education and Research (NIPER), Sector 67, S.A.S Nagar, Mohali 160062, Punjab, India

Received 9 December 2006; received in revised form 31 January 2007; accepted 15 March 2007
Available online 5 April 2007

Abstract

Influenza virus is a major global threat that impacts the world in one form or another as flu infections. Neuraminidase, one of the targets for these viruses, has recently been exploited in the treatment of these infections. Quantitative structure activity relationship studies were performed on thiourea analogues using spatial, topological, electronic, thermodynamic and E-state indices. Genetic algorithm based genetic function approximation method of variable selection was used to generate the model. Highly statistically significant model was obtained when number of descriptors in the equation was set to 5. The atom type log *P* and shadow indices descriptors showed enormous contributions to neuraminidase inhibition. The validation of the model was done by cross validation, randomization and external test set prediction. The model gives insight on structural requirements for designing more potent analogues against influenza virus neuraminidase.

Keywords: Neuraminidase; Thiourea; Influenza; QSAR; GFA; Descriptors

1. Introduction

The influenza virus is a pathogenic agent for humans and has been socializing in the human population since at least the sixteenth century leading to repeated febrile respiratory disease every year. One of the lethal outbreaks was the Spanish flu caused by influenza (H1N1) virus killing around 40–50 million people [1]. Taubenberger et al. sequenced nine fragments of viral RNA from the coding regions of hemagglutinin, neuraminidase, nucleoprotein, matrix protein 1 and matrix protein 2 isolated from the lung tissue sample of the victim [2]. The sequences were consistent with a novel H1N1 influenza A virus which belongs to the subgroup of strains that infect humans and swine but not the avian subgroup [2]. Even large-scale sequencing has been adopted to provide a more comprehensive picture of the evolution of influenza viruses and of their pattern of transmission through human and animal populations, but still there are many loopholes in understanding its mutational behavior which makes present therapy ineffective [3]. Some pandemics can be seen rapidly progressing to involve all parts of the world due to emergence of a unique virus to which the overall population holds no immunity. These pandemics show characteristics which include occurrence outside the usual season [4], extremely rapid transmission from person to person with concurrent outbreaks throughout the globe and high attack rates in all age groups [5].

Chemotherapy for influenza virus fails due to newly discovered drug resistance in mutant strains. Several probable molecular targets for influenza virus include M2 proteins, endonuclease, hemagglutin and neuraminidase [6–10]. Currently available classes of antivirals include M2 proton channel inhibitors (amantadine and rimantadine) and the neuraminidase inhibitors (zanamivir and oseltamivir) [9,11]. The genetic basis for resistance of M2 proton channel inhibitors is single nucleotide changes and corresponding single amino acid substitutions in the trans-membrane of the M2 ion channel

* Corresponding author. Tel.: +91 172 2214 682; fax: +91 172 2214 692.
E-mail address: mesophia@niper.ac.in (M.E. Sobhia).

0223-5234/$ - see front matter © 2007 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2007.03.020

protein. Clinical studies highlight that the resistant variants emerged rapidly in rimantadine treated patients and these resistant variants were able to spread widely in closed populations when the M2 inhibitors were used for treatment and prophylaxis [12]. The apparent choice remains to be the development of effective drugs which are not susceptible to such mutations.

Release of progeny virus from infected cells by cleaving sugars which bind the mature viral particles is facilitated by an enzyme called neuraminidase. Targeting influenza virus neuraminidase was one of the breakthroughs in developing agents against flu; developing such agents provides valuable perspectives applicable to the field of antiviral chemotherapy. Crystal structure of influenza A neuraminidase was solved in the year 1983 [13] and its complex with its natural substrate sialic acid was reported in 1992 [14]. Selective and potent neuraminidase inhibitors were achieved by structure-based drug design, a well known rational drug design approach [15]. Oseltamivir is the drug of choice for treating influenza in current scenario but it has been reported to cause vomiting and nausea. Tremendous antiviral activity is displayed by zanamivir when administered intranasally but it is less effective when delivered systemically. So, there is still an enormous need to design and identify new agents for the chemotherapy of influenza virus infection and formulate effective drugs for systemic administration.

Quantitative Structure Activity Relationships (QSAR) for different sets of compounds have been reported by Verma and Hansch [16]. Seventeen different QSAR equations to understand chemical–biological interactions governing their activities toward influenza neuraminidase were presented in their work. Out of 17, 8 contain a correlation between activity and hydrophobicity indicating that it is one of the most important properties to look upon while designing drug for flu virus [16]. Few of the equations also suggest bilinear relationship between hydrophobicity and activity i.e. activity is decreased if the hydrophobicity is increased after a certain limit [16]. We report here different QSAR models developed on newly synthesized thiourea analogues reported to be potent inhibitors of influenza virus neuraminidase. Genetic algorithm based genetic function approximation method of variable selection was used to generate these models. The best model was used for predicting the test molecules which were not included in the training set. Randomization test at various intervals of confidence levels was done to validate the final model. The model gave good insights for developing new analogues as influenza virus neuraminidase inhibitors.

2. Experimental

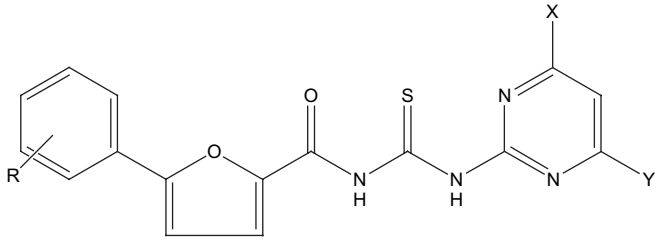
2.1. Dataset

In vitro inhibitory activity data (IC_{50} μM) of the thiourea analogues on influenza virus neuraminidase, reported by Sun et al. were taken for the study [17]. IC_{50} values were measured spectrofluorometrically using 2′-(4-methylumbelliferyl)-α-D-acetylneuraminic acid as substrate for neuraminidase to yield a fluorescent product which was quantified [17]. Out of 40 thiourea analogues reported, 30 molecules were selected for developing the model and 10 molecules for which the precise data were not available were discarded for the analysis. The IC_{50} (μM) values were taken in molar (M) range and converted to pIC_{50} according to the formula.

$$pIC_{50} = -\log IC_{50}$$

The dataset was randomly segregated into training and test sets comprising 24 and 6 molecules, respectively (Tables 1–3).

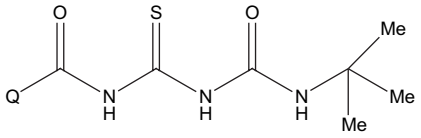
Table 1
Structure with actual and predicted activities of polysubstituted pyrimidinyl acyl(thio)urea analogues



Molecules	R	X	Y	IC_{50} (μM)	Actual pIC_{50} (M)	Predicted pIC_{50} (M)	Residual
1[a]	2-Cl	OEt	Me	1.65	5.78	6.69	−0.91
2	2-Cl	OEt	OEt	0.08	7.10	6.96	0.14
3	2-Cl	OH	Me	0.32	6.49	6.08	0.41
4	2-Cl	OMe	OMe	1.77	5.75	5.85	−0.10
5[a]	2-Cl	Cl	Cl	14.5	4.84	6.20	−1.36
6[b]	2-Cl	Cl	OEt	>20	–	–	–
7[b]	2-Cl	OMe	Me	>20	–	–	–
8	4-NO_2	Cl	Cl	1.66	5.78	5.85	−0.07
9[a]	4-NO_2	OEt	Me	2.30	5.64	6.41	−0.77
10	4-NO_2	OH	Me	0.36	6.44	5.87	0.57
11[b]	4-NO_2	OEt	OEt	>20	–	–	–

[a] Molecules used in test set.
[b] Molecules of discarded set.

Table 2
Structure with actual and predicted activities of *tert*-butylaminocarbonyl acyl(thio)urea analogues



Molecules	Q	IC_{50} (μM)	Actual pIC_{50} (M)	Predicted pIC_{50} (M)	Residual
12	5-(2-Cl-Ph)-2-furyl	1.42	5.85	6.25	−0.40
13	5-(4-NO_2-Ph)-2-furyl	1.30	5.89	5.87	0.02
14[a]	Ph	1.79	5.75	5.50	0.25
15	OMe	1.83	5.74	5.53	0.21
16	(2,4-Cl_2-Ph)-OCH_2	1.67	5.78	5.82	−0.04
17	2,6-F_2-Ph	1.43	5.84	5.91	−0.07
18	*S*-(+) 2-Me-1-(4-Cl-Ph)-Pr	1.35	5.87	5.94	−0.07
19	*cis*-(−) CFPC[b]	0.51	6.29	6.18	0.11
20[a]	*trans*-(−) DCPC[c]	0.26	6.59	6.46	0.13

[a] Molecule used in test set.
[b] CFPC, 3-(2-chloro-3,3,3-trifluoropropenyl)-2,2-dimethyl cyclopropyl.
[c] DCPC, 3-(2,2-dichloro ethenyl)-2,2-dimethyl cyclopropyl.

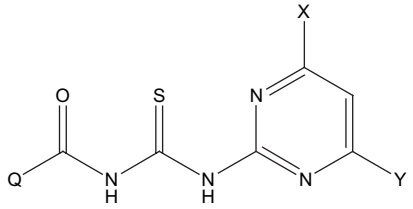
2.2. Molecular modeling

Molecular modeling studies were performed using the molecular modeling package SYBYL 7.1 [18] installed on a Silicon Graphics Fuel Work station. The most active molecule in the training set was chosen as the template and the rest of the molecules were derived from it. Tripos force field and Gasteiger–Huckel partial atomic charges were used for minimizing the molecules [19]. Powell's conjugate gradient method was used for minimization [20]. The minimum energy difference of 0.001 kcal/mol was set as a convergence criterion.

2.3. Descriptor calculation

Descriptors were calculated using the Cerius^2 4.9 software package [21]. Different descriptor classes such as E-state indices, electronic, spatial, structural, thermodynamic and topological descriptors were calculated for the molecules in the dataset [22–24]. More than 140 descriptors were calculated and some were rejected because they contain a value of zero for all the compounds. The inter-correlation of descriptors was taken into account and highly correlated descriptors were grouped together and the descriptor with the highest correlation with biological activity was taken from the group. Description of descriptors included in the model is listed in

Table 3
Structure with actual and predicted activities of aryl and chrysanthemoyl thiourea analogues



Molecules	Q	X	Y	IC_{50} (μM)	Actual pIC_{50} (M)	Predicted pIC_{50} (M)	Residual
21	5-(4-NO_2-Ph)-2-furyl	OMe	Me	1.22	5.91	5.81	0.10
22	5-(2-Cl-Ph)-2-furyl	OMe	Cl	1.29	5.89	5.91	−0.02
23	6-Cl-3-py	Me	OH	8.58	5.07	5.42	−0.35
24[c]	6-Cl-3-py	OMe	Me	>20	–	–	–
25[c]	6-Cl-3-py	OMe	Cl	>20	–	–	–
26[c]	6-Cl-3-py	OEt	OEt	>20	–	–	–
27[c]	2-Cl-3-py	OEt	OEt	>20	–	–	–
28[c]	2-Cl-3-py	OEt	Cl	>20	–	–	–
29	2-Cl-3-py	Me	Me	7.19	5.14	5.19	−0.05
30	2-Cl-3-py	OMe	Cl	2.59	5.59	5.35	0.24
31[c]	3-py	OMe	OMe	>20	–	–	–
32	5,6-Cl_2-3-py	OMe	OMe	18.5	4.73	4.88	−0.15
33[a]	Ph	Me	Me	2.10	5.68	5.95	−0.27
34	2-Me-1-(4-Cl-Ph)-Pr	OEt	OEt	0.31	6.51	6.69	−0.18
35[c]	2-Me-1-(4-Cl-Ph)-Pr	Cl	Cl	>20	–	–	–
36	CFPC[b]	OMe	OMe	0.97	6.01	5.93	0.08
37	CFPC[b]	Me	Me	0.58	6.24	6.44	−0.20
38	2-F-4-Cl-Ph	Me	Me	1.36	5.87	5.97	−0.10
39	2-F-4-Cl-Ph	OMe	Cl	5.10	5.29	5.32	−0.03
40	(2,4-Cl_2-Ph)-OCH_2	OMe	OMe	1.89	5.72	5.81	−0.09

[a] Molecules used in test set.
[b] CFPC, 3-(2-chloro-3,3,3-trifluoropropenyl)-2,2-dimethyl cyclopropyl.
[c] Molecules of discarded set.

Table 4
Descriptors and their types used in the study

Type	Descriptors
Structural	Molecular weight, number of chiral centers, number of rotatable bonds, number of hydrogen-bond acceptors, number of hydrogen-bond donors
Electronic	Sum of atomic polarizabilities, sum of partial charges, sum of formal charges, dipole moment, energy of highest occupied orbital (HOMO), energy of lowest unoccupied orbital (LUMO), superdelocalizability
Topological	Kier and Hall molecular connectivity index, Wiener index, Zagreb index, Hosoya index, Balaban indices
E-state indices	Electrotopological-state indices
Spatial	Jurs descriptors, radius of gyration, PMI, area, shadow indices, density, V_m
Thermodynamic	Molar refractivity, heat of formation, log of the partition coefficient, log of the partition coefficient atom type value, desolvation free energy of water, desolvation free energy of octanol

Table 4. The descriptors which remained after removing the ones which had zero values for all compounds were subjected to genetic function approximation (GFA) for selection of variables to obtain the QSAR models using genetic algorithm principles.

2.4. Regression analysis

GFA is a genetics based method of variable selection, which combines Holland's genetic algorithm (GA) with Friedman's multivariate adaptive regression splines (MARS) [25,26]. The GFA method works by generating equations (set at 100 by default in the Cerius^2 software) randomly. Then pairs of ''parent'' equations are chosen for ''crossover'' operations from this set of 100 equations randomly. 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 described by the following formula:

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

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 [25]. The smoothing parameter that 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. 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

The plot between actual and predicted pIC50 values for training set is shown in Fig. 1 and the histogram for residual is shown in Fig. 2. Number of descriptors necessary and sufficient for the QSAR equation was first determined. A brute force approach was followed to increase the number of terms in the QSAR equation one by one and the effect of addition of new term on the statistical quality of the model. As the square correlation coefficient (r^2) can be easily increased by number of terms in the QSAR equation, the cross validation square correlation coefficient (q^2) as the limiting factor for number of descriptors was used in the model. As shown in Fig. 3 the q^2 value increases till the number of descriptors in the equation reaches up to 5 and q^2 value starts decreasing as the number of descriptors increases further. Thus, 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. The molecule **5** in the dataset occasionally disturbed the robustness of the model and it was kept in the test set. The same molecule **5** turned out to be an outlier with a residual of −1.36 (Fig. 2). The probable reason for the high residual of **5** is due to its very low activity ($\mathrm{pIC_{50}} = 4.84$) in comparison to other compounds.

$$\begin{aligned}\mathrm{pIC_{50}} = {} & 3.2747 + 3.5165 * \mathrm{shadow\_XYfrac} - 0.9702 \\ & * \mathrm{Atype\_C\_25} + 0.0098 * \mathrm{shadow\_XY} + 0.1182 \\ & * \mathrm{Atype\_H\_47} - 0.6835 * \mathrm{Atype\_C\_5} \end{aligned} \quad (1)$$

$$N = 24;\ \mathrm{LOF} = 0.106;\ r^2 = 0.854;\ r^2_{\mathrm{adj}} = 0.814;$$
$$F\text{-test} = 21.081;\ \mathrm{LSE} = 0.036;\ r = 0.924;\ q^2 = 0.748$$

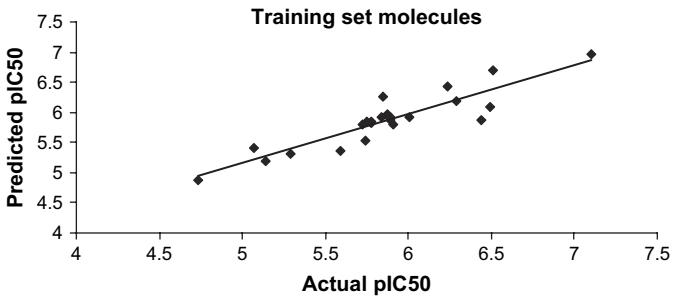


Fig. 1. Plot of actual versus predicted $\mathrm{pIC_{50}}$ of the training set molecules.

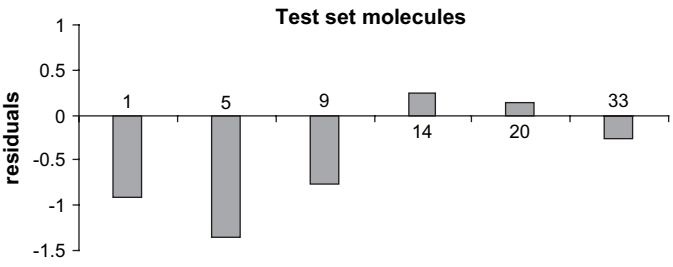


Fig. 2. Histogram of residuals of the test set molecules.

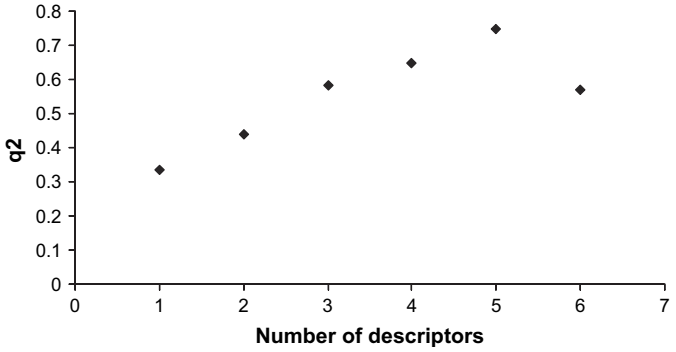


Fig. 3. Plot of cross-validated r^2 (q^2) as a function of number of descriptors.

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

Further statistical significance of the relationship between the antiviral activity and chemical structure descriptors was demonstrated by randomization procedure. The randomization test was done by repeatedly permuting the activity values of the dataset 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, then its score should be significantly better than those from permuted data. The randomization test was performed at different confidence intervals at 90%, 95%, 98% and 99%. There are more randomization tests run for higher confidence levels. For a 90% confidence level, there are 9 trials run, 19 trials for 95%, 49 trials for 98% and 99 trials for 99%. The r value of the original model was much higher than any of the trials using permuted data as shown in Table 6. Hence, it shows that the model developed is statistically significant and robust. The inter-correlation of the descriptors used in the final model was checked and the descriptors were found to be reasonably orthogonal. The correlation matrix based on correlation coefficient for the descriptors used in the final model is shown in Table 7.

As displayed in Table 5, all of the equations in different models show presence of either a spatial descriptor or a thermodynamic descriptor or both indicating that these descriptors play an important role in determining the neuraminidase inhibitory activity. The best model with 5 descriptors shows shadow_XY and shadow_XY fraction. Shadow indices represent the area of the projections of the molecular volume in three perpendicular planes: XY, YZ, XZ. Shadow_XY descriptor is calculated by projecting the molecular surface on XY plane while the shadow_XY fraction is a geometric spatial descriptor related to the breadth of a molecule. Docking studies by Sun et al. suggest that most of the active compounds shared similar binding modes when docked to neuraminidase, highlighting the importance of their spatial pose in the neuraminidase inhibition. Shadow indices are a set of geometric descriptors which characterize the shape of the molecules based on their conformation and their spatial orientation. Shadow_XY and shadow_XY fraction show a positive contribution to anti influenza (H1N1) activity suggesting that the inhibitory effect can be enhanced by increasing these values.

As Verma and Hansch analyzed that hydrophobicity is one of the important properties to be taken into account while designing anti influenza agents, different models developed by us also showed the presence of atom type descriptors which characterize hydrophobicity. Various atom type AlogP descriptors were used to calculate the log P of molecules. Ghose and Crippen [27] used the atomic contribution of individual atom types toward the overall hydrophobicity of molecules where carbon, hydrogen, oxygen, nitrogen, sulfur and halogens were classified into 120 atom types [28]. In this strategy halogens and hydrogen are classified by the hybridization and oxidation state of the carbon they are bonded to; carbon atoms were classified by their hybridization state and the chemical nature of their neighboring atoms. There are 44 carbon types alone classified by this system. In the present study Atype_C_5 is one of the

Table 5
Statistical evaluation of equations with varying number of descriptors

Descriptor	Equation	LOF	r2	r2adj	F-test	LSE	r	q2
1	$pIC_{50} = 4.2522 + 0.0005$ * Jurs-DPSA-2	0.164	0.443	0.418	17.522	0.137	0.666	0.335
2	$pIC_{50} = 3.5535 + 2.6815$ * Jurs-FPSA-1 + 0.0005 * Wiener	0.158	0.557	0.515	13.195	0.109	0.746	0.441
3	$pIC_{50} = 1.19026 - 0.3162$ * Atype_C_5 + 5.4668 * shadow_XYfrac + 0.0132 * Zagreb	0.167	0.621	0.564	10.906	0.094	0.788	0.582
4	$pIC_{50} = 5.9117 + 0.1736$ * Atype_H_47 − 0.7913 * Atype_C_5 − 0.9750 * Atype_C_25 + 0.00001 * Jurs-SASA	0.124	0.777	0.729	16.507	0.055	0.881	0.650
5	$pIC_{50} = 3.2747 + 3.5165$ * shadow_XYfrac − 0.9702 * Atype_C_25 + 0.0098 * shadow_XY + 0.1182 * Atype_H_47 − 0.6835 * Atype_C_5	0.106	0.854	0.814	21.081	0.036	0.924	0.748
6	$pIC_{50} = 5.906 - 0.2294$ * Atype_N_75 + 0.0101 * Atype_C_13 + 0.0008 * Wiener − 0.0007 * PMI-mag − 0.2423 * Atype_C_5 − 0.0687 * Jurs-FNSA-1	0.212	0.785	0.710	10.367	0.053	0.886	0.569

Table 6
Statistical results of randomization test performed to check the validation of model

Confidence level	90%	95%	98%	99%
Total trials	9	19	49	99
r From non-random	0.924	0.924	0.924	0.924
Random r's > non-random	0	0	0	0
Random r's < non-random	9	19	49	99
Mean value of r from random trial	0.4525	0.4202	0.4234	0.4556
Standard deviation of random trials	0.1386	0.1172	0.1350	0.1214
Standard deviation from non-random r to mean	3.4044	4.2997	3.7087	3.8611

atom type AlogP descriptors which was chosen by genetic function approximation method of variable selection indicating its crucial role in predicting antiviral activity (Eq. (1)). Many of the equations developed in GFA runs showed its presence (Table 5). The negative slope of Atype_C_5 in this equation represents that activity decreases with an increase in lipophilicity related to C_5 atom type for these inhibitors. The atom type C_5 is C in CH_3X where X represents any heteroatom (O, N, S, and halogens).

Another atom type descriptor Atype_C_25 showed a negative contribution to anti influenza activity. Here the C is part of R- -CR- -R, where R represents any group linked through carbon and "- -" represents aromatic bonds as in benzene or delocalized bonds as the N−O bond in nitro group. Hydrophobicity associated with C atom as part of the aromatic ring or N−O bonded in nitro group is unfavorable for neuraminidase inhibition. A bilinear relationship between hydrophobicity and activity was discovered by Verma and Hansch for different classes of anti-flu agents, indicating that it is not always the case that the increase in the hydrophobicity of molecule will behave linearly with the activity. Atype_H_47 shows a positive slope representing that its presence is essential for these analogues to inhibit neuraminidase of the flu virus. In the H atom type 47 descriptor, the hydrogen atom corresponds to H bonded to carbon $C^1_{sp^3}$, $C^0_{sp^2}$ where 0 and 1 represent the formal oxidation numbers while sp^2 and sp^3 denote their hybridization states. The hydrophobicity associated with atom type H with sp^3 carbon having a formal oxidation number of 1 and atom type H with sp^2 carbon having a formal oxidation number of 0 is favorable for neuraminidase inhibition. Thus it can be seen that atomwise contribution for hydrophobicity plays a role in determining the activity against influenza virus. Appropriate presence and absence of such atom type descriptors will probably help to modulate the hydrophobic character of the molecule, giving suggestions to design more potent agents against influenza infections.

Table 7
The correlation matrix for the descriptors used in the final model

	Shadow_ XYfrac	Shadow_ XY	Atype_ H_47	Atype_ C_5	Atype_ C_25	pIC_{50}
Shadow_XYfrac	1					
Shadow_XY	0.144	1				
Atype_H_47	0.389	0.571	1			
Atype_C_5	0.199	−0.123	0.555	1		
Atype_C_25	0.021	0.523	0.264	−0.308	1	
pIC_{50}	0.370	0.425	0.192	−0.376	−0.138	1

4. Conclusion

The QSAR model was developed with spatial, topological, electronic, thermodynamic and E-state indices on 30 thiourea analogues. The model developed was validated by cross validation techniques, randomization and external test set prediction. The spatial and thermodynamic descriptors were found to play a major role in determining inhibitory activity for neuraminidase. The shadow indices highlight the spatial importance for designing inhibitors for neuraminidase. The atom type log P descriptors explain the hydrophobic contributions of different atom types which must be taken into account while designing new inhibitors against neuraminidase. The atomwise contributions to hydrophobicity will probably help to appropriately take into account those atom types which are essential for determining neuraminidase inhibitory activity and come out with potent leads against influenza viruses.

References

[1] K.D. Patterson, G.F. Pyle, Bull. Hist. Med. 65 (1991) 4–21.
[2] J.K. Taubenberger, A.H. Reid, A.E. Krafft, K.E. Bijwaard, T.G. Fanning, Science 275 (1997) 1793–1796.
[3] E. Ghedin, N.A. Sengamalay, M. Shumway, J. Zaborsky, T. Feldblyum, V. Subbu, D.J. Spiro, J. Sitz, H. Koo, P. Bolotov, D. Dernovoy, T. Tatusova, Y. Bao, K. St George, J. Taylor, D.J. Lipman, C.M. Fraser, J.K. Taubenberger, S.L. Salzberg, Nature 437 (2005) 1162–1166.
[4] M.F. Boni, J.R. Gog, V. Andreasen, M.W. Feldman, Proc. Biol. Sci. 273 (2006) 1307–1316.
[5] N.J. Cox, K. Subbarao, Annu. Rev. Med. 51 (2000) 407–421.
[6] A.J. Hay, A.J. Wolstenholme, J.J. Skehel, M.H. Smith, EMBO J. 4 (1985) 3021–3024.
[7] J.C. Hastings, H. Selnick, B. Wolanski, J.E. Tomassini, Antimicrob. Agents Chemother. 40 (1996) 1304–1307.
[8] M. Mammen, G. Dahmann, G.M. Whitesides, J. Med. Chem. 38 (1995) 4179–4190.
[9] P.M. Colman, Expert Rev. Anti Infect. Ther. 3 (2005) 191–199.
[10] J. Wilson, M. von Itzstein, Curr. Drug Targets 4 (2003) 389–408.
[11] A. Moscona, N. Engl. J. Med. 353 (2005) 1363–1373.
[12] F.G. Hayden, R.B. Belshe, R.D. Clover, A.J. Hay, M.G. Oakes, W. Soo, N. Engl. J. Med. 321 (1989) 1696–1702.
[13] P.M. Colman, J.N. Varghese, W.G. Laver, Nature 303 (1983) 41–44.
[14] J.N. Varghese, J.L. McKimm-Breschkin, J.B. Caldwell, A.A. Kortt, P.M. Colman, Proteins 14 (1992) 327–332.
[15] M. von Itzstein, W.Y. Wu, G.B. Kok, M.S. Pegg, J.C. Dyason, B. Jin, P.T. Van, M.L. Smythe, H.F. White, S.W. Oliver, Nature 363 (1993) 418–423.
[16] R.P. Verma, C. Hansch, Bioorg. Med.Chem. 14 (2006) 982–996.
[17] C. Sun, X. Zhang, H. Huanga, P. Zhoua, Bioorg. Med. Chem. 14 (2006) 8574–8581.
[18] SYBYL Molecular Modeling System, version 7.1, Tripos Inc., St. Louis, MO, 63144-2913.
[19] J. Gasteiger, M. Marsili, Tetrahedron 36 (1980) 3219–3228.
[20] M.J.D. Powell, Math. Program. 12 (1977) 241–254.
[21] Cerius^2 version 4.9, Accelrys Inc., 6985 Scranton Road, San Diego, CA, USA.

[22] C. de Gregorio, L.B. Kier, L.H. Hall, J. Comput. Aided Mol. Des. 12 (1998) 557–561.
[23] L.B. Kier, L.H. Hall, Pharm. Res. 7 (1990) 801–807.
[24] E. Estrada, E. Uriarte, Curr. Med. Chem. 8 (2001) 1573–1588.
[25] D. Rogers, A.J. Hopfinger, J. Chem. Inf. Comput. Sci. 34 (1994) 854–866.
[26] 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.
[27] A.K. Ghose, G.M. Crippen, J. Chem. Inf. Comput. Sci. 7 (1987) 21–35.
[28] A.K. Ghose, V.N. Viswanadhan, G.R. Revankar, R.K. Robins, J. Chem. Inf. Comput. Sci. 29 (1989) 163–172.