RESEARCH ARTICLE

QSAR STUDIES, AND IN SILICO ADME PREDICTION OF P-AMINOSALICYLIC ACID DERIVATIVES AS NEURAMINIDASE INHIBITORS

Mithilesh Kumar Dwivedi*, Purushottam Das Soni#, Shailja Sachan⊥, Santosh Tiwari•

*Department of Chemistry,Holkar Science College,Indore,(M.P.),INDIA

#Department of Chemistry,M.G.C.G.University,Chitrakoot,Satna(M.P.),INDIA

⊥Department of Chemistry,M. S. Golvalkar College,Rewa ,(M.P.),INDIA

•Department of Chemistry,A. P. S. University,Rewa (M.P.),INDIA

*Corresponding Author's E-mail: dwivedimk12@gmail.com

ABSTRACT:

QSAR analysis on a set of synthesized p-Aminosalysilic Acid derivatives analogues tested growth inhibitory antiviral activity was performed by using MLR procedure. The activity contribution of these compounds were determined from regression equation and the validation procedures to analyze the predictive ability of QSAR models were described.The results are discussed on the basis of statistical data. High agreement between experimental and predicted antiviral activity inhibitory values are obtained. The results revealed the significant roles of topological, geometrical and substituent electronic descriptor parameters on the inhibitory activitypIc$_{50}$ of coumarin derivative analogues of the studied molecules

Key Words: QSAR, Antiviral Activity, MLR

1 INTRODUCTION

Influenza is worldwide one of the deadliest infectious diseases that can affect millions of people every year [1]. Hemagglutinin (HA) and neuraminidase (NA) are essential for viral replication, infectivity, and the infective cycle of influenza. NA catalyzes removal of terminal sialic acid (SA) linked to glycoproteins and glycolipids. Scientific research suggested that NA is not only crucial in the release of virion progeny away from infected cells [2], but also important in the movement of the virus through mucus of respiratory tract and reducing the propensity of the virus particles to aggregate. Despite the homology identity of NA in different strains is only about 30%, the catalytic site of NA in all influenza A and B virus is completely conserved [3]. Therefore, inhibitors that were designed to block the active site of NA could become an efficient treatment against the influenza. Rational design of neuraminidase inhibitors (NAIs) is now in the clinic and is effective for the treatment of influenza. Recently, research of structure-based NAIs is becoming an interesting field, leading to a breakthrough in the control of influenza [4]. The NA active site is highly polar, with ten Arg, Asp angd Glu residues and four hydrophobic residues [5]. To facilitate the discussion of the inhibitor-binding modes, the active site of NA is divided into five regions: S1–S5 (Figure 1). These subsites are diagrammed and numbered in counterclockwise fashion using the crystal structure of the substrate-based inhibitor dehydrodeoxy-N-acetylneuraminic acid (DANA) bound to the active site [6]. However, NAIs based on the structure of DANA were found to be rapidly excreted from physiological systems. Therefore, Jedrzejas et al. [7] developed a novel series of specific NAIs that were benzoic acid derivatives on the basis of NA-DANA complex. They found that the aromatic NAIs occupied the same site as DANA in the active site of NA. The coordinates of such inhibitors complexed with NA could be used as the starting model for further design of more potent benzoic acid NAIs. Over the past years, we have engaged in a project aimed at developing novel NAIs as antiviral agents. We have described the design and synthesis of a novel series of pyrrolidine derivatives as potent NAIs [8]. As a further development of this project, we used benzene ring as a replacement of the dihydropyran ring of DANA (Figure 2). The benzene ring scaffold had advantages of non-chirality, chemical and metabolic stability, and increased lipophilicity compared with the dihydropyran ring. In our previous work, we have described the synthesis and evaluation of a novel class of p-aminosalicylic acid (PAS) derivatives as potent influenza NAIs (Figure 2) [9]. Molecular docking, quantitative structure–activity relationship (QSAR) studies and pharmacokinetics absorption, distribution, metabolism, and excretion (ADME) prediction for any given scaffold of interest are the most popular methods

of computer aided drug design [10]. In our ongoing project, molecular docking, 3D-QSAR studies, and in silico ADME prediction were carried out to better understand the structural basis for PAS NAIs. Herein, Surflex-Dock, comparative molecular field analysis (CoMFA), comparative molecular similarity indices analysis (CoMSIA) studies, and Volsurf analysis of pharmacokinetic properties of 40 PAS NAIs are described.
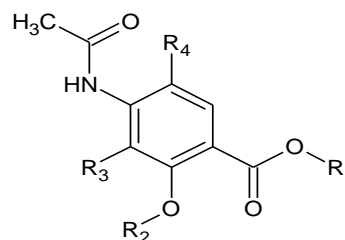

Figure 1: The chemical structure of compound

Table 1: Structure and activity of compounds

S.No	Compound	R_1	R_2	R_3	R_4	pIC50
1	A1	Me	Et-	H	H	6.658
2	A2	Me	Et-	NO_2	NO_2	5.939
3	A3	Me	Et-	NH_2	NH_2	6.509
4	A4	Me	Et-	$=R_4$	$N=C(NH_2)_2$	7.444
5	A5	H	Et-	$=R_4$	$N=C(NH_2)_2$	7.495
6	B1	Me	i-pr	H	H	5.274
7	B2	Me	i-pr	H	NO_2	6.244
8	B3	Me	i-pr	H	NH_2	6.638
9	B4	Me	i-pr	H	$N=C(NH_2)_2$	6.921
10	B5	H	i-pr	H	$N=C(NH_2)_2$	7.309
11	C1	Me	n-pr	H	H	5.033
12	C2	Me	n-pr	H	NO_2	5.447
13	C3	Me	n-pr	H	NH_2	5.899
14	C4	Me	n-pr	H	$N=C(NH_2)_2$	5.996
15	C5	H	n-pr	H	$N=C(NH_2)_2$	6.143
16	D1	Me	s-Bu	H	H	5.527
17	D2	Me	s-Bu	H	NO_2	5.799
18	D3	Me	s-Bu	H	NH_2	6.022
19	D4	Me	s-Bu	H	$N=C(NH_2)_2$	6.131
20	D5	H	s-Bu	H	$N=C(NH_2)_2$	7.143
21	E1	Me	n-Bu	H	H	6.638
22	E2	Me	n-Bu	H	NO_2	6.854
23	E3	Me	n-Bu	H	NH_2	7.131
24	E4	Me	n-Bu	H	$N=C(NH_2)_2$	7.284
25	E5	H	n-Bu	H	$N=C(NH_2)_2$	7.409
26	F1	Me	isopentil	H	H	5.419
27	F2	Me	isopentil	H	NO_2	5.529
28	F3	Me	isopentil	H	NH_2	6.013
29	F4	Me	isopentil	H	$N=C(NH_2)_2$	6.745
30	F5	H	isopentil	H	$N=C(NH_2)_2$	7.387
31	G1	Me		H	H	6.308
32	G2	Me		H	NO_2	6.48
33	G3	Me		H	NH_2	6.74
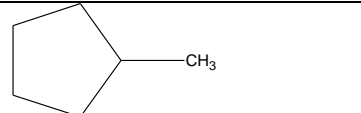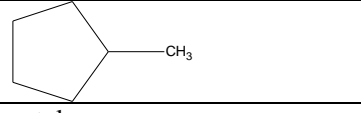
34	G4	Me	(cyclopentyl–CH_3)	H	$N=C(NH_2)_2$	7.032
35	G5	H	(cyclopentyl–CH_3)	H	$N=C(NH_2)_2$	7.252
36	H1	Me	cetyl	H	H	5.57
37	H2	Me	cetyl	H	NO_2	5.727
38	H3	Me	cetyl	H	NH_2	5.995
39	H4	Me	cetyl	H	$N=C(NH_2)_2$	6.059
40	H5	H	cetyl	H	$N=C(NH_2)_2$	6.159

Table 2: Calculated Topological descriptors used in QSAR/QSPR Modeling

Compd. no.	pIC50	MR	MV	Parachor	IR	Surface T	Density	X2	X3
1	6.658	63.6	200.9	512.3	1.545	42.2	1.18	6.907	4.989
2	5.939	76.7	224.6	623.2	1.598	59.3	1.457	9.76	7
3	6.509	72.08	205.4	563.9	1.618	56.7	1.3	7.867	6.016
4	7.444	85.26	235.6	655.1	1.643	59.7	1.49	10.714	7.088
5	7.495	80.3	211.2	610.5	1.685	69.7	1.59	10.573	6.57
6	5.274	66.48	210.4	530.9	1.544	40.7	1.117	7.266	5.1
7	6.244	73.02	222.3	586.3	1.57	48.3	1.26	8.726	6.136
8	6.638	70.71	212.7	556.7	1.579	46.8	1.176	7.796	5.579
9	6.921	76.8	230	594.3	1.582	44.5	1.27	9.208	6.146
10	7.309	71.84	205.6	549.7	1.615	51	1.35	9.066	5.627
11	5.033	68.24	217.4	552.1	1.54	41.5	1.155	7.26	5.258
12	5.447	74.78	229.2	607.5	1.566	49.3	1.292	8.721	6.295
13	5.899	72.47	219.9	577.9	1.573	47.8	1.212	7.79	5.737
14	5.996	78.38	237.4	621	1.574	46.8	1.29	9.202	6.304
15	6.143	73.41	213	576.4	1.605	53.6	1.38	9.061	5.786
16	5.527	72.83	234.3	589.2	1.534	40	1.132	7.867	5.731
17	5.799	79.37	246.1	644.7	1.558	47	1.26	9.327	6.768
18	6.022	77.06	236.5	615.1	1.565	45.6	1.184	8.397	6.211
19	6.131	82.8	252.6	652.1	1.569	44.4	1.27	9.809	6.778
20	7.143	77.84	228.2	607.5	1.597	50.2	1.35	9.668	6.259
21	6.638	72.87	233.9	591.8	1.535	40.9	1.134	7.614	5.508
22	6.854	79.42	245.7	647.3	1.559	48.1	1.262	9.074	6.545
23	7.131	77.11	236.2	617.7	1.566	46.7	1.186	8.144	5.987
24	7.284	82.99	253.5	659.6	1568	45.8	1.27	9.556	6.554
25	7.409	78.02	229	615	1.596	51.9	1.34	9.414	6.036
26	5.419	77.46	250.8	629	1.529	39.5	1.113	8.443	5.686
27	5.529	84.01	262.6	684.5	1.552	46.1	1.234	9.903	6.723
28	6.013	81.7	253	654	1.558	44.8	1.163	8.973	6.165
29	6.745	87.41	268.7	690.7	1.564	43.6	1.25	10.385	6.732
30	7.387	82.45	244.2	646.1	1.59	48.9	1.31	10.243	6.214
31	6.308	75.39	229.1	599.9	1.571	46.9	1.21	8.504	6.44
32	6.48	81.94	241	655.4	1.595	54.6	1.337	9.965	7.477
33	6.74	79.63	231.4	625.7	1.604	53.4	1.262	9.034	6.92
34	7.032	85.23	241.9	651	1.622	52.4	1.38	10.446	7.487
35	7.252	80.27	217.5	606.4	1.659	60.4	1.47	10.305	6.968
36	5.57	130	429.5	1086	1.517	40.8	1.046	12.21	8.758
37	5.727	136.54	441.3	1141.5	1.53	44.7	1.12	13.67	9.795
38	5.995	137.33	450.7	1134.9	1.521	40.1	1.026	12.74	9.237
39	6.059	139.33	444.2	1136.1	2	42.7	1.14	14.152	9.804
40	6.159	134.37	419.8	1091.5	1.553	45.6	1.17	14.011	9.286

In the present study, we have performed the quantitative structure activity relationship analysis by various regression methods. Regression methods are used to build a QSAR model in the form of a mathematical equation. This equation explains variation of one or more dependent variables (usually activity) in terms of independent variables. The QSAR model can be used to predict activities for new molecules, for screening a large set of molecules whose activities are not known. Multiple regression is the standard method for multivariate data analysis. Stepwise multiple regression (SMR) is an approach to select a subset of variables, when the numbers of independent variables (descriptors) are much more than the number of data points (molecules). SMR is a way of computing OLS regression in stages [11-16]. It is also a procedure to examine the impact of each variable to the model step by step.

2 MATERIAL AND METHODS

Data set

A data set of 40 molecules have been taken from the reputed published results [17]. Anticancer activity was expressed as pIC_{50} values. [Table-1] It is essential to assess the predictive power of the models by using a test set of compounds. This was achieved by arbitrarily setting aside some compounds as a test set. The structures and anticancer activity data of compounds are listed in Table 1.

Molecular structure generation

All the molecular modeling and statistical analysis were performed using Vlife MDS software [18-20]. The structures of the compounds were built using molecular sketching facilities provided in the modeling environment of Virus life. Energy minimization and batch optimization was carried out using Merck Molecular force field. All the molecules were initially optimized and then used for the calculation of descriptors and further QSAR study.

Table 3: Correlation Matrix

	pIC50	MR	MV	PC	IR	ST	Density	X2	X3
pIC50	1.0000								
MR	-0.1643	1.0000							
MV	-0.2748	0.9860	1.0000						
Parachor	-0.2223	0.9965	0.9946	1.0000					
IR	0.2125	-0.0115	-0.0089	-0.0140	1.0000				
ST	0.5497	-0.2243	-0.3696	-0.2733	-0.0562	1.0000			
DEN	0.6046	-0.3706	-0.5084	-0.4241	0.0222	0.9174	1.0000		
X2	0.0535	0.9227	0.8571	0.8970	0.0011	0.0634	-0.0194	1.0000	
X3	-0.0857	0.9469	0.9019	0.9347	-0.0118	-0.0152	-0.1611	0.9497	1.0000

QSAR study

All the 2D descriptors (thermodynamic, spatial, electronic and topological parameters) were calculated for QSAR analysis using Vlife MDS software. Thermodynamic parameters describe free energy change during drug receptor complex formation. Spatial parameters are the quantified steric features of drug molecules required for its complimentary fit with receptor. Electronic parameters describe weak non-covalent bonding between drug molecules and receptor [21-22]. Random Selection method and Sphere Exclusion Method were used for the selection of the training and test set. For variable selection Stepwise forward-backward method was used. A suitable statistical method coupled with a variable selection method allows analyses of this data in order to establish a QSAR

model with the subset of descriptors that are most statistically significant in determining the biological activity [23-25]. The QSAR models were generated by Multiple Regression Analysis method.

3 RESULTS AND DISCUSSION

For QSAR analysis regression was performed using IC50 values as dependent variables and calculated parameters as independent variables. In any thorough investigation of the effects of molecular properties, it is essential to prove that the results are both statistically valid and make chemical sense. It would be appropriate to obtain insight into the physical meaning of the correlation obtained as an output of the regression analysis. The magnitude of a descriptor could be used as a guideline to improve the anticancer activity of molecules.

Table 4: Results of regression analysis

Eq. No.	QSAR/QSPR Models	N	R^2	R^2adj	MSE	PRESS	R^2cv		Ratio
1	pIC50=2.0860+ 3.4277DEN	40	0.3656	0.3489	0.3081	12.84265	0.3043	0.870	21.898
2	pIC50= 5.19097+0 .2805MR -3.3334E-02*PC	40	0.5218	0.4959	0.2386	10.0181	0.4573	0.0765	20.186
3	pIC50= 5.1771+ 5.2541E-04*IR +0.2782MR-3.3057E-02*PC	40	0.5575	0.5207	0.2269	493768.8	0.0000	0.0746	15.121
4	pIC50= 6.1369+ 6.1647E-04*IR + 0.1932MR -4.9529E-02*MV -0.5040X3	40	0.5853	0.5379	0.2187	124978.1	0.0000	0.0733	12.349
5	pIC50= 7.5527+ 5.664E-04*IR +0.2477MR -0.0676MV -4.0291E-02*ST -0.4135X3	40	0.6063	0.5484	0.2138	463744.1	0.0000	0.0724	10.471
After deletion of compound no. 21,22 and 23 as outlier									
6	pIC_{50} = 7.5539+0.2784MR -7.6851E-02*MV + 6.0620E-04*IR -5.0431E-02*ST -0.3874X3	37	**0.7258**	0.6815	0.1552	608814.8	0.0000	0.621	16.407

With reference to table 3, the selected descriptors are used for mono paramatric QSAR model no.1 development which show the importance of physicochemical descriptor density which is directly proportional with the antiviral activity as Neuraminidase Inhibitors positively correlated. The monoparametric QSAR model No.1 is given below;

pIC50=2.0860+ 3.4277DEN   **Eq…1**

low statistical results indicates needs for the development of multiparametic and more QSAR models follow rule of thumb. The QSAR model no.2 has significant importance in which MR has positive contribution with the antiviral activity while the Pc show inverse contribution with antiviral activity. The statistical descriptors are given in Table no.4 (Model No.2).

pIC50= 5.1909+0.2805MR-3.3334E-02*PC

                                        Eq…2

The QSAR model no. 3 show their significant statistical importance with tri parametric model in which IR and MR are directly proportional with the antiviral activity while Pc are inversely proportional with the antiviral activity (QSAR model No.3).

pIC50= 5.1771+ 5.254E-04*IR+ 0.2782MR-3.3057E-02*PC   **Eq…3**

The above described all models are not statistically excellent indicates the deletion of outliers compound whose activity are not uniform and After deleting Comp No.,21,22 and 23 resulting the development of high statistically significant qsar model no.5 indicates that the MR play a major role in the antiviral activity.

pIC50= 7.5527+ 5.6647E-04*IR+ 0.2477MR-0.0676MV-4.0291E-02*ST-0.4135X3   **Eq….4**

Among the generated QSAR models; three models were selected on the basis of various statistical parameters such as squared correlation co-efficient (r2) which is relative measure of quality of fit. Fischer's value (F test) which represents F-ratio between the variance of calculated and observed activity.

These models were generated in stepwise manner by forward-backward selection method starting with best single variable and adding further significant variable according to their contribution to the model. Various models of the data set were obtained which showed individual correlation of all calculated parameters with IC_{50} of anticancer activity.

Table 5: Actual and predicted antiviral activity with residuals

Row	pIC$_{50}$ actual	Predicted pIC$_{50}$
1	6.658	0.893
2	5.939	-0.011
3	6.509	-0.143
4	7.444	0.01
5	7.495	-0.13
6	5.274	-0.596
7	6.244	0.252
8	6.638	0.26
9	6.921	0.281
10	7.309	0.301
11	5.033	-0.687
12	5.447	-0.392
13	5.899	-0.304
14	5.996	-0.338
15	6.143	-0.54
16	5.527	-0.065
17	5.799	0.047
18	6.022	-0.11
19	6.131	-0.203
20	7.143	0.406
21	7.284	0
22	7.409	0.683
23	5.419	-0.237
24	5.529	-0.309
25	6.013	-0.202
26	6.745	0.306
27	7.387	0.513
28	6.308	0.226
29	6.48	0.279
30	6.74	0.168
31	7.032	-0.123
32	7.252	-0.195
33	5.57	0.272
34	5.727	0.113
35	5.995	0.436
36	6.059	-0.206
37	6.159	-0.654

4 CONCLUSIONS

In summary, from the derived QSAR model it can be concluded that antiviral activity by the Neuraminidase inhibitors is strongly influenced by the physicochemical descriptors interactions and electro-topological nature of substituents. Pattern of substitution can be extracted from the developed model. The descriptors showed by QSAR study can be used further for study and designing of new compounds. Consequently this study may prove to be helpful in development and optimization of existing antiviral activity of this class of compounds.
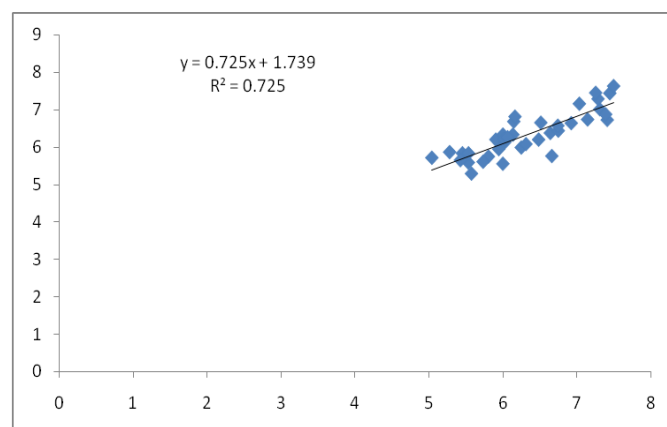

$$y = 0.725x + 1.739$$
$$R^2 = 0.725$$

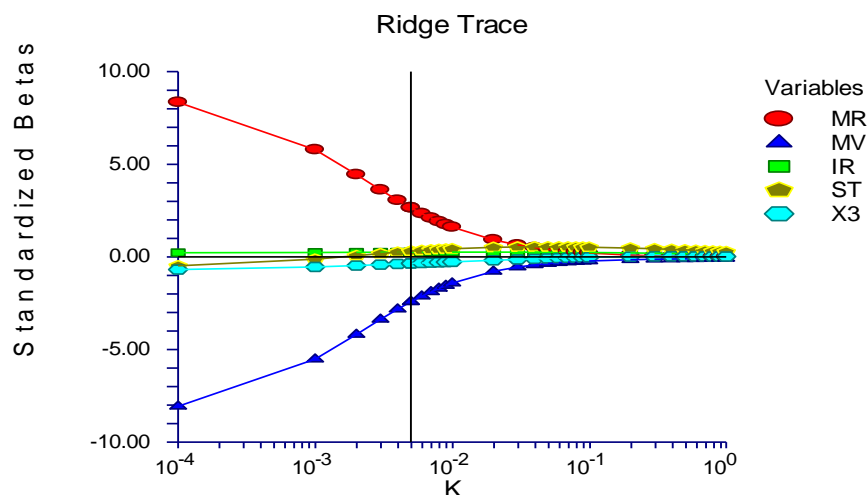Figure 2: Graph plotted between predicted IC$_{50}$ and actual IC$_{50}$

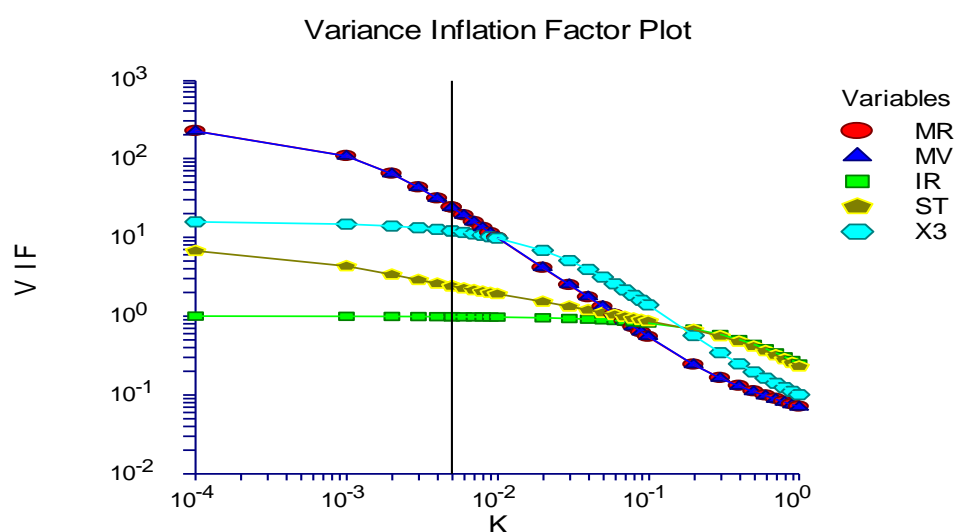Figure 3: Graph plotted between k and Standardized beats of used descriptors in QSAR modeling

Figure 4: Graph plotted between VIF and K of used descriptors in QSAR/QSPR modeling

REFERENCES

1. Pennisi E. (1995) Science; 270:1916–1917.
2. Liu C.G., Eichelberger M.C., Compans R.W. (1995) J Virol; 69:1099–1106.
3. Ghate A.A., Air G.M. (1998). Eur J Biochem; 258:320–331.
4. Magano J. (2009) Chem Rev; 109:4398–4438.
5. Vincent S., Kent D., Stewart C.J., Maring S.M., Vincent G., Yugui G.,Gary W. et al. (2003). Biochemistry; 42:718–727.
6. Bossart W.P., Carson M., Babu Y.S., Smith C.D., Laver W.G., Air G.M. (1993) J Mol Biol;232:1069–1083.
7. Jedrzejas M.J., Singh S., Brouillette W.J., Laver W.G., Air G.M., Luo M. (1995) Biochemistry;34:3144–3151.
8. Zhang J., Wang Q., Fang H., Xu W., Liu A., Du G. (2007) Bioorg Med Chem;15:2749–2758.
9. Zhang J., Wang Q., Fang H., Xu W., Liu A., Du G. (2008) Bioorg Med Chem;16:3839–3847.
10. Sun C., Zhang X., Huang H., Zhou P. (2006) Bioorg Med Chem;14:8574–8581.
11. Sun J., Cai S., Yan N., Mei H. (2010) Eur J Med Chem;45:1008–1014.
12. Tintori C., Magnani M., Schenone S., Botta M. (2009) Eur J Med Chem;44:990–1000.
13. Pan J., Liu G.Y., Cheng J., Chen X.J., Ju X.L. (2010) Eur J Med Chem;45:967–972.
14. Singh S., Jedrzejas M.J., Air G.M., Luo M., Laver W.G., Brouillette W.J. (1995) J Med Chem;38:3217–3225.
15. da Cunha E.F., Sippl W., de Castro R.T., Ceva-Antunes O.A., de Alencastro R.B., Albuquerque M.G. (2009) Eur J Med Chem;44:4344–4352.
16. Puntambekar D.S., Giridhar R., Yadav M.R. (2006) Eur J Med Chem;41:1279–1292.
17. Xiao J., Guo Z., Guo Y., Chu F., Sun P. (2006) Protein Eng Des Sel;19:47–54.
18. Crivori P., Poggesi I. (2006) Eur J Med Chem;41:795–808.
19. Falc_ J.L., Lloveras M., Buira I., Teixid_ J., Borrell J.I., M_ndez E., Terencio J., Palomer A., Guglietta A. (2005) Eur J Med Chem;40:1179–1187.
20. Hu X. (2006) Bioorg Med Chem Lett;16:6321–6327.
21. Gueto C., Ruiz J.L., Torres J.E., M_ndez J., Vivas-Reyes R. (2007) Bioorg Med Chem;16:2439–2447.
22. Sivan S.K., Manga V. (2010) J Mol Model;16:1169–1178.
23. Chen Y., Li H., Tang W., Zhu C., Jiang Y., Zou J., Yu Q., You Q. (2008) Eur J Med Chem;44:2868–2876.
24. Lu P., Wei X., Zhang R. (2010) Eur J Med Chem;45:1792–1798.
25. Gupta P., Roy N., Garg P. (2009) Eur J Med Chem;44:4276–4287.