

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/41427495>

3D-QSAR studies on quinazoline antifolate thymidylate synthase inhibitors by CoMFA and CoMSIA models

ARTICLE *in* EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY · APRIL 2010

Impact Factor: 3.45 · DOI: 10.1016/j.ejmech.2009.12.065 · Source: PubMed

CITATIONS

18

READS

23

4 AUTHORS, INCLUDING:



Satya Prakash Gupta

National Institute of Technical Teachers' Tr...

164 PUBLICATIONS 1,272 CITATIONS

SEE PROFILE



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: <http://www.elsevier.com/locate/ejmech>

Original article

3D-QSAR studies on quinazoline antifolate thymidylate synthase inhibitors by CoMFA and CoMSIA models

Vivek Srivastava^{a,d}, S.P. Gupta^{b,*}, M.I. Siddiqi^c, B.N. Mishra^{d,**}^a Department of Biotechnology, Meerut Institute of Engineering and Technology, Meerut 250005, India^b Department of Pharmaceutical Technology, Meerut Institute of Engineering and Technology, Meerut 250005, India^c Computational Biology and Bioinformatics Lab., MSB Division, CDRI, Lucknow 226001, India^d Department of Biotechnology, Institute of Engineering and Technology, UP Technical University Campus, Sitapur Road, Lucknow 226021, India

ARTICLE INFO

Article history:

Received 13 June 2009

Received in revised form

12 December 2009

Accepted 23 December 2009

Available online 13 January 2010

Keywords:

CoMFA

CoMSIA

QSAR

Quinazolines

Thymidylate synthase inhibitors

ABSTRACT

Thymidylate synthase (TS) is a crucial enzyme for DNA biosynthesis and many nonclassical lipophilic antifolates targeting this enzyme are quite efficient and encouraging as antitumor drugs. Herein, we report some 3D-QSAR analyses using CoMFA and CoMSIA on quinazoline antifolates in order to have a better understanding of the mechanism of action and structure-activity relationship of these compounds. By applying leave-one-out (LOO) cross-validation study, we obtained cross-validated q^2 value of 0.573 for CoMFA and 0.445 for CoMSIA, while the non-cross-validated r^2 values for them were found to be 0.935 and 0.893, respectively. The models were graphically interpreted using CoMFA and CoMSIA contour plots. The results obtained from this study could be used for rational design of potent inhibitors against thymidylate synthase.

© 2010 Elsevier Masson SAS. All rights reserved.

1. Introduction

Folate metabolism has long been recognized as an attractive target for cancer chemotherapy because of its indispensable role in the biosynthesis of nucleic acid precursors [1,2]. The crucial enzyme involved in folate metabolism is thymidylate synthase (TS). The folate-binding site in TS is believed to offer better opportunities for the design of highly specific inhibitors than the pyrimidine (dUMP) binding site. This belief led to the design of drugs such as CB3717, ZD1694 (Thomudex, Raltitrexed), LY231514, AG337 (Thymitaq), and GW1843U89, a quinazoline-based drug which had antitumor activity in clinical studies [3]. The antifolate compounds evaluated in this investigation are derivatives of quinazoline antifolate, having structures similar to the Tomudex/ZD1694 class of antifolates, where quinazoline is a compound made up of two fused six-membered simple aromatic rings, a benzene ring and a pyrimidine ring. Its chemical formula is $C_8H_6N_2$. Due to the interest in design of new anticancer drugs, several quinazoline antifolate inhibitors were chosen from the ICI Pharmaceutical and Institute of

Cancer Research, England, for screening against human thymidylate synthase [4–7]. In our 3D-QSAR study, we have taken one hundred six molecules; whose structures of and the biological activity (IC_{50}) are given in Table 1.

The three-dimensional quantitative structure-activity relationship (QSAR) study based on only the 3D-structures of the ligands involves two methods: comparative molecular field analysis (CoMFA) [8] and comparative molecular similarity indices analysis (CoMSIA) [9]. QSAR studies incorporate 3D information for the ligands and provide a more detailed analysis of ligand–receptor interactions [10]. The present 3D CoMFA and CoMSIA studies of one hundred six quinazoline antifolate thymidylate synthase inhibitors will not only illustrate the conformation or spatial orientation of anti cancer antifolates but also provide useful rationale for further design of new drug candidates for cancer.

2. Materials and method

2.1. Molecular structures and optimization

One hundred six molecules selected for the present study were taken from the work of Hughes et al. [4–7]. The structures of the compounds and their biological data are given in Table 1. The IC_{50} values were converted into the corresponding pIC_{50} ($-\log IC_{50}$) and

* Corresponding author. Fax: +91 121 2439058; +91 9368222297 (mobile).

** Corresponding author. Fax: +91 522 2365730; +91 9415012030 (mobile).

E-mail addresses: spgbits@gmail.com (S.P. Gupta), profbnmishra@gmail.com (B.N. Mishra).

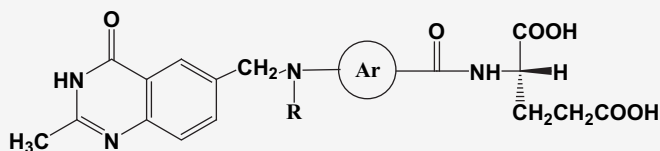
Table 1
Quinazoline antifolate derivatives and their observed and predicted TS inhibition activities.

Compound	Quinazoline antifolate derivatives with modifications		Observed Activity (pIC ₅₀)	Predicted Activity (CoMFA)	Residual	Predicted Activity (CoMSIA)	Residual
	R ¹	R ²					
1	CH ₃	CH ₂ C≡CH	7.4	7.05	0.35	7.01	0.39
2 ^a	CH₃	H	5.35	5.27	0.08	5.12	0.23
3 ^a	CH₃	CH₃	6.52	6.33	0.19	6.44	0.08
4	CH ₃	CH ₂ CH ₃	6.77	6.91	−0.14	6.89	−0.12
5	CH ₃	CH ₂ CH=CH ₂	6.32	6.39	−0.07	6.37	−0.05
6 ^a	CH₃	(CH₂)₂F	6.62	6.60	0.02	6.58	0.04
7	CH ₃	(CH ₂) ₂ Br	5.89	5.90	−0.01	6.01	−0.12
8	CH ₃	(CH ₂) ₂ SH	5.75	5.77	−0.02	5.77	−0.02
9	CH ₃	(CH ₂) ₂ OH	6.3	6.38	−0.08	6.19	0.11
10	CH ₃	(CH ₂) ₃ OH	6.27	6.09	0.18	6.02	0.25
11	CH ₃	(CH ₂) ₂ OCH ₃	4.79	4.90	−0.11	5.02	−0.23
12 ^a	CH₃	(CH₂)₃OCH₃	4.89	4.66	0.23	5.08	−0.19
13	CH ₃	CH ₂ COCH ₃	4.6	4.83	−0.23	5.41	−0.81
14	CH ₂ CH ₃	CH ₂ C≡CH	6.85	6.75	0.10	6.79	0.06
15	CH ₂ CH ₃	H	4.7	4.83	−0.13	4.90	−0.20
16	CH(CH ₃) ₂	CH ₂ C≡CH	6.21	6.63	−0.42	6.76	−0.55
17	CH ₂ F	CH ₂ C≡CH	7	6.94	0.06	6.70	0.30
18	CH ₂ F	CH ₃	5.54	5.08	0.46	5.34	0.20
19	CH ₂ F	CH ₂ CH ₃	6.43	6.64	−0.21	6.49	−0.06
20	CH ₂ F	(CH ₂) ₂ F	6.47	6.40	0.07	6.29	0.18
21	CH ₂ F	CH ₂ C≡CH	6.24	6.82	−0.58	6.65	−0.41
22 ^a	CF₃	CH₂C≡CH	5.24	5.02	0.22	4.80	0.44
23 ^a	CH₂OH	CH₃	6.19	6.60	−0.41	5.69	0.50
24	CH ₂ OH	CH ₂ CH ₃	6.58	6.75	−0.17	6.75	−0.17
25	CH ₂ OH	CH ₂ CH=CH ₂	6.04	6.06	−0.02	6.06	−0.01
26	CH ₂ OH	(CH ₂) ₂ F	6.66	6.51	0.15	6.53	0.13
27	CH ₂ OH	(CH ₂) ₂ OH	6.11	6.25	−0.14	6.25	−0.14
28	CH ₂ NHCOCH ₃	CH ₂ C≡CH	6.32	6.40	−0.08	6.24	0.08
29	CH ₂ S-2-pyrimidine	CH ₂ C≡CH	6.62	6.34	0.28	6.39	0.23
30	Phenyl	CH ₂ C≡CH	6.66	6.46	0.20	6.70	−0.04
	X	Y					
31	NH ₂	H	7.7	7.57	0.14	7.90	−0.20
32	NH ₂	F	8.05	7.88	0.17	8.22	−0.17
33	H	H	6.77	7.06	−0.29	7.09	−0.32
34	H	F	7.24	7.49	−0.25	7.34	−0.10
35	CH ₃	F	7.7	7.49	0.21	7.27	0.43
36	CH ₂ OH	H	7.01	6.92	0.09	7.06	−0.05
37	CH ₂ OH	F	7.24	7.35	−0.11	7.33	−0.09
38	OCH ₃	H	7.17	7.22	−0.05	7.29	−0.12
39	OCH ₃	F	7.44	7.23	0.21	7.15	0.29
	X	R					
40	H	H	5.35	5.39	−0.04	5.10	0.25
41	H	CH ₃	6.52	6.59	−0.07	6.47	0.05

(continued on next page)

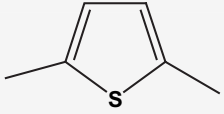
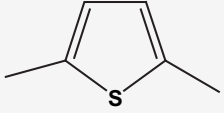
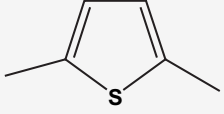
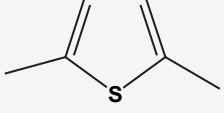
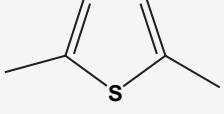
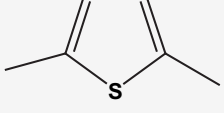
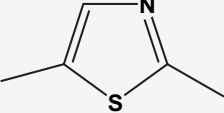
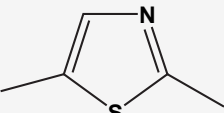
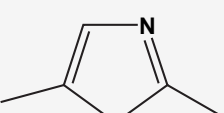
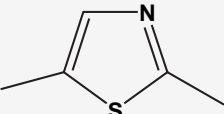
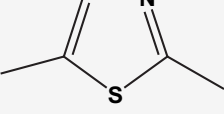
Table 1 (continued)

42	H	CH ₂ CH ₃	6.77	6.89	−0.12	6.70	0.07
43	H	CH ₂ CH=CH ₂	6.32	6.20	0.13	6.02	0.30
44	H	(CH ₂) ₂ F	6.62	6.57	0.05	6.45	0.17
45	H	(CH ₂) ₂ OH	6.3	6.38	−0.08	6.19	0.11
46	H	(CH ₂) ₃ OH	6.27	6.09	0.18	6.02	0.25
47 ^a	2'-F	CH ₂ C≡CH	7.7	7.49	0.21	7.24	0.46
48	2'-F	H	5.42	5.26	0.16	5.18	0.24
49 ^a	2'-F	CH ₃	6.92	6.71	0.21	6.75	0.17
50	2'-F	CH ₂ CH ₃	7.35	7.00	0.35	6.95	0.40
51 ^a	2'-F	CH ₂ CH=CH ₂	7.12	7.50	−0.38	6.25	0.87
52 ^a	2'-F	(CH ₂) ₂ F	7.37	7.26	0.11	7.44	−0.06
53	2'-F	(CH ₂) ₂ OH	7.11	6.78	0.33	6.32	0.80
54	2'-F	(CH ₂) ₃ OH	6.23	6.34	−0.11	6.22	0.01
55 ^a	2'-F	(CH ₂) ₂ NH ₂	4.93	5.03	−0.10	5.24	−0.31
56 ^a	2'-F	CH ₂ CN	6.77	7.00	−0.23	7.43	−0.66
57	2'-F	CH ₂ CONH ₂	5.87	5.77	0.10	5.68	0.19
58 ^a	3'-F	CH ₂ C≡CH	5.84	6.06	−0.22	6.06	−0.22
59	3'-F	H	5.03	5.40	−0.37	5.28	−0.25
60	3'-F	CH ₃	5.94	5.65	0.29	5.85	0.09
61	3'-F	CH ₂ CH ₃	6.17	6.10	0.07	6.52	−0.35
62	3'-F	CH ₂ CH=CH ₂	5.29	5.32	−0.03	5.58	−0.29
63	3'-F	(CH ₂) ₂ F	5.76	5.79	−0.03	5.52	0.24
64	3'-F	(CH ₂) ₂ OH	5.45	5.43	0.02	5.75	−0.30
65	3'-F	(CH ₂) ₃ OH	6.01	6.09	−0.08	6.11	−0.10
66	2'-Cl	CH ₂ C≡CH	7.13	7.17	−0.04	7.24	−0.11
67	2'-Cl	CH ₂ CH ₃	6.92	6.95	−0.03	6.87	0.05
68	2'-CF ₃	CH ₂ C≡CH	6.32	6.08	0.25	6.25	0.08
69 ^a	2'-CH ₃	CH ₂ C≡CH	6.3	6.85	−0.54	6.76	−0.46
70	2'-CH ₃	CH ₃	6.03	6.85	−0.06	6.08	−0.05
71 ^a	2'-CH ₃	CH ₂ CH ₃	6.5	6.85	−0.01	6.24	0.26
72	2'-NH ₂	CH ₂ C≡CH	6.54	6.85	−0.13	6.55	−0.01
73 ^a	2'-NH ₂	CH ₃	6.07	6.85	−0.11	6.09	−0.02
74	2'-NH ₂	CH ₂ CH ₃	6.51	6.85	−0.01	6.47	0.04
75	2'-OH	CH ₂ C≡CH	7.05	6.85	−0.19	7.14	−0.09
76	2'-OCH ₃	CH ₂ C≡CH	7.17	6.85	−0.04	7.25	−0.08
77	2'-OCH ₃	CH ₂ CH ₃	6.72	6.85	−0.10	6.88	−0.16
78	2'-NO ₂	CH ₂ C≡CH	6.89	6.85	−0.11	6.81	0.08
79	2'-NO ₂	CH ₂ CH ₃	6.84	6.85	−0.01	6.84	0.00



R		Ar					
80 ^a	CH ₂ C≡CH		6.36	6.58	−0.22	6.77	−0.41
81 ^a	H		4.61	4.43	0.18	4.72	−0.11
82 ^a	CH ₃		6.17	6.03	0.14	6.20	−0.03
83	CH ₂ CH ₃		6.24	6.21	0.03	6.27	−0.03

Table 1 (continued)

84 ^a	(CH ₂) ₂ CH ₃		5.73	6.07	−0.34	6.14	−0.41
85	CH ₂ CH=CH		5.78	5.85	−0.07	5.84	−0.06
86 ^a	(CH ₂) ₂ F		6.26	6.27	−0.01	5.94	0.32
87	(CH ₂) ₂ OH		5.92	5.91	0.01	5.91	0.01
88 ^a	(CH ₂) ₃ OH		5.47	6.05	−0.58	6.06	−0.59
89	CH ₂ CN		5.51	5.54	−0.03	5.61	−0.10
90	CH ₂ C≡CH		6.64	6.73	−0.09	6.79	−0.15
91 ^a	H		5.15	5.49	−0.34	5.46	−0.31
92	CH ₃		6.38	6.54	−0.16	6.50	−0.12
93	CH ₂ CH ₃		6.64	6.40	0.24	6.73	−0.09
94 ^a	CH ₂ CH=CH ₂		5.97	6.67	−0.70	5.68	0.29

(continued on next page)

Table 1 (continued)

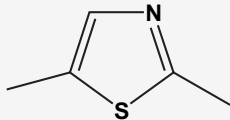
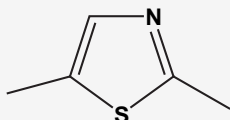
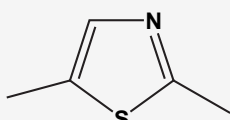
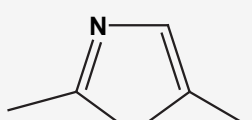
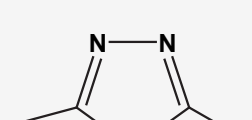
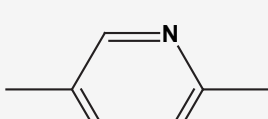
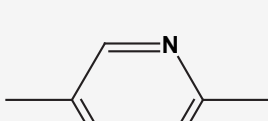
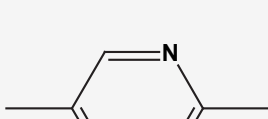
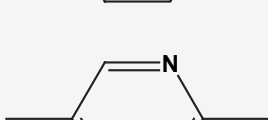
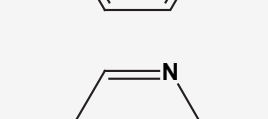
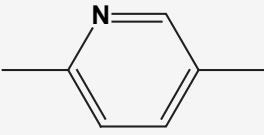
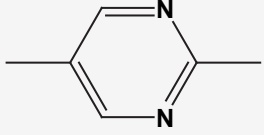
95	(CH ₂) ₂ F		6.72	6.73	−0.01	6.72	0.00
96	(CH ₂) ₂ OH		6.38	6.48	−0.10	6.57	−0.19
97	(CH ₂) ₃ OH		5.76	5.78	−0.02	5.67	0.09
98	CH ₂ CH ₃		6.12	6.29	−0.17	6.14	−0.02
99	CH ₂ CH ₃		6.03	5.99	0.04	6.00	0.03
100	CH ₂ C≡CH		7.4	7.14	0.26	7.18	0.22
101^a	CH₃		6.46	6.33	0.13	5.61	0.85
102^a	CH₂CH₃		6.82	7.22	−0.40	6.65	0.17
103	CH ₂ CH=CH ₂		6.12	6.04	0.08	5.93	0.19
104	(CH ₂) ₂ F		6.8	6.67	0.13	6.72	0.08

Table 1 (continued)

105	CH ₃		5.37	5.32	0.05	5.45	−0.08
106	CH ₃		5.63	5.57	0.06	5.63	0.00

^a Compounds taken for the test set.

used as dependent variables in CoMFA and CoMSIA analysis. The pIC₅₀ values have a span of 3 log units providing a broad and homogenous data set for 3D-QSAR study. The 3D-QSAR models were generated using a training set of eighty molecules (Table 1). Predictive power of the resulting models was evaluated using a test set of twenty six molecules (compounds marked with ^a in Table 1). The test compounds were selected randomly such that the structural diversity and wide range of activity in the data set were included.

2.2. Molecular alignment

CoMFA and CoMSIA results may be extremely sensitive to a number of factors such as alignment rules, over all orientation of the aligned compounds, lattice shifting step size and probe atom type [11]. The accuracy of the prediction of CoMFA and CoMSIA models and the reliability of the contour models depend strongly on the structural alignment of the molecules [11] and thus we applied molecular alignment to align all the molecules used in present study in space. The molecular alignment was achieved by the SYBYL routine database align. The most active compound (compound **32**) was used as an alignment template and the rest of the molecules were aligned to it by using the common substructure as shown in Fig. 1.

2.3. 3D QSAR studies

In order to have better understanding and explore the contributions of electrostatic, steric, and hydrophobic fields of the data set, and to build predictive 3D-QSAR models, CoMFA and CoMSIA studies were performed based on the molecular alignment as

described. CoMFA calculates steric and electrostatic properties according to Lennard-Jones and coulomb potentials, respectively, whereas CoMSIA calculates similarity indices in the space surrounding each of the molecules in the dataset.

2.3.1. CoMFA studies

Steric and electrostatic interactions were calculated using the Tripos force field [12] with a distance-dependent dielectric constant at all interactions in a regularly spaced (2Å) grid taking a sp³ carbon atom as steric probe and a +1 charge as electrostatic probe. The cutoff was set to 30 kcal/mol. With standard options for scaling of variables, the regression analysis was carried out using the full cross-validated partial least squares (PLS) method (leave-one-out) [13]. The minimum sigma (column filtering) was set to 2.0 kcal/mol to improve the signal to noise ratio by omitting those lattice points whose energy variation was below this threshold. The final model, non cross-validated conventional analysis, was developed with the optimum number of components to yield a non cross-validated *r*² value.

2.3.2. CoMSIA studies

In CoMSIA, a distance-dependent Gaussian-type physicochemical property has been adopted to avoid singularities at the atomic positions and dramatic changes of potential energy for grids being in the proximity of the surface. With the standard parameters and no arbitrary cutoff limits, three fields associated to three physicochemical properties, namely, steric, electrostatic, and hydrophobic, were calculated. The steric contribution was reflected by the third power of the atomic radii of the atoms. Electrostatic properties were introduced as atomic charges resulting from molecular docking. An atom-based hydrophobicity was assigned according to

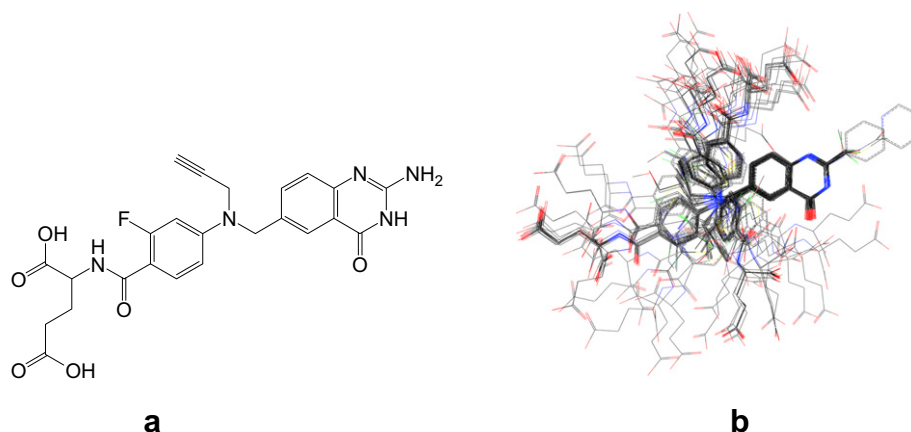


Fig. 1. (a) Template used for molecular alignment of quinazoline antifolate derivatives (compound **32**). (b) Molecular alignment.

the parameterization developed by Ghose et al. [14] The lattice dimensions were selected with a sufficiently large margin ($>4\text{\AA}$) to enclose all the binding conformations of the inhibitors. In general, similarity indices, $A_{F,K}$, between the compounds of interest were computed by placing a probe atom at the intersections of the lattice points and using Equation (1),

$$A_{F,K}^q(j) = - \sum_{i=1}^n W_{\text{probe},k} W_{ik} e^{-a_{iq} r_{iq}^2} \quad (1)$$

where q represents a grid point, i is the summation index over all atoms of the molecule j under computation, W_{ik} is the actual value of the physicochemical property k of atom i , and $W_{\text{probe},k}$ is the value of the probe atom. In the present study, we used a probe atom ($W_{\text{probe},k}$) with charge +1, radius 1Å, hydrophobicity +1, and attenuation factor of 0.3 for the Gaussian type distance. The statistical evaluation for the CoMSIA analysis was performed in the same way as described for CoMFA.

2.4. Partial least square (PLS) analysis

PLS method [15] was used to linearly correlate the CoMFA and CoMSIA fields to biological activity values. The cross-validation was performed using leave-one-out (LOO) method in which one compound is removed from the dataset and its activity is predicted using the model derived from the rest of the molecules in the dataset. Equal weights for CoMFA were assigned to steric and electrostatic fields using CoMFA STD scaling option. To speed up the analysis and reduce noise, a minimum column filter value of 2.0 kcal/mol was used for the cross-validation. Non-cross-validation was performed to calculate conventional r^2 using the same number of components. To further assess the robustness and statistical confidence of the derived models, bootstrapping analysis for 100 runs was performed [15,16]. Bootstrapping involves the generation of many new data sets from original data set and is obtained by randomly choosing samples from the original data set. The statistical calculation is performed on each of these bootstrapping samplings. The difference between the parameters calculated from the original data set and the average of the parameters calculated from the many bootstrapping samplings is a measure of the bias of the original calculations. The entire cross-validated results were analyzed considering the fact that a value of q^2 above 0.3 indicates that probability of chance correlation is less than 5% [15].

2.5. Hardware and software

InsightII 2000.1 [17] and Sybyl 7.1 [18] were used for molecular modeling on a SGI Origin 300 workstation equipped with 4 * 600 Mhz R12000 processors.

3. Results and discussion

3.1. 3D-QSAR analysis

CoMFA and CoMSIA 3D-QSAR models were derived using previously reported human thymidylate synthase (TS) inhibitors. The chemical structures of molecules and their experimental biological activity (IC_{50}) values are shown in Table 1.

3.1.1. CoMFA analysis

Eighty compounds out of the total one hundred six TS inhibitors were used as training set and rest twenty six compounds were used as test set. The test set compounds were selected randomly in such a way that the structural diversity and wide range of activity in the data set were included. PLS analysis was carried out for the training set and the result is listed in Table 2, which shows a cross-validated

Table 2

Statistical results of CoMFA and CoMSIA models.

	CoMFA	CoMSIA
<i>PLS Statistics</i>		
q^2	0.573	0.445
r^2	0.935	0.893
S	0.187	0.239
F	147.631	86.231
$r^2_{\text{bootstrap}}$	0.955	0.929
$S_{\text{bootstrap}}$	0.152	0.192
Optimal Components	7	7
<i>Field Distribution %</i>		
Steric	37.4	13.6
Electrostatic	62.6	35.8
Hydrophobic	–	–
Hydrogen Bond Donor	–	26.6
Hydrogen Bond Acceptor	–	23.9
r^2_{pred}	0.781	0.583

q^2 value of 0.573 for seven components used. The non cross-validated PLS analysis with the optimum components revealed a conventional r^2 value of 0.935, with an F -value (Fisher ratio) of 147.631 and an estimated standard error of prediction (SEP) equal to 0.187. As obvious from Table 2, the steric field descriptors explain 37.4% of the variance, while the electrostatic descriptors explain 62.6% of the variance. Bootstrap analysis for 100 runs was then carried out for further validation of the model by statistical sampling of the original data set to create new data sets. Thus, the difference in the parameters calculated from the original data and the average of the parameters calculated from the $N (=100)$ runs of bootstrapping sampling is a measure of the bias of the original calculation. This yielded higher $r^2_{\text{bootstrap}}$ value 0.955 for CoMFA with standard error value of 0.152, supporting the statistical validity of the developed models. The predicted activities for the inhibitors versus their experimental activities are listed in Table 1 and the correlation between the predicted activities and the experimental activities is depicted in Fig. 2. Table 1 and Fig. 2 demonstrate that the predicted activities by the constructed CoMFA model are in good agreement with the experimental data, suggesting that the CoMFA model should have a satisfactory predictive ability.

3.1.2. CoMSIA analysis

The CoMSIA analysis was performed using four descriptor fields: steric, electrostatic, hydrogen bond donor, and hydrogen bond acceptor. The CoMSIA study revealed $q^2 = 0.445$ (optimum number

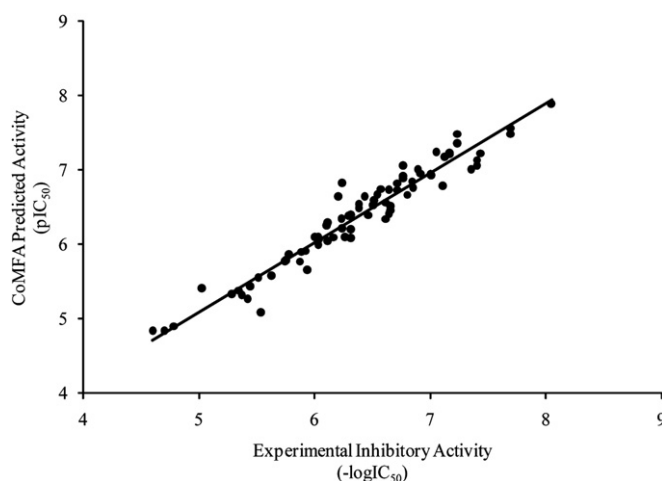


Fig. 2. Correlation between experimental and CoMFA predicted inhibitory activities of the training set compounds.

of components 7), $r^2 = 0.893 < \text{SEP } 0.239$, and $F = 86.231$ (Table 2) for training set. As obvious from the Table 2, the percentage of the variance explained by steric, electrostatic, hydrogen bond donor, and hydrogen bond acceptor field descriptors are 13.6, 35.8, 26.6, and 23.9, respectively. Bootstrap analysis for 100 runs was then carried out for further validation of the model by statistical sampling of the original data set to create new data set. Thus, the difference in the parameters calculated from the original data and the average of the parameters calculated from the $N (=100)$ runs of bootstrapping sampling is a measure of the bias of the original calculation. This yielded higher $r^2_{\text{bootstrap}}$ value 0.929 for CoMSIA with standard error value of 0.192, supporting further the statistical validity of the developed models. The predicted inhibitory activities are listed in Table 1. The correlation between the experimental and predicted bioactivities is shown in Fig. 3. All the results indicate that the CoMSIA model is also fairly predictive.

3.1.3. Validation of 3D-QSAR models

The twenty six randomly selected compounds (marked with ^a in Table 1) were used as test set to verify the stability and predictive ability of the constructed CoMFA and CoMSIA models. The predicted pIC_{50} values with the QSAR models are found to be in good agreement with the experimental data within a statistically tolerable error range, with a predicted correlation coefficient, r^2_{pred} , equal to 0.781 for CoMFA and 0.583 for CoMSIA (Table 2). The correlation between the CoMFA and CoMSIA predicted activities and the experimental activities of the test set compounds are shown in Figs. 4 and 5, respectively. The test results indicate that the CoMFA and CoMSIA models can be reliably used in the design of new TS inhibitors.

3.1.4. Contour analysis

The visualization of the results of the CoMFA and CoMSIA models have been performed using the StDev*Coeff mapping option contoured by contribution. The default level of contour by contribution, 80 for favored region and 20 for disfavored region, was set during contour analysis.

3.1.4.1. CoMFA contour maps. CoMFA steric contour maps are shown in Figs. 6 and 7. The steric interaction is represented by green and yellow contours, while electrostatic interaction is denoted by red and blue contours. A large green contour was found near the plane of fluoro substituted phenyl ring of

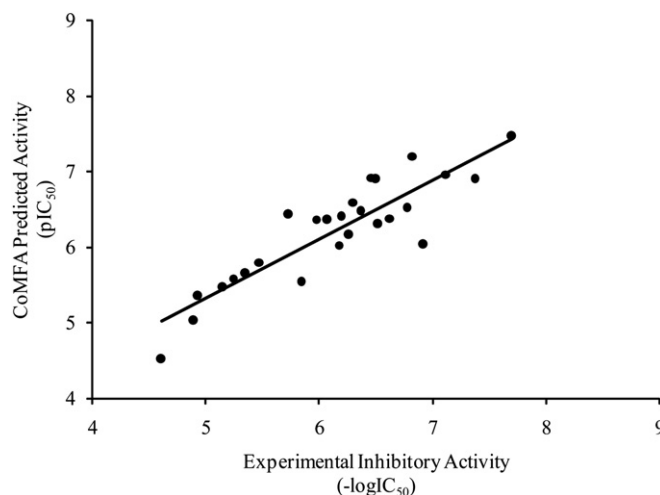


Fig. 4. Correlation between the experimental and CoMFA predicted activities of test set compounds.

compound **32** indicating that bulky substituents were preferred in this region (Fig. 6a). Similarly, in case of compound **35**, fluoro substituent group extends towards the large green contour which is the favorable region for bulkier groups, and hence compound **35** has high binding affinity (Fig. 6b). This may be the reason why compounds with halogen substituents in this area, e.g., compounds **32**, **34**, **35**, **37** and **39**, are more potent than molecules without any substituent at this particular position, such as compounds **7**, **8**, **11**, **13**, and **15**.

A large yellow contour is located around the unsubstituted phenyl ring of compound **13** and its CH_2COCH_3 substitution in between phenyl and quinazoline ring, suggesting that groups with low steric factor are required in this region to increase the activity (Fig. 7a). In Fig. 7b, a large yellow contour is found adjacent to ethyl substitution on the quinazoline ring of compound **15**, which indicates that groups with low steric factor are desired here to increase the binding affinity of the compound. This is possibly the reason as to why compounds **11**, **13**, **15** are less potent.

The CoMFA electrostatic contour plots for highly active compounds **32** and **35** are displayed in Figs. 8 and 9. A blue contours indicates that substituents should be electron deficient

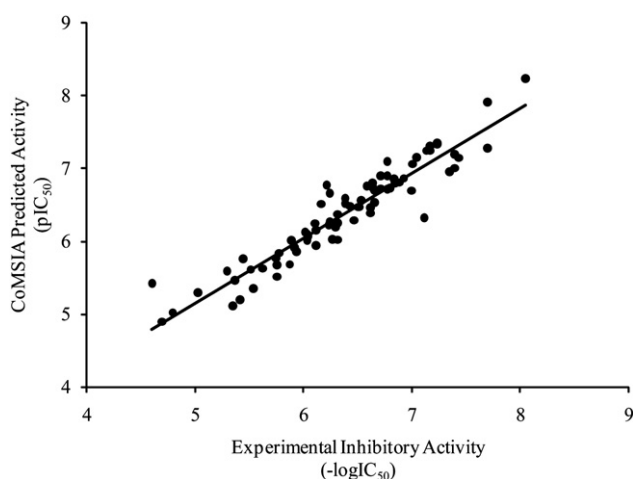


Fig. 3. Correlation between the experimental and CoMSIA predicted inhibitory activities of the training set compounds.

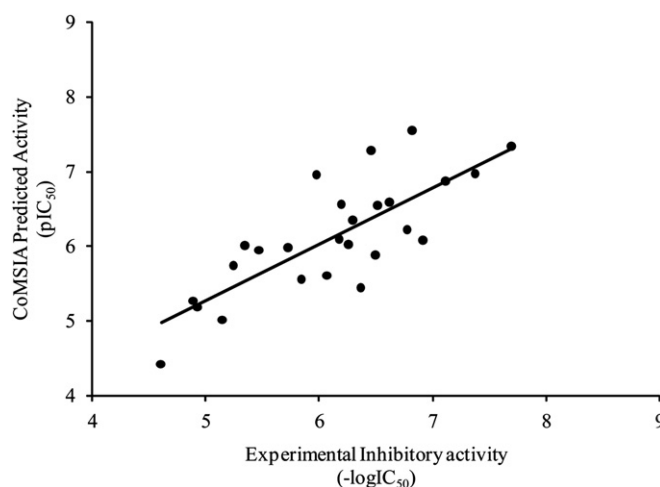


Fig. 5. Correlation between the experimental and CoMSIA predicted activities of test set compounds.

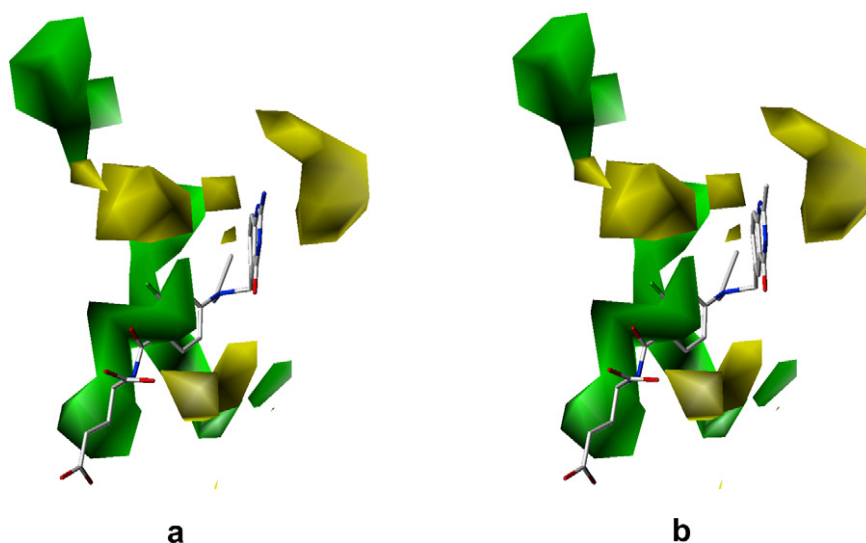


Fig. 6. CoMFA steric contours for highly potent compounds: (a) compound **32**, (b) compound **35**.

for high binding affinity with protein, while red color indicates that they should be electron rich for the same. Since a red contour was found overlapping with fluoro substituent at 3-position of compounds **32** and **35**, which is an electron rich functionality, compounds **32** and **35** exhibit high inhibitory potencies (Fig. 8a and b). Compounds with electron-rich substituents at this position are highly active as compared to those not having such substitutions, e.g., compound **13** and **15**. Similarly, a large blue contour is found overlapping with nitrogen between the quinazoline and phenyl ring in most of the compounds, indicating that the presence of the nitrogen is essential for the activity.

A large blue contour was found near the $-\text{CH}_2\text{COCH}_3$ substituent between phenyl ring and quinazoline ring of compound **13**, indicating that negatively charged groups are disfavored at this position, and that is why compound **13** is less potent than compound **32**, which has no electron-deficient substituent at this position (Fig. 9a). Additionally, compound **15**, which does not

have any substituent at the above mentioned position, is more active than compound **13** (Fig. 9b). The presence of positive charge favoring blue contours around the phenyl ring indicates that the addition of electron deficient groups may increase the binding affinity.

3.1.4.2. CoMSIA contour maps. The CoMSIA contours maps, derived using steric, electrostatic, hydrogen bond donor and hydrogen bond acceptor fields, are presented in Figs. 10–13. CoMSIA steric and electrostatic contours are more or less similar to those of the CoMFA. As in case of CoMFA, in case of CoMSIA, too, a large green contour was found overlapping the plane of fluoro-substituted phenyl ring of compound **32** to indicate that bulky substituents were preferred in this region (Fig. 10a). In case of compound **13**, a large yellow contour is found overlapping ethyl substituent at quinazoline ring. This is the region where bulkier substituents would not be preferred and that is why compounds **11–13**, **15** and **16** that have bulkier substituents are either less active or moderately

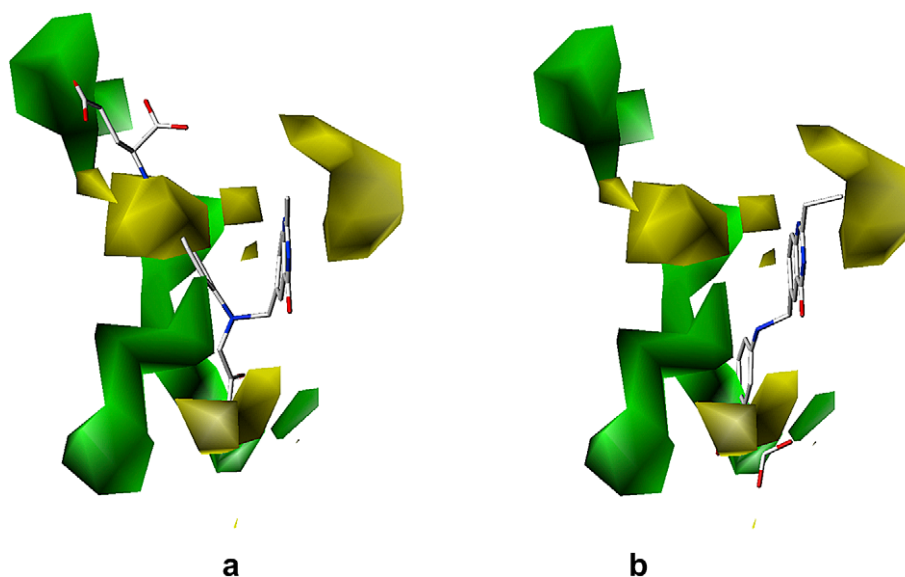


Fig. 7. CoMFA steric contour maps for less potent compounds: (a) compound **13**, (b) compound **15**.

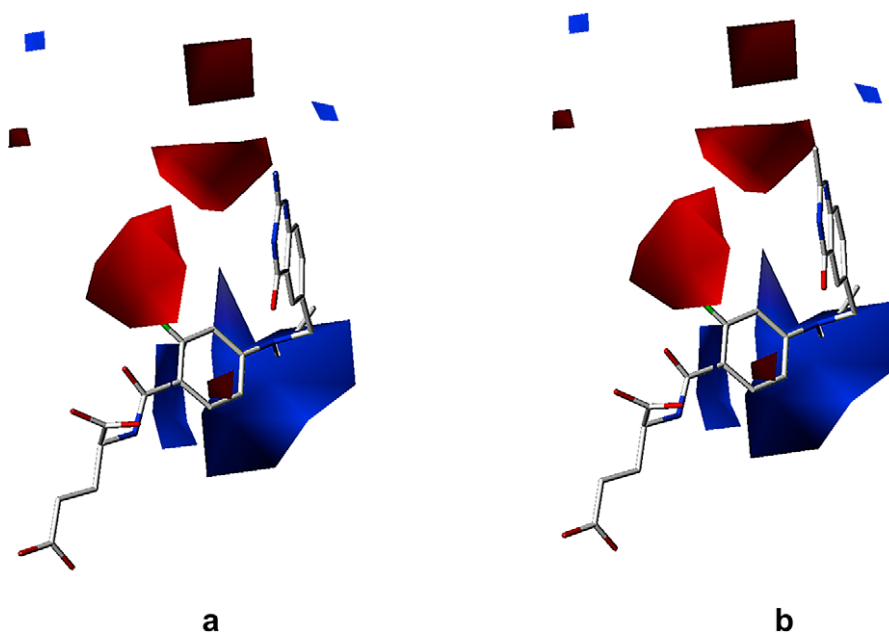


Fig. 8. CoMFA electrostatic contour maps of highly active compounds: (a) compound **32**, (b) compound **35**.

active (Fig. 10b). A yellow contour is also found near bulkier $-\text{CH}_2\text{COCH}_3$ substituent of compound **13** which is also not preferred at this position.

Fig. 11 shows the CoMSIA electrostatic fields denoted by red and blue contours. Red contours represent regions where negatively charged substituents are preferred on ligands and blue contours indicate regions where electron-rich substituents are unfavorable for the activity. There is one large blue contour overlapping the $-\text{CH}_2\text{COCH}_3$ group of compound **13**, which means that this group is not favored in this region and will lead to a decrease in the inhibitory potency (Fig. 11a). In Fig. 11b, a small

red contour is found overlapping fluorine substituent at phenyl ring in compound **32** and this is the favored position for electro-negative groups.

Hydrogen-bond donor contour maps from CoMSIA are shown in Fig. 12. Hydrogen bond donor-favored regions are represented by cyan contours and unfavorable regions by purple contours. CoMSIA hydrogen-bond donor contour map showed two cyan contours covering $-\text{NH}_2$ substituent at quinazoline ring in compound **32**, suggesting that substitution of hydrogen-bond donor groups in this region can be expected to improve the activity of molecule (Fig. 12a). In Fig. 12b, a small cyan contour is present near the

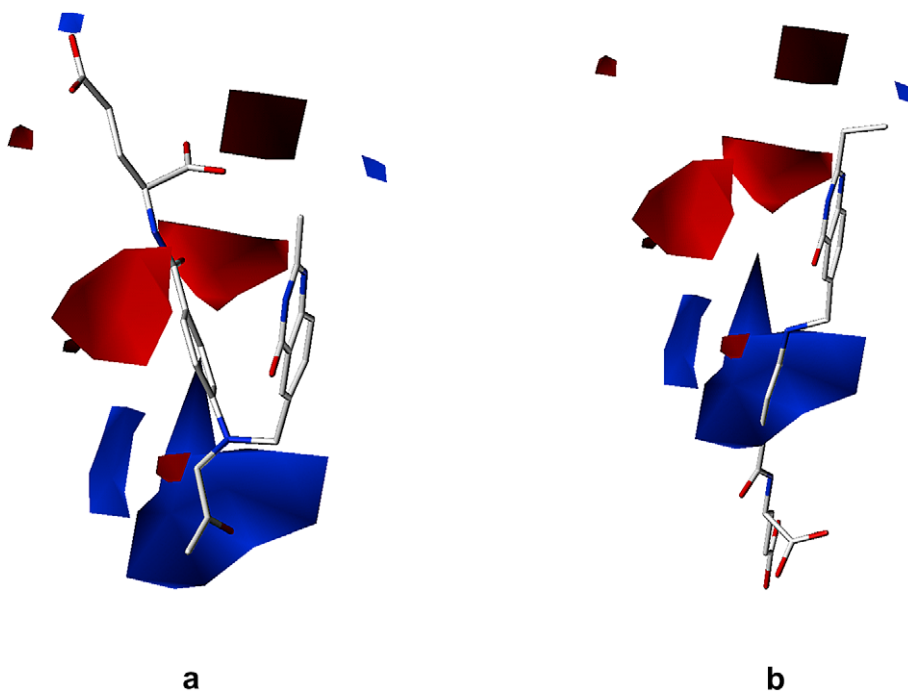


Fig. 9. CoMFA electrostatic contour maps for less active compounds: (a) compound **13**, (b) compound **15**.

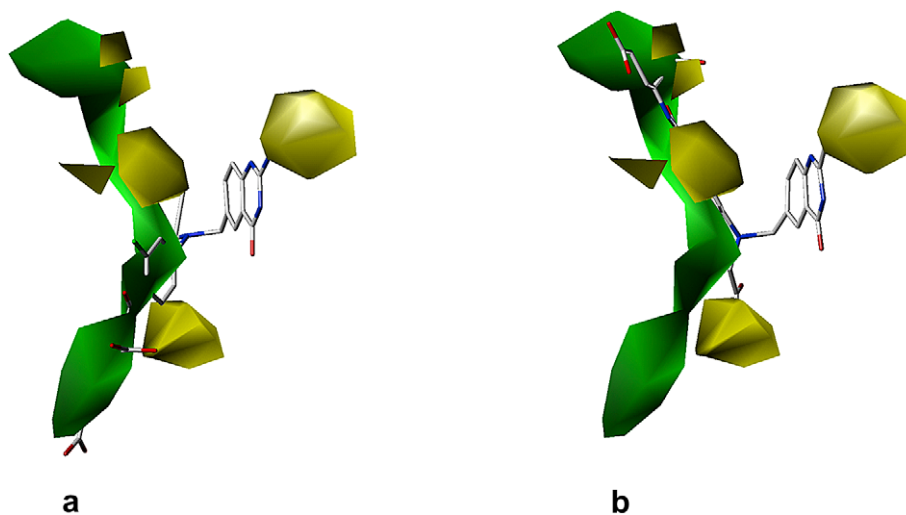


Fig. 10. CoMSIA steric contour maps: (a) compound 32, (b) compound 13.

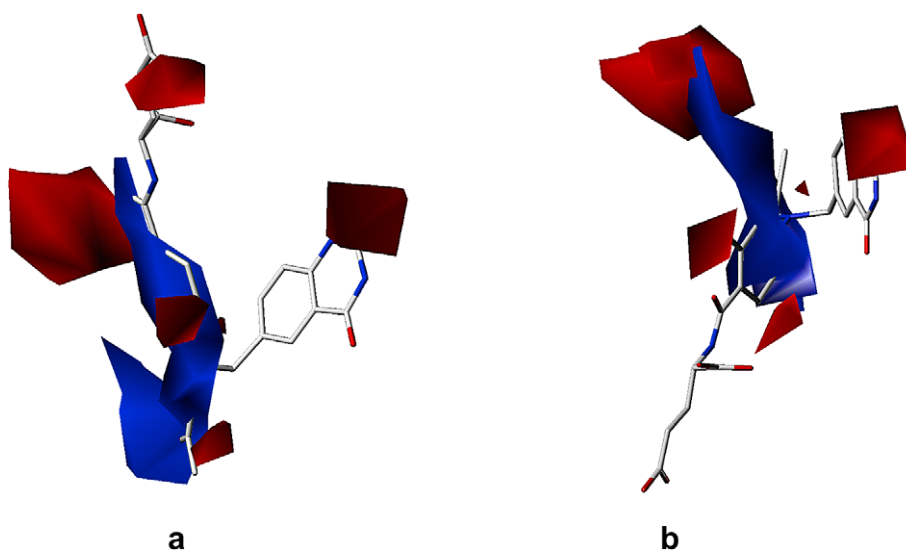


Fig. 11. CoMSIA electrostatic contour maps: (a) compound 13, (b) compound 32.

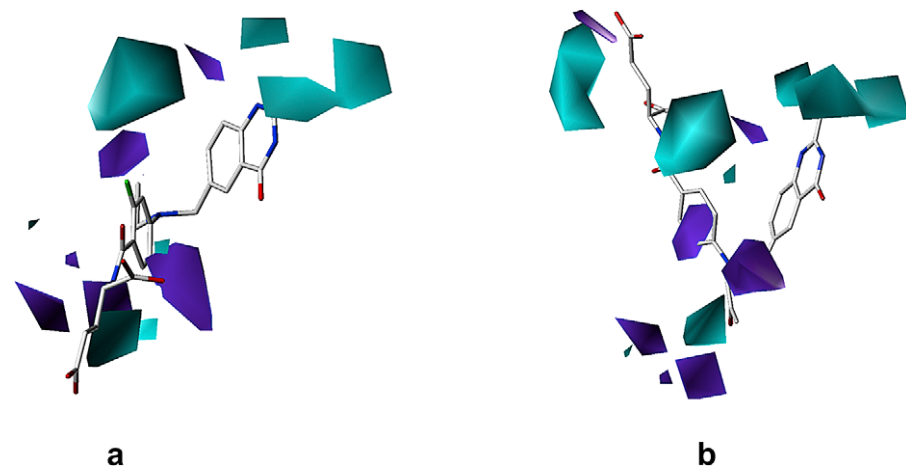


Fig. 12. CoMSIA hydrogen-bond donor contour maps: (a) compound 32, (b) compound 13.

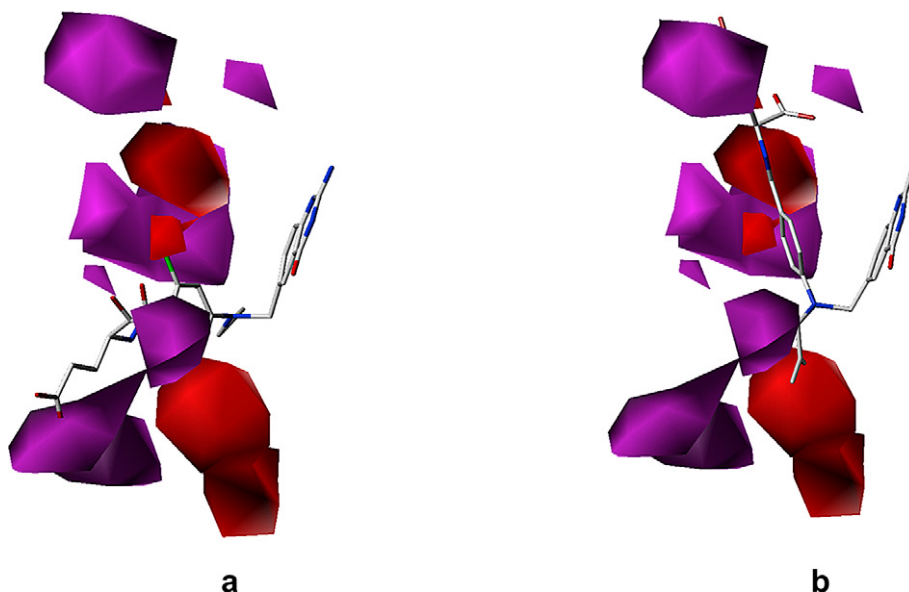


Fig. 13. CoMSIA hydrogen-bond acceptor contour maps: (a) compound 32, (b) compound 13.

–CH₂COCH₃ group of compound 13, indicating that hydrogen-bond acceptors are not preferred in this region, and that is why compound 13 is less active than compounds 15–18. Substituents which act as hydrogen-bond donors may lead to an increase in the activity of the molecule.

Fig. 13 displays the hydrogen-bond acceptor contour maps represented by magenta and red contours. Magenta contours indicate regions where hydrogen-bond acceptor substituents on ligands can be favored and the red ones represent areas where such substituents on inhibitors may be disfavored. In both Fig. 13a and b, lots of magenta contours are visible which display the importance of the presence of hydrogen-bond acceptor groups for TS inhibitory activity of quinazoline antifolate inhibitors. In Fig. 13b hydrogen-bond acceptor group CH₂COCH₃ of compound 13 extends towards disfavored red contour that may contribute to low activity of this compound.

4. Conclusion

In conclusion, our present studies have established that CoMFA and CoMSIA models are quite reliable to efficiently guide further modification in the molecules for obtaining better drugs. They have provided good statistical results in terms of q^2 and r^2 values for quinazoline antifolate derivatives. Both CoMFA and CoMSIA models provided the significant correlations of biological activities with steric and electrostatic fields, establishing the significance of steric and electrostatic fields in the selectivity and activity of the compounds. However, in comparison to CoMSIA the CoMFA method was found to provide a slightly better statistical model. However, the reliability of both the models was verified by the compounds in the testing set. The 3D-QSAR results revealed some important sites, where steric, electrostatic, hydrogen-bond donor, and hydrogen-bond acceptor modifications should significantly affect the bioactivities of the compounds. Thus useful clues to designing novel inhibitors of thymidylate synthase with high affinity for the treatment of cancer chemotherapy are provided.

Acknowledgement

The financial assistance in the form of an SRF to VS at Institute of Engineering and Technology, Lucknow, by CSIR, New Delhi, is thankfully acknowledged. Authors also acknowledge the facilities availed at Advanced Computer System Bioinformatics (ACS-bioINFORMATICS), Lucknow and wish to thank Mr. Ashutosh Kumar, CDRI, Lucknow, for his cooperation in many ways.

References

- [1] E.M. Berman, L.M. Werbel, *J. Med. Chem.* 34 (1991) 479–485.
- [2] V. Srivastava, A. Kumar, B.N. Mishra, M.I. Siddiqi, *Bioinformation* 3 (2008) 180–188.
- [3] A.L. Jackman, A.H. Calvert, *Ann. Oncol.* 6 (1995) 871–881.
- [4] L.R. Hughes, L. Ann Jackman, J. Oldfield, R.C. Smith, K.D. Burrows, P.R. Marsham, A.M. Bishop Joel, T.R. Jones, B.M. O'Connor, A.H. Calvert, *J. Med. Chem.* 33 (1990) 3060–3067.
- [5] L. Ann Jackman, P.R. Marsham, T.J. Thornton, A.M. Bishop Joel, B.M. O'Connor, L.R. Hughes, A.H. Calvert, T.R. Jones, *J. Med. Chem.* 33 (1990) 3067–3071.
- [6] P.R. Marsham, L. Ann Jackman, J. Oldfield, L.R. Hughes, T.J. Thornton, M.F. Bisset Graham, B.M. O'Connor, A.M. Bishop Joel, A.H. Calvert, *J. Med. Chem.* 33 (1990) 3072–3078.
- [7] P.R. Marsham, L.R. Hughes, L. Ann Jackman, J. Hayter, J. Oldfield, J.M. Wardleworth, A.M. Bishop Joel, B.M. O'Connor, A.H. Calvert, *J. Med. Chem.* 34 (1991) 1594–1605.
- [8] R.D. Cramer III, D.E. Patterson, J.D. Bunce, *J. Am. Chem. Soc.* 110 (1988) 5959–5967.
- [9] G. Klebe, U. Abraham, T. Mietzner, *J. Med. Chem.* 37 (1994) 4130–4146.
- [10] V. Srivastava, A. Kumar, B.N. Mishra, M.I. Siddiqi, *Bioinformation* 2 (2008) 384–391.
- [11] S.J. Cho, A. Tropsha, *J. Med. Chem.* 38 (1995) 1060–1066.
- [12] M.C. Clark, R.D. Cramer III, N.V. Opden Bosch, *J. Comput. Chem.* 10 (1989) 982–1012.
- [13] B.L. Bush, *J. Comput. Aided Mol. Des.* 7 (1993) 587–619.
- [14] H. Gohlke, M. Hendlich, G. Klebe, *J. Mol. Biol.* 295 (2000) 337–356.
- [15] R.D. Cramer, J.D. Bunce, D.E. Patterson, *Quant. Struct. Act. Relat.* 7 (1988) 18–25.
- [16] M. Clark, R.D. Cramer III, D.M. Jones, D.E. Patterson, P.E. Simeroth, *Tetrahed. Comput. Methodol.* 3 (1990) 47–59.
- [17] InsightII 2000.1, Accelrys Inc., San Diego, CA. <http://www.accelrys.com>.
- [18] Sybyl 7.1, Tripos Inc. 1699 South Hanley Road, St. Louis, Missouri 63144, USA. www.tripos.com.