Chem Biol Drug Des 2012; **79:** 935–942

Research Article

doi: 10.1111/j.1747-0285.2012.01351.x

3D-QSAR Studies on the Inhibitory Activity of Trimethoprim Analogues against *Escherichia coli* Dihydrofolate Reductase

Ramadoss Vijayaraj[1], Mekapothula Lakshmi Vasavi Devi[1], Venkatesan Subramanian[1,*] and Pratim Kumar Chattaraj[2,*]

[1]*Chemical Laboratory, Central Leather Research Institute, Council of Scientific and Industrial Research, Adyar, Chennai 600 020, India*
[2]*Department of Chemistry, Indian Institute of Technology, Kharagpur 721 302, India*
**Corresponding authors: Venkatesan Subramanian, subuchem@hotmail.com; Pratim Kumar Chattaraj, pkc@chem.iitkgp.ernet.in*

Three-dimensional quantitative structure activity relationship (3D-QSAR) study has been carried out on the *Escherichia coli* DHFR inhibitors 2,4-diamino-5-(substituted-benzyl)pyrimidine derivatives to understand the structural features responsible for the improved potency. To construct highly predictive 3D-QSAR models, comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) methods were used. The predicted models show statistically significant cross-validated and non-cross-validated correlation coefficient of r^2_{CV} and r^2_{nCV}, respectively. The final 3D-QSAR models were validated using structurally diverse test set compounds. Analysis of the contour maps generated from CoMFA and CoMSIA methods reveals that the substitution of electronegative groups at the first and second position along with electropositive group at the third position of R2 substitution significantly increases the potency of the derivatives. The results obtained from the CoMFA and CoMSIA study delineate the substituents on the trimethoprim analogues responsible for the enhanced potency and also provide valuable directions for the design of new trimethorpim analogues with improved affinity.

Key words: CoMFA, CoMSIA, DHFR, pyrimidine derivatives, QSAR, trimethoprim

Received 2 July 2011, revised 23 December 2011 and accepted for publication 21 January 2012

Dihydrofolate reductase (DHFR) catalyzes the reduction of 5,6-dihydrofolate (DHF) to 5,6,7,8-tetrahydrofolate (THF) using NADPH as a cofactor (1,2). The DHFR also carries out the reduction of folate to DHF; however, the reduction rate is comparatively slower than that of DHF to THF. The resulting THF and other reduced folates act as cofactors for various one-carbon transfer reactions such as the biosynthesis of purines, thymidylate, panthothenate, and methionine that are involved in the cell cycle. The DHFR enzyme is a potential drug target for various conditions such as bacterial infection, malaria, and cancer because of its association with the DNA synthesis (3,4). The essential role of DHFR in DNA synthesis and in a variety of anabolic pathways makes it a common target for antiproliferative therapeutics. Because of its biological and pharmacological importance, DHFR has been the subject of intensive structural and kinetic investigation over many years, serving as a paradigm of enzymatic systems in many experimental and theoretical studies (5–12). The thymidylate synthase catalyzes the conversion of deoxyuridylate (dUMP) to deoxythymidylate (dTMP) in which the dihydrofolates are generated as by-product (13). The drugs targeted to DHFR are known as antifolates, and trimethoprim (TMP) is an antifolate that has high specificity to bacterial DHFR over mammalian DHFR (3,14–17). The exposure of *Escherichia coli* with TMP leads to rapid accretion of dihydrofolate, and it is further oxidized into folates (1,8). The inhibitory activity of TMP analogous of 65 pyrimidine derivatives has been experimentally measured against *E. coli* DHFR (ecDHFR) (18). The X-ray crystallography technique has been extensively used to study the interactions of various TMP analogous with the ecDHFR (19,20).

The majority of DHFR inhibitors described in the literature are derived from various substitutions on 2,4-diaminopyrimidine derivatives. Several QSAR studies have been performed on the pyrimidine derivatives of ecDHFR (21–25). Czaplinski *et al.* have synthesized TMP analogues with various 4-anilinoalkoxy moieties at the para position of phenyl ring of pyrimidine derivatives and tested against ecDHFR (26). In addition, various QSAR models have also been developed for the same set of derivatives. The nonlinear QSAR study on the pyrimidine derivatives unveils important features of the data set and identifies three distinct classes of molecules based on the mean activity values (27). King *et al.* (28) have carried out QSAR study using the inductive logic programming (ILP) procedure to discover new indicator variables for pyrimidine derivatives against ecDHFR. The 3D-QSAR study on the TMP analogues using the tensor representation method has unveiled the geometrical information on the ecDHFR indirectly at the early stages when the experimental geometries of the target are unavailable (29). The application of an adaptive neuro-fuzzy inference system (ANFIS)-based QSAR models has been successfully derived for the pyrimidine derivatives (30). The same derivatives have been used to develop several neural network QSPR models using the Hansch

substituent constant predictor (31). Summerfield *et al.* have carried out high-throughput screening of structurally diverse set of small molecules to identify new set of inhibitors for ecDHFR. They have identified several novel compounds that make interaction with the M20 loop of DHFR, and further the interaction was characterized using X-ray crystallography and SAR analysis (9,11). Recently, Shine *et al.* (32) have estimated the relative binding free energy of pyrimidine derivatives (TMP and its analogues) with the ecDHFR based on a free energy variational principle, and further they have derived a QSAR model using the experimental free energy of binding.

Despite numerous QSAR studies have been carried out with different approaches on the pyrimidine derivatives of ecDHFR, the application of 3D-QSAR study that yields valuable information on the design and development of new ligands is scarce. Despite there are numerous inhibitors successfully targeted against DHFR to combat a variety of bacterial infection and caner, still there is a need to improve the inhibitory activity of TMP analogues. Hence, the 3D-QSAR approach has been applied on these compounds using the comparative molecular field analysis (CoMFA) (33,34) and comparative molecular similarity indices analysis (CoMSIA) (35) methods. The results obtained from this study could help in understanding various structural features responsible for the activity of pyrimidine derivatives and also provide useful information for the design and synthesis of novel inhibitors.

Materials and Methods

Data set

The 2,4-diamino-5-(substituted-benzyl) pyrimidine derivatives of ecDHFR inhibitors were used for the 3D-QSAR study (18). The data set of 65 compounds with the corresponding inhibition constant (K_i) values converted into pK_i is shown in the Table 1. The data set is further divided into training set and test set which includes 55 and 10 compounds, respectively. The test set compounds were randomly chosen to represent the diversity in the structure and activity values. The training set compounds were used to build the 3D-QSAR models by considering the pK_i values as dependent variables, and the predicted models were evaluated using the test set compounds.

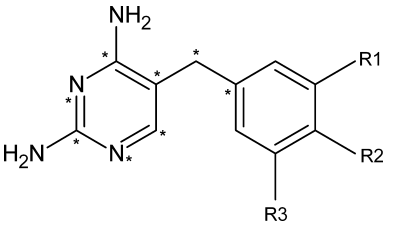
Molecular modeling and alignment

The three-dimensional structures of all compounds were prepared using the SYBYL 8.1.1 software of Tripos (36). Partial atomic charges were calculated by the Gasteiger–Huckel method, (37) and energy minimizations were performed using the Tripos force field and the Powell conjugate gradient algorithm with a convergence criterion of 0.005 kcal/mol Å. The compound **65** that is having lowest energy was chosen as template molecule. The template molecule was used to superimpose all the compounds with the application of database alignment method. The atoms selected for the common substructure alignment are shown with asterisks in Table 1.

CoMFA and CoMSIA analysis

The following standard procedures were used to calculate the CoMFA and CoMSIA descriptors. Each compound was inserted into a

Table 1: Data set of pyrimidine derivatives used for the 3D-QSAR study



Compound	R1	R2	R3	pK_i
1	CH_2CH_3	CH_2CH_3	CH_2CH_3	7.82
2[a]	OCH_3	$OCH_2CH_2OCH_3$	OCH_3	8.35
3	OCH_3	OCH_3	OCH_3	8.08
4[a]	OCH_3	$N(CH_3)_2$	OCH_3	7.71
5	OCH_3	Br	OCH_3	8.18
6	OCH_3	SCH_3	OCH_3	8.07
7	OCH_3	$O(CH_2)_7CH_3$	OCH_3	7.20
8	CH_2OH	H	CH_2OH	6.31
9	OCH_3	H	OCH_3	7.71
10[a]	OCH_2CH_3	H	OCH_2CH_3	7.69
11	OCH_2CH_3	H	$OCH_2CH_2CH_3$	7.69
12[a]	$OCH_2CH_2CH_3$	H	$OCH_2CH_2CH_3$	7.41
13	CH_3	H	CH_3	7.04
14	H	OH	OH	6.46
15	H	$NHCOCH_3$	NO_2	6.97
16[a]	H	$OCH_2CH_2OCH_3$	$OCH_2CH_2OCH_3$	7.22
17	H	OCH_3	OCH_3	7.72
18[a]	H	OCH_3	OH	6.84
19	H	OSO_2CH_3	OCH_3	7.94
20[a]	H	OH	OCH_3	7.54
21[a]	H	$OCH_2CH_2OCH_3$	OCH_3	7.77
22[a]	H	$OCH_2C_6H_5$	OCH_3	7.53
23	H	OCH_3	OSO_2CH_3	7.80
24	$OCH_2C_6H_5$	OCH_3	H	7.66
25	CF_3	OCH_3	H	7.69
26	$O(CH_2)_7CH_3$	OCH_3	H	7.16
27	OCH_2CH_3	$OCH_2C_6H_5$	H	7.35
28	H	H	OCH_2CONH_2	6.57
29	H	H	CH_2OH	6.28
30	H	H	OSO_2CH_3	6.92
31	H	H	CH_2OCH_3	6.59
32	H	H	OH	6.47
33	H	H	$OCH_2CH_2OCH_3$	6.53
34	H	H	OCH_3	6.93
35	F	H	H	6.23
36	CH_3	H	H	6.70
37	Cl	H	H	6.65
38	Br	H	H	6.96
39	CF_3	H	H	7.02
40[a]	$CO(CH_2)_3CH_3$	H	H	6.55
41	I	H	H	7.23
42	$O(CH_2)_3CH_3$	H	H	6.82
43	$OCH_2C_6H_5$	H	H	6.99
44	$O(CH_2)_5CH_3$	H	H	6.86
45	$O(CH_2)_6CH_3$	H	H	6.39
46	$O(CH_2)_7CH_3$	H	H	6.25
47	H	NH_2	H	6.30
48	H	$NHCOCH_3$	H	6.89
49	H	OSO_2CH_3	H	6.60
50	H	OH	H	6.45
51	H	$OCH_2CH_2OCH_3$	H	6.40

Table 1: Continued

Compound	R1	R2	R3	pK_i
52	H	NO_2	H	6.20
53	H	OCH_3	H	6.82
54	H	F	H	6.35
55	H	$N(CH_3)_2$	H	6.78
56	H	CH_3	H	6.48
57	H	Cl	H	6.45
58	H	Br	H	6.82
59	H	OCF_3	H	6.57
60	H	$O(CH_2)_3CH_3$	H	6.89
61	H	$OCH_2C_6H_5$	H	6.89
62	H	$O(CH_2)_5CH_3$	H	6.07
63	H	$O(CH_2)_6CH_3$	H	6.10
64	H	C_6H_5	H	6.93
65	H	H	H	6.18

[a]Test set compounds.

three-dimensional lattice with grid points separated by 2.0 Å in *x*, *y*, and *z* directions and an extension of 4 Å beyond the aligned molecules in all directions. For the CoMFA analysis, the steric (van der Waals) and electrostatic (Coulombic) field descriptors were calculated at all lattice points by adding the individual energy interactions between each atom of the compound using an sp^3 carbon probe atom with +1 charge and 1.52 Å van der Waals radius. The CoMFA cutoff values were set to 30 kcal/mol for both steric and electrostatic fields. The similarity indices were computed using a probe atom with 1 Å radius and +1 atomic charge. The distance dependence between the grid point and each atom of molecule was determined by Gaussian function through the similarity indices calculated at all grid points, and a default value of 0.3 was used as an attenuation factor α. The value displayed for each CoMSIA column is the root mean square deviation of the points that constitute the field.

Partial least squares analysis

The descriptors calculated from the CoMFA and CoMSIA analysis were used as independent variables, and the pK_i values were used as dependent variables in partial least squares (PLS) (38) analysis to derive 3D-QSAR models using the standard implementation in the SYBYL package. Linear regression equations were obtained by the PLS technique, correlating changes in the steric and electrostatic fields (for CoMFA) and similarity fields (for CoMSIA) with changes in the pK_i values of the inhibitor sets. The cross-validation in PLS analysis was performed by leave-one-out (LOO) procedure to find the optimal number of components in building the regression models and to check the statistical significance of the model. The LOO method effectively evaluates the model by excluding one compound from the data set at a time and measuring the internal predictive ability of the newly created model. The quality of the model was expressed as the cross-validated correlation coefficient r^2_{CV}. The optimal number of components obtained from the LOO method was used to derive the final model using non-cross-validated PLS analysis. The correlation coefficient (r^2_{nCV}) was used to measure the quality of the obtained final model. Column filtering, for any column of computed energies with a variation less than 2.0 kcal/mol, was applied as needed to reduce computation time without negatively affecting the quality of the models. The predictive power of the models was evaluated first by leave-one-out (LOO) cross-validation (7,8). The non-cross-validated coefficient, r^2_{CV}, was calculated using eqn 1.

$$r^2_{nCV} = 1 - \frac{\sum_{i=1}^{n} (y_{i,\text{observed}} - y_{i,\text{predicted}})^2}{\sum_{i=1}^{n} (y_{i,\text{observed}} - y_{\text{mean}})^2} \quad (1)$$

where $y_{i,\text{predicted}}$, $y_{i,\text{observed}}$, and y_{mean} are predicted, actual, and mean values of the target property (pK_i), respectively. To test the utility of the model as a predictive tool, an external set of compounds with known activities but not used in model generation (the test set) was predicted. The predictive squared correlation coefficient, r^2_{pred}, calculated by using eqn 2, was based on molecules from the test set, and it was used to evaluate the predictive power of the CoMFA and CoMSIA models.

$$r^2_{\text{pred}} = (\text{SD} - \text{PRESS})/\text{SD} \quad (2)$$

where SD is the sum of the squared deviations between the actual activities of the compounds in the test set and the mean activity of the compounds in the training set, and 'PRESS' is the sum of the squared deviations between predicted and actual activities for every compound in the test set. A r^2_{pred} value of 1 signifies that the CoMFA model is perfectly predictive for the test set, while prediction of a mean value of the training set for every member in the test set yields a predictive $r^2 = 0$.

Results and Discussion

CoMFA and CoMSIA statistical results

The CoMFA and CoMSIA models were developed using the steric and electrostatic descriptors. The statistical results obtained from these models are presented in Table 2. The leave-one-out cross-validated correlation coefficients (r^2_{CV}) for CoMFA and CoMSIA models are 0.75 and 0.73, respectively. The supporting information Table S1 shows that the model attains significant r^2_{CV} value with six components. The non-cross-validated correlation coefficients (r^2_{nCV}) for CoMFA and CoMSIA models are 0.95 and 0.88, respectively. The low standard error of estimation (SEE) and high F values confirm that the models are statistically significant to predict the inhibitory activity of pyrimidine derivatives. The predicted steric and electrostatic contributions from the CoMFA and CoMSIA analysis are 64.5% and 35.5% as well as 57.1% and 42.9%, respectively. Evidences indicate that the steric contributions are slightly higher than the electrostatic contributions. To validate the predicted models, the external test set compounds were used to assess the predictive power of the models. The test set compounds are not included in the prediction of CoMFA and CoMSIA models. The Tables 3 and 4 contain the predicted activities (pK_i) for the training set and test set compounds. The predictive correlation coefficient values (r^2_{pred}) for the CoMFA and CoMSIA models are 0.69 and 0.59, respectively. The graphical results for the experimental verses predicted pK_i values of both training and test sets are displayed in Figure 1.

Table 2: Statistical summary of CoMFA and CoMSIA models

Parameter	CoMFA	CoMSIA
r^2_{CV}	0.75	0.73
r^2_{nCV}	0.95	0.88
r^2_{pred}	0.69	0.59
SEE	0.14	0.21
F-test	145.01	59.93
NC	6	6
Steric contribution (%)	64.50	57.10
Electrostatic contribution (%)	35.50	42.90

r^2_{CV}, cross-validated correlation coefficient by PLS LOO method, r^2_{nCV}, non-cross-validated correlation coefficient, $r^2_{pred.}$, predicted correlation coefficient for the test set compounds, SEE, standard error of estimate, NC, optimum number of components as obtained from the cross-validation study; CoMFA, comparative molecular field analysis; CoMSIA, comparative molecular similarity indices analysis.

Electrostatic contour map of CoMFA and CoMSIA

The electrostatic red contours derived from the CoMFA and CoMSIA imply the presence of negative charge favors the inhibitory activity. On the other hand, the blue contours indicate the regions where the occurrence of positive charge enhances the activity. In the electrostatic field, the red and blue contours represent 80% and 20% levels of contributions, respectively. The electrostatic contour map obtained for the CoMFA and CoMSIA models is displayed in Figure 2. The large blue contours observed at the second and third positions of R1 and R3 substitutions indicate that the presence of electropositive groups at these positions increases the potency (Figure 2A). This observation is in agreement with the previous QSAR study based on the hydrophobic and molar refractivity constants of various R1, R2, and R3 substituents (18). The optimum value for the hydrophobic constant is derived as 0.73 for R1 and R3 substitution. In Figure 2, the first positions of the R1 and R2 substitutions are surrounded by a large red contour. A small red contour is located at the first position of R3 and second position of R2 substitutions (Figure 2A). It shows that the presence of electronegative groups at these positions favors the activity of the inhibitors. Figure 2A also illustrates a small blue contour at the third position of the R2 substitution. The CoMSIA electrostatic contour shows (Figure 2B) small and large red surfaces at the first position of R3 and R1 substitutions, respectively. Figure 2B also displays a large blue surface at the second and third positions of the R3 substitution and a small blue surface at the third position of the R2 substitution. The observed electrostatic contour maps from the CoMFA and CoMSIA analysis for R1 and R3 substitutions clearly demonstrate that the first position is highly preferred by the electronegative groups, whereas the second and third positions are predominantly favored by the electropositive groups. The substitution of $-OCH_3$ group at the R1 and R3 substitutions induces positive potential at the second position, and further lengthening of the chain with hydrocarbons enhances the positive potential at this position. This can be further interpreted that the presence of O atom at the first position attracts the electron clouds from the $-CH_3$ as well as from the adjacent phenyl group and thus leading to negative potential at the O atom and positive potential at the $-CH_3$ group. In fact, the

Table 3: Observed and predicted activities (pK_i) of training set CoMFA and CoMSIA models

		CoMFA		CoMSIA	
Compound	Exp. pK_i	Predicted	Residual	Predicted	Residual
1	7.82	7.72	0.10	7.69	0.13
3	8.08	8.31	0.23	8.37	0.29
5	8.18	8.01	0.18	7.90	0.28
6	8.07	8.22	0.15	8.24	0.17
7	7.20	7.21	0.01	7.21	0.01
8	6.31	6.40	0.09	6.62	0.31
9	7.71	7.68	0.03	7.72	0.01
11	7.69	7.69	0.00	7.82	0.13
13	7.04	6.98	0.06	6.84	0.20
14	6.46	6.47	0.01	6.69	0.23
15	6.97	6.90	0.07	6.98	0.00
17	7.72	7.67	0.05	7.60	0.12
19	7.94	7.61	0.33	7.70	0.24
23	7.80	7.78	0.02	7.53	0.27
24	7.66	7.46	0.20	7.69	0.03
25	7.69	7.72	0.03	7.57	0.12
26	7.16	7.11	0.05	7.05	0.11
27	7.35	7.48	0.13	7.26	0.09
28	6.57	6.59	0.02	6.63	0.06
29	6.28	6.12	0.16	6.40	0.12
30	6.92	6.99	0.07	6.95	0.02
31	6.59	6.36	0.23	6.42	0.17
32	6.47	6.40	0.07	6.50	0.03
33	6.53	6.61	0.08	6.49	0.04
34	6.93	6.94	0.01	7.05	0.12
35	6.23	6.31	0.08	6.60	0.37
36	6.70	6.87	0.17	6.61	0.09
37	6.65	6.61	0.04	6.54	0.11
38	6.96	6.83	0.13	6.54	0.42
39	7.02	7.12	0.10	6.95	0.07
41	7.23	6.91	0.32	6.56	0.67
42	6.82	6.90	0.08	6.98	0.16
43	6.99	7.04	0.05	7.06	0.07
44	6.86	6.74	0.12	6.64	0.22
45	6.39	6.54	0.15	6.48	0.09
46	6.25	6.30	0.04	6.38	0.13
47	6.30	6.31	0.01	6.56	0.26
48	6.89	6.90	0.01	6.92	0.03
49	6.60	6.85	0.25	7.00	0.40
50	6.45	6.33	0.12	6.46	0.01
51	6.40	6.69	0.29	6.44	0.04
52	6.20	6.23	0.03	6.21	0.01
53	6.82	6.82	0.00	6.91	0.09
54	6.35	6.35	0.00	6.49	0.14
55	6.78	6.74	0.04	6.75	0.03
56	6.48	6.49	0.01	6.63	0.15
57	6.45	6.49	0.04	6.48	0.03
58	6.82	6.64	0.18	6.49	0.33
59	6.57	6.70	0.13	6.67	0.10
60	6.89	6.77	0.12	6.71	0.18
61	6.89	7.08	0.19	6.88	0.01
62	6.07	6.25	0.18	6.18	0.10
63	6.10	6.11	0.01	5.96	0.14
64	6.93	6.95	0.02	7.06	0.13
65	6.18	6.12	0.06	6.36	0.18

Exp., experimental activity; CoMFA, comparative molecular field analysis; CoMSIA, comparative molecular similarity indices analysis.

Table 4: Observed and predicted activities (pK_i) of test set CoMFA and CoMSIA models

		CoMFA		CoMSIA	
Compound	Exp. pK_i	Predicted	Residual	Predicted	Residual
2	8.35	7.96	0.39	7.87	0.48
4	7.71	8.32	0.61	8.51	0.80
10	7.69	7.79	0.10	7.81	0.12
12	7.41	7.64	0.23	7.79	0.38
16	7.22	7.09	0.13	6.62	0.60
18	6.84	7.12	0.28	7.06	0.22
20	7.54	7.10	0.44	7.29	0.25
21	7.77	7.39	0.38	7.14	0.63
22	7.53	7.72	0.19	7.56	0.03
40	6.55	7.35	0.80	7.14	0.58

CoMFA, comparative molecular field analysis; CoMSIA, comparative molecular similarity indices analysis.

negative potential at the O atom of $-OCH_3$ group is further strengthened by the mixing of O atom lone pair electrons with the π-electron clouds of the phenyl ring. The absence of this intrinsic property is evidenced by the lower activity for the compounds **13** and **65** that consist of $-CH_3$ and H substitutions at R1 and R3 position. The absence of O atom at the first position leads to less positive potential at the $-CH_3$ and H groups of R1 and R3 substitution which significantly reduces the potency of these compounds. It is further evident by the presence of negative potential at the first position and absence of positive potential at the second position or negative potential at the second position of R1 and R3 substitutions such as compounds **8**, **14**, **15**, **18**, **25**, **29**, **32**, **35**, **37**–**39**, and **41** that show considerably lower activity. In addition, the substitution of $-OCH_3$ group only at the R3 position retains moderate activity (compounds **17**, **19**–**22**) when compared to the $-OCH_3$ groups substituted at both R1 and R3 positions (compounds **2**–**7**, **9**). It is interesting to note that compounds **47**–**65** have hydrogen substitution at the R1 and R3 positions, and hence, it shows significantly lower activity. It can be seen from the compound **8** that the substitution of $-CH_2OH$ at the R1 and R3 position, which is deficient of the electropositive group at the second position, leads to lower activity. This observation reiterates that the R1 and R3 substitutions play a dominant role in the activity of pyrimidine derivatives that prefers electronegative groups at the first position and electropositive groups at the second and third positions.

The observed red contour at the first position of the R2 substitution shows that the presence of electronegative group improves the potency, and further a small red contour is located at the second position of the R2 substitution. The presence of electronegative group at the second position of R2 substitution gives moderate activity that is apparent from the activity of the compound **19**. In addition, a small blue surface located at the third position of R2 substitution indicates the presence of electropositive group enhances the activity. It can be noted that the substitution of different halogen groups at R1 (compounds **35**, **37**, **38**, and **41**) and R2 (compounds **5**, **54**, **57**, and **58**) positions exhibits variable activity. It can be observed that the activity decreases with the increase in the electronegativity of the halogen substituent. The order of the halogen groups at R1 and R2 position that favors the activity is

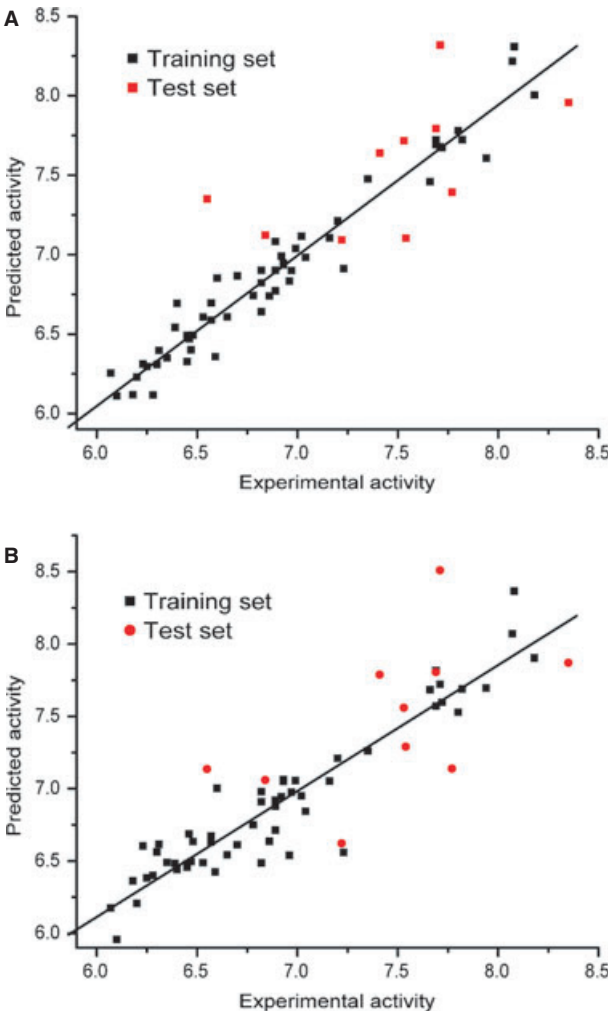


Figure 1: Comparison of experimental versus predicted pK_i values obtained from CoMFA (A) and CoMSIA (B) study.

I > Br > Cl > F. This observation can be correlated with the absence of the positive potentials at the second position of R1 substitution. In the case of halogen group substitution at the R2 position, the compound **5** is substituted with Br at the R2 first position and $-OCH_3$ at the R1 and R3 positions, which shows the inhibitory activity of 8.18, while the other compounds with halogen substitutions at the R2 position, compounds **54**, **57**, and **58**, contain H substitution at the R1 and R3 with comparatively less activity than that of compound **5**. It is obvious on the basis of favored order of halogen groups at R1 and R2 with improved potency, the present study suggests that the substitution of I group at the R2 position of compound **5** could further improve the inhibitory potency.

Steric contour map of CoMFA and CoMSIA

The steric contour maps observed using CoMFA and CoMSIA analysis are presented in Figure 2. The green and yellow contours stand for the sterically favorable (80% contribution) and unfavorable (20% contribution) regions, respectively. The large green contours

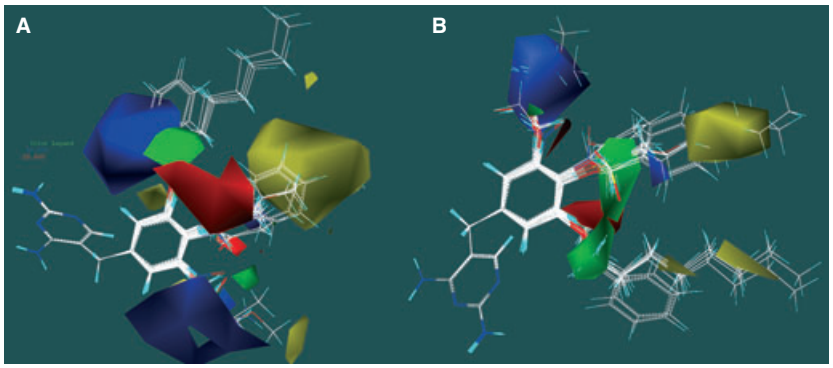


Figure 2: Electrostatic and steric contour maps of CoMFA (A) and CoMSIA (B). Electrostatic fields are represented by blue and red contours. Blue contours indicate regions where the electropositive groups increase activity, while red contours indicate regions where electronegative groups increase activity. In steric fields, green contours indicate regions where bulky groups increase activity, while yellow contours indicate regions where bulky groups decrease activity.

observed at second and third positions of R1 and R2 substitutions reveal that the bulky group substituents at these regions improve the inhibitory activity. It is further supported by the moderate activity observed for the bulky group substituents at the R1 and R2 positions such as compounds **4**, **19**, **22**, **24**, and **27**. Several yellow contours located at the terminals of the R1 substitution (Figure 2A,B) highlights that the lengthening of the chain does not favor the potency of the inhibitors. It is clear from the lower activity of the compounds **26**, **44**–**46** with extended chains at the R1 substitution. The large yellow surfaces (Figure 2) located at the terminal of the R2 substitution reveals that the extended chain substituents at this position reduce the potency of the compounds. The extended chain substituents at the R2 position such as compounds **51**, **60**–**63** possess lower activity. In contrary, the compounds **7**, **16**, and **21** with extended chain substituents at the R2 position exhibit moderate activity. The reduced activity observed for the compounds **51**, **60**–**63** can be attributed because of the presence of hydrogen substitutions at the R1 and R3 positions. Figure 2A shows a large yellow contour at the first position of the R3 substitution which signifies that the bulky group substituents are not favored at this position. In the case of compound **15**, the substitution of branched chain substituent $-NO_2$ at the R3 position does not favors the potency of the compound. A small green contour placed at the second position of the R3 substitution highlights that the occurrence of the bulky group at this position favors the activity of the compound.

Conclusion

In the present work, 3D-QSAR study has been carried out to understand the structural features responsible for the potency of the pyrimidine derivatives against ecDHFR. The CoMFA and CoMSIA models were established using the steric and electrostatic descriptors, and the statistical significance of the obtained models is ensured with the high values of r^2_{CV} and r^2_{nCV}. The reliability of the predicted models is validated using the test set compounds, and the predictive correlation coefficient (r^2_{pred}) confirms the high predictive power of the predicted models. It can be observed from the statistical results that the CoMFA model has better predictive ability than the CoMSIA model. The predicted CoMFA/CoMSIA field distributions show reasonable consistency. The interpretation of various contour maps reveals that the substitution of electronegative groups at the first and second position along with electropositive group at the third position of R2 substitution significantly increases the potency of the derivatives. The presence of negative potential at the first position and positive potential at the second position of R1 and R3 substitution favors the inhibitory activity. The substitution of extended aliphatic chains at the R2 position does not favor the inhibitory activity. The investigation of pyrimidine derivatives of ecDHFR using 3D-QSAR method provides significant structural information required to optimize the inhibitory activity.

Acknowledgment

We thank the Council of Scientific and Industrial Research (CSIR), New Delhi for financial assistance.

References

1. Blakley RL, Benkovic SJ (Eds.), (1984) Chemistry and Biochemistry of Folates: Dihydrofolate Reductase, Vol. 1. New York, NY: Wiley; pp. 191–251.
2. Schnell JR, Dyson HJ, Wright PE (2004) Structure, dynamics and catalytic function in dihydrofolate reductase. Annu Rev Biophys Biomol Struct;33:119–140.
3. Hitchings AH (1989) Selective inhibitors of dihydrofolate reductase (nobel lecture). Angew Chem Int Ed;28:879–885.
4. Schweitzer BI, Dicker AP, Bertino JR (1990) Dihydrofolate reductase as a therapeutic target. FASEB J;4:2441–2452.
5. Roth B, Aig E, Rauckman BS, Srelitz JZ, Phillips AP, Ferone R, Bushby SRM, Siegel CW (1981) 2,4-diamino-5-benzylpyrimidines and analogs as antibacterial agents. 5. 3′,5′-dimethoxy-4′-substituted-benzyl analogs of trimethoprim. J Med Chem;24:933–941.
6. Li RL, Dietrich SW, Hansch C (1981) Quantitative structure-selectivity relationships. Comparison of the inhibition of *Escherichia*

coli and bovine liver dihydrofolate reductase by 5-(substituted-benzyl)-2,4-diaminopyrimidines. J Med Chem;24:538–544.
7. Roth B, Rauckman BS, Ferone R, Baccanari DP, Champness JN, Hyde RM (1987) 2,4-Diamino-5-benzylpyrimidines as antibacterial agents. 7. Analysis of the effect of 3,5-dialkyl substituent size and shape on binding to four different dihydrofolate reductase enzymes. J Med Chem;30:348–356.
8. Quinlivan EP, McPartlin J, Weir DG, Scott J (2000) Mechanism of the antimicrobial drug trimethoprim revisited. FASEB J;14:2519–2524.
9. Zolli-Juran M, Cechetto JD, Hartlen R, Daigle DM, Brown ED (2003) High throughput screening identifies novel inhibitors of *Escherichia coli* dihydrofolate reductase that are competitive with dihydrofolate. Bioorg Med Chem Lett;13:2493–2496.
10. Then RL (2004) Antimicrobial dihydrofolate reductase inhibitors – achievements and future options. J Chemother;16:3–12.
11. Summerfield RL, Daigle DM, Mayer S, Mallik D, Hughes DW, Jackson SG, Sulek M, Organ MG, Brown ED, Junop MS (2006) A 2.13 Å structure of *E. coli* dihydrofolate reductase bound to a novel competitive inhibitor reveals a new binding surface involving the M20 loop region. J Med Chem;49:6977–6986.
12. Kwon YK, Lu W, Melamud E, Khanam N, Bognar A, Rabinowitz JD (2008) A domino effect in antifolate drug action in *Escherichia coli*. Nat Chem Biol;4:602–608.
13. Carreras CW, Santi DV (1995) The catalytic mechanism and structure of thymidylate synthase. Annu Rev Biochem;64:721–762.
14. Finland M, Kass EH (1973) Trimethoprim-sulfamethoxazole – summary and comments on conference. J Infect Dis;128:S792–S816.
15. Koetzle TF, Williams GJB (1976) The crystal and molecular structure of the antifolate drug trimethoprim (2,4-diamino-5-(3,4,5-trimethoxybenzyl) pyrimidine). A neutron diffraction study. J Am Chem Soc;98:2074–2078.
16. Bowden K, Harris NV, Watson CA (1993) Structure-activity-relationships of dihydrofolate-reductase inhibitors. J Chemother;5:377–388.
17. Zinner SH (2007) Antibiotic use: present and future. New Microbiol;30:321–325.
18. Selassie CD, Li RL, Poe M, Hansch C (1991) Optimization of hydrophobic and hydrophilic substituent interactions of 2,4-diamino-5-(substituted-benzyl)pyrimidines with dihydrofolate reductase. J Med Chem;34:46–54.
19. Champness JN, Stammers DK, Beddell CR (1986) Crystallographic investigation of the cooperative interaction between trimethoprim, reduced cofactor and dihydrofolate-reductase. FEBS Letts;199:61–67.
20. Matthews DA, Bolin JT, Burridge JM, Filman DJ, Volz KW, Kaufman BT, Beddell CR, Champness JN, Stammers DK, Kraut J (1985) Refined crystal-structures of *Escherichia-coli* and chicken liver dihydrofolate-reductase containing bound trimethoprim. J Biol Chem;260:381–391.
21. Li R, Poe M (1988) Quantitative structure-activity relationships for the inhibition of *Escherichia coli* dihydrofolate reductase by 5-(substituted benzyl)-2,4-diaminopyrimidines. J Med Chem;31: 366–370.
22. King RD, Muggleton S, Lewis RA, Sternberg MJ (1992) Drug design by machine learning - the use of inductive logic programming to model the structure-activity-relationships of trimethoprim analogs binding to dihydrofolate-reductase. Pro Natl Acad Sci USA;89:1322–11326.
23. So SS, Richards WG (1992) Application of neural networks: quantitative structure-activity relationships of the derivatives of 2,4-diamino-5-(substituted-benzyl)pyrimidines as DHFR inhibitors. J Med Chem;35:3201–3207.
24. Hirst JD, King RD, Sternberg MJ (1994) Quantitative structure-activity-relationships by neural networks and inductive logic programming .1. The inhibition of dihydrofolate-reductase by pyrimidines. J Comput Aided Mol Des;8:405–420.
25. King RD, Hirst JD, Sternberg MJE (1995) Comparison of artificial-intelligence methods for modeling pharmaceutical QSARS. Appl Artif Intell;9:213–233.
26. Czaplinski KH, Hasel W, Wiese M, Seydel JK (1995) New benzylpyrimidines: inhibition of DHFR from various species. QSAR, comfa and PC analysis. Eur J Med Chem;30:779–787.
27. Hirst JD (1996) Nonlinear quantitative structure-activity relationship for the inhibition of dihydrofolate reductase by pyrimidines. J Med Chem;39:3526–3532.
28. King RD, Srinivasan A (1997) The discovery of indicator variables for QSAR using inductive logic programming. J Comput Aided Mol Des;11:571–580.
29. Dunn WJ III, Hopfinger AJ (1998) 3D QSAR of flexible molecules using tensor representation. Perspect Drug Discovery Des;12: 167–182.
30. Loukas YL (2001) Adaptive neuro-fuzzy inference system: an instant and architecture-free predictor for improved QSAR studies. J Med Chem;44:2772–2783.
31. Chiu TL, So SS (2004) Development of neural network QSPR models for Hansch substituent constants. 1. Method and validations. J Chem Inf Comput Sci;44:147–153.
32. Shine Y, Kikuchi T (2009) Estimation of relative binding free energy based on a free energy variational principle for quantitative structure activity relationship analyses. Chem Phys;365:53–59.
33. Cramer RD III, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc;110:5959–5967.
34. Cramer RD III, Patterson DE, Bunce JD (1988) Cross-validation, bootstrapping, and partial least-squares compared with multiple-regression in conventional qsar studies. Quant Struct Act Relat;7:18–25.
35. Klebe G, Abraham U, Mietzner T (1994) Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem;37: 4130–4146.
36. Tripos Inc. (2009) Sybyl Molecular Modeling Software. MO, USA.
37. Gasteiger J, Marsili M (1980) Iterative partial equalization of orbital electronegativity – a rapid access to atomic charges. Tetrahedron;36:3219–3228.
38. Wold S., Johansson A., Cochi M., (1993) 3D QSAR in drug design. Vol. 1: theory, methods and applications. Kubinyi H (Ed.), ESCOM science publishers, Lieden; pp. 523–550.

Supporting Information

Additional Supporting Information can be found in the online version of this article:

Table S1. Number of components and the corresponding cross-validated r^2 value of CoMFA and CoMSIA models.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.