Brief Report

3D QSAR investigations on locomotor activity of 5-cyano-*N*1,6-disubstituted 2-thiouracil derivatives

Bhanudas S. Kuchekar[1], Yogesh V. Pore[2,*]

[1] *Department of Pharmaceutical Chemistry, MAEER's Maharashtra Institute of Pharmacy, MIT Campus, Pune, Maharashtra, India;*
[2] *Department of Pharmaceutical Chemistry, Government College of Pharmacy, Karad, Maharashtra, India.*

ABSTRACT: Three dimensional quantitative structure activity relationship (3D QSAR) investigations were carried out on a series of 5-cyano-*N*1,6-disubstituted 2-thiouracil derivatives for their locomotor activity. The structures of all compounds were built on a workspace of VlifeMDS3.5 molecular modeling software and 3D QSAR models were generated by applying a partial least square (PLS) linear regression analysis coupled with a stepwise variable selection method. Both derived models were found to be statistically significant in terms of regression and internal and external predictive ability (r^2 = 0.9414 and 0.8511, q^2 = 0.8582 and 0.6222, pred_r^2 = 0.5142 and 0.7917). The QSAR models indicated that both electrostatic and steric interaction energies were contributing significantly to locomotor activity of thiouracil derivatives.

Keywords: 3D QSAR, thiouracil derivatives, locomotor activity, predictive ability, PLS

1. Introduction

It is well known that pyrimidine compounds are associated with a large number of biological activities like antimicrobial (*1*,*2*), anticancer (*3*), antiviral (*4*), antioxidant (*5*) and CNS activities (*6*). There has always been growing interest in design and development of pyrimidine derivatives. In recent years computational chemistry, especially quantitative structure activity relationship studies (QSAR) have become an integral part of drug discovery processes. Significant attention has been focused on QSAR investigations of a variety of heterocyclic compounds and pyrimidines are not exceptional cases (*7-11*).

Thus, recognizing the biological significance of pyrimidines, some novel 5-cyano-*N*1,6-disubstituted 2-thiouracil derivatives were synthesized and evaluated for locomotor activity using actophotometer. QSAR relationships of these compounds were also further established with locomotor activity.

The objective of the present work was to perform three dimensional quantitative structure activity relationship (3D QSAR) studies on 5-cyano-*N*1,6-disubstituted 2-thiouracil derivatives. Partial least square (PLS) linear regression analysis was used to derive various QSAR models. The models were further validated for their regression coefficient, internal and external predictive ability and statistical significance. The models were interpreted to investigate the contribution of various 3D descriptors in locomotor activity.

2. Materials and Methods

2.1. *Molecular modeling software*

VLifeMDS (Version 3.5 VLife Sciences Technologies Pvt. Ltd., Pune, India) molecular modeling software was used for construction of molecules and generation of 3D QSAR models.

2.2. *Synthesis and biological evaluation of thiouracil derivatives (P1-P24)*

The thiouracil derivatives were synthesized and evaluated for locomotor activity by actophotometer according to a reported procedure (*12*).

2.3. *Biological data*

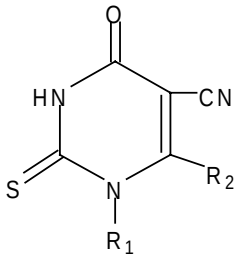
A data set of pEC_{50} values (locomotor activity) of twenty four compounds was used for 3D QSAR investigations (Table 1). The negative logarithmic values of micromolar concentrations of compounds required to produce a fifty percent response (EC_{50}) in animals were used as dependent variables.

2.4. *Energy minimization and alignment of molecules*

The molecules were optimized for energy minimization

**Address correspondence to:*
Dr. Yogesh Pore, Department of Pharmaceutical Chemistry, Government College of Pharmacy, Karad, Maharashtra 415124, India.
e-mail: dryogeshpore@rediffmail.com

Table 1. Nature of R group and pEC_{50} values of thiouracil derivatives



Compound code	Nature of R_1	Nature of R_2	pEC_{50} (– log EC_{50} or log 1/EC_{50}) (μmol)
P1	C_6H_5	C_6H_5	0.5534
P2	C_6H_5	4-OH-C_6H_5	0.7598
P3	C_6H_5	4-OCH_3-C_6H_5	0.8260
P4	4-Cl-C_6H_5	C_6H_5	0.5678
P5	4-Cl-C_6H_5	4-OH-C_6H_5	0.6683
P6	4-Cl-C_6H_5	4-OCH_3-C_6H_5	0.7735
P7	4-CH_3-C_6H_5	C_6H_5	0.7927
P8	4-CH_3-C_6H_5	4-OH-C_6H_5	0.7749
P9	4-CH_3-C_6H_5	4-OCH_3-C_6H_5	0.7761
P10	4-OCH_3-C_6H_5	C_6H_5	0.7958
P11	4-OCH_3-C_6H_5	4-OH-C_6H_5	0.7657
P12	4-OCH_3-C_6H_5	4-OCH_3-C_6H_5	0.8535
P13	4-F-C_6H_5	C_6H_5	0.5712
P14	4-F-C_6H_5	4-OH-C_6H_5	0.6928
P15	4-F-C_6H_5	4-OCH_3-C_6H_5	0.7814
P16	2, 4-CH_3-C_6H_5	C_6H_5	0.7975
P17	2, 4-CH_3-C_6H_5	4-OH-C_6H_5	0.7601
P18	2, 4-CH_3-C_6H_5	4-OCH_3-C_6H_5	0.8045
P19	3-CH_3-C_6H_5	C_6H_5	0.7629
P20	3-CH_3-C_6H_5	4-OH-C_6H_5	0.7610
P21	3-CH_3-C_6H_5	4-OCH_3-C_6H_5	0.7667
P22	4-NO_2-C_6H_5	C_6H_5	0.5581
P23	4-NO_2-C_6H_5	4-OH-C_6H_5	0.6638
P24	4-NO_2-C_6H_5	4-OCH_3-C_6H_5	0.7899

using MMFF (Merck Molecular Force Field) in the MOPAC module of VLifeMDS software. The threshold value for root mean square (rms) gradient was kept at 0.001 kcal/mol•Å. All molecules were subsequently aligned by a template based alignment technique using a common structure as a template. The most active molecule was selected as a template for alignment of the molecules. The alignment is useful for studying shape variation with respect to the base structure selected for alignment.

2.5. *Descriptor calculation*

For 3D QSAR analysis, the VLife Molecular Design Suite (VLifeMDS) allows the user to choose probe, grid size, and grid interval for the generation of descriptors. After suitable alignment of a given set of molecules, a common rectangular grid (lattice) was generated around the molecules. The steric, electrostatic and hydrophobic interaction energies were computed at the lattice points of the grid using a methyl probe of charge +1. These interaction energy values were considered for relationship generation and utilized as descriptors.

2.6. *Statistical analysis*

The descriptors were taken as independent variables and biological activity as dependent variables. The Partial Least Squares Regression (PLSR) method of analysis was used to derive the 3D QSAR equations. Statistical parameters employed were the number of compounds in regression *n*, the regression coefficient r^2, the *F*-test (Fischer's value) for statistical significance *F*, the cross-validated correlation coefficient q^2 and the standard error of estimation r^2 and q^2. The regression coefficient r^2 represents the part of variation in the observed data that is explained by the regression. Correlation coefficient values closer to 1.0 represent the better fit of the regression. The *F*-test is the ratio of the variance explained by the model and the variance due to error in the regression. High values of the *F*-test indicate that the model is statistically significant. The predictive ability (internal) of the generated models was evaluated by a cross validation method using a 'leave-one-out' scheme. Validation parameters considered were cross validated q^2. The predictive ability (external) of the selected model was also confirmed by external

validation of test set compounds which is denoted with pred_r^2.

3. Results

The QSAR investigations of 5-cyano-*N*1,6-disubstituted 2-thiouracil analogues resulted in several 3D QSAR equations. The two best equations are discussed.

3.1. *Validation of QSAR models*

Table 2 illustrates validation parameters for 3D QSAR models. Model 1 was obtained by a classical sphere exclusion type algorithm of training and test data selection (*13-16*) whereas model 2 was obtained by a random method of training and test set data selection (*17*). As indicated, both QSAR models were found to be statistically significant and predictive in terms of r^2, q^2, *F* and pred_r^2 values (*18-22*). Table 3 shows comparative predicted activities along with residuals of the two models. Figures 1 and 2 represent the fitness plots (R^2) of observed *vs.* predicted biological activity for models 1 and 2, respectively.

Model 1 has shown excellent fit (r^2 = 0.9414), good internal predictive ability (q^2 = 0.8542) and a good fitness plot (R^2 = 0.8519) with an optimal ability

Table 2. Comparative data of validation parameters employed for 3D QSAR equations

Model/Parameters	Model 1	Model 2
Equation	pEC_{50} = + 0.4431 E_86 – 187.1126 S_753 + 0.0044 S_284 + 0.0104 S_606 – 7.6961	pEC_{50} = – 104.9810 S_753 + 0.6250 E_86 + 0.0031 S_284 + 0.0161 E_524 + 0.0312 S_434 – 3.8624
Training set size (n)	19	18
Test set size	5 (P1, P20, P2, P9, and P15)	6 (P6, P8, P10, P15, P22, and P24)
Degree of freedom	15	14
r^2	0.9414	0.8511
r^2 se	0.0238	0.0385
F test	80.3745	26.6740
q^2	0.8582	0.6222
q^2 se	0.0370	0.0613
pred_r^2	0.5142	0.7917
pred_r^2 se	0.0681	0.0428
R^2 for fitness plot	0.8519	0.8365

Table 3. Comparative observed and predicted activities (LOO) of thiouracil derivatives by 3D QSAR models

Compounds	Observed activity pEC_{50}	Predicted activity pEC_{50}[a]			
		3D Model 1	Residuals	3D Model 2	Residuals
P1	0.5534	0.6158[b]	– 0.0624	0.6196	– 0.0662
P2	0.7598	0.6939[b]	0.0659	0.7285	0.0313
P3	0.8260	0.8256	0.0004	0.8137	0.0123
P4	0.5678	0.5826	– 0.0148	0.5758	– 0.008
P5	0.6683	0.6617	0.0066	0.6947	– 0.0264
P6	0.7735	0.7303	0.0432	0.7153[b]	0.0582
P7	0.7927	0.7685	0.0242	0.7247	0.068
P8	0.7749	0.8124	– 0.0375	0.8114[b]	– 0.0365
P9	0.7761	0.8509[b]	– 0.0748	0.8304	– 0.0543
P10	0.7958	0.7754	0.0204	0.8037[b]	– 0.0079
P11	0.7657	0.7775	– 0.0118	0.8096	– 0.0439
P12	0.8535	0.8585	– 0.005	0.8366	0.0169
P13	0.5712	0.5730	– 0.0018	0.5521	0.0191
P14	0.6928	0.6531	0.0397	0.6786	0.0142
P15	0.7814	0.7782[b]	0.0032	0.7415[b]	0.0399
P16	0.7975	0.7991	– 0.0016	0.7994	– 0.0019
P17	0.7601	0.7702	– 0.0101	0.7680	– 0.0079
P18	0.8045	0.7952	0.0093	0.7687	0.0358
P19	0.7629	0.7798	– 0.0169	0.7764	– 0.0135
P20	0.7610	0.6922[b]	0.0688	0.7234	0.0376
P21	0.7667	0.7802	– 0.0135	0.7552	0.0115
P22	0.5581	0.5895	– 0.0314	0.5574[b]	0.0007
P23	0.6638	0.6730	– 0.0092	0.6924	– 0.0286
P24	0.7899	0.7844	0.0055	0.7372[b]	0.0527

[a] Indicates predicted activity by leave one out cross validation; [b] Indicates molecules of test set.

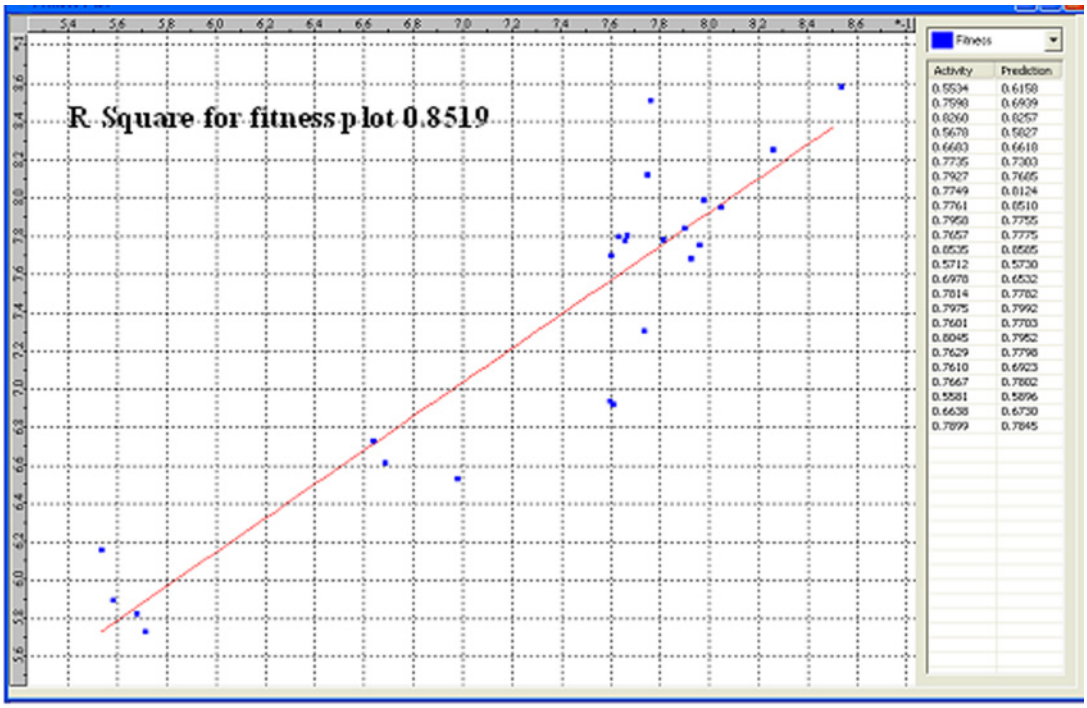


Figure 1. The plot of observed *versus* predicted activity for 3D model 1.

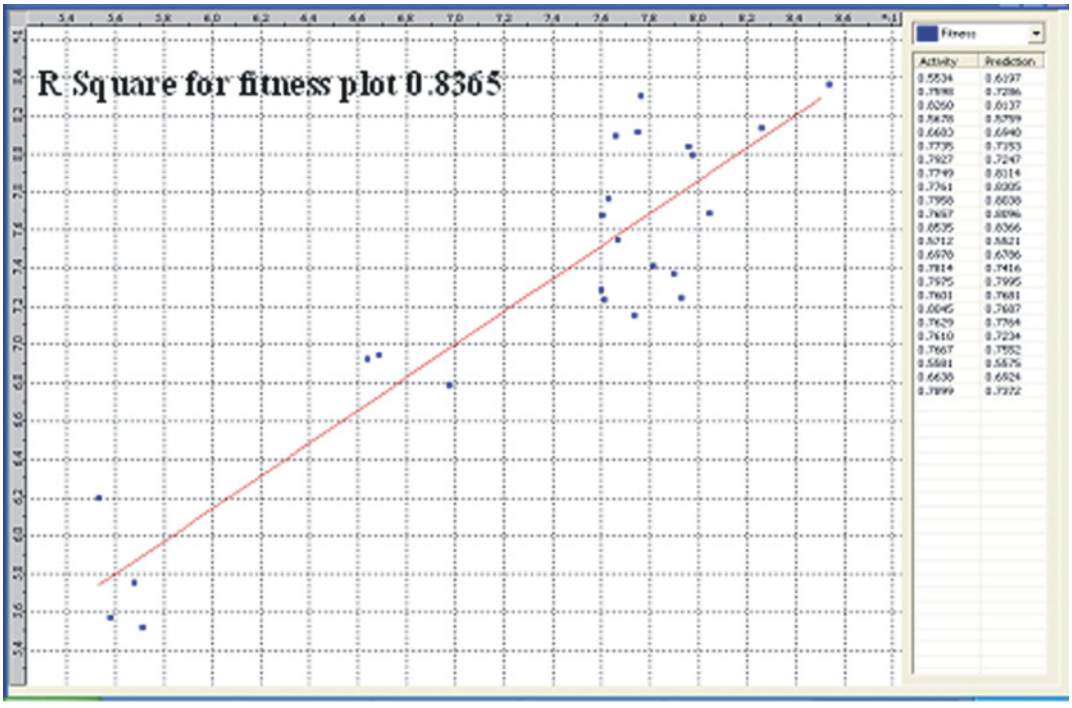


Figure 2. The plot of observed *versus* predicted activity for 3D model 2.

to predict the activities of test set molecules (pred_r^2 = 0.5182) which have not been included to build the QSAR model.

Model 2 has also displayed a good correlation coefficient (r^2 = 0.8511), optimal internal predictive ability (q^2 = 0.6222) with a fitness plot of (R^2 = 0.8109). However, the ability of model 2 to predict the activities of test set molecules is certainly excellent (pred_r^2 = 0.7917) as compared to model 1.

From the equations, it could be concluded that 94.14% (r^2 = 0.9414) and 85.11% (r^2 = 0.8511) of the variation in the biological activity was accounted for by the parameters used in equations 1 and 2, respectively. This signifies that in both the models, a good correlation exists between their corresponding descriptors and biological activity (*23*). Further, in

both cases the high values of *F* tests indicated that the statistical significance of 99.99% of the models meant that probability of failure of the models was 1 in 10,000.

3.2. *Interpretation of QSAR models*

The local fields around aligned molecules of those found to be important for activity variation in model 1 and model 2 are shown in Figures 3 and 4, respectively.

As shown in the figures, two electrostatic field descriptors E_86 and E_524 (blue points) and four steric field descriptors S_753, S_284, S_606, and S_434 (green points) are contributing significantly for locomotor activity of the molecules. The electrostatic

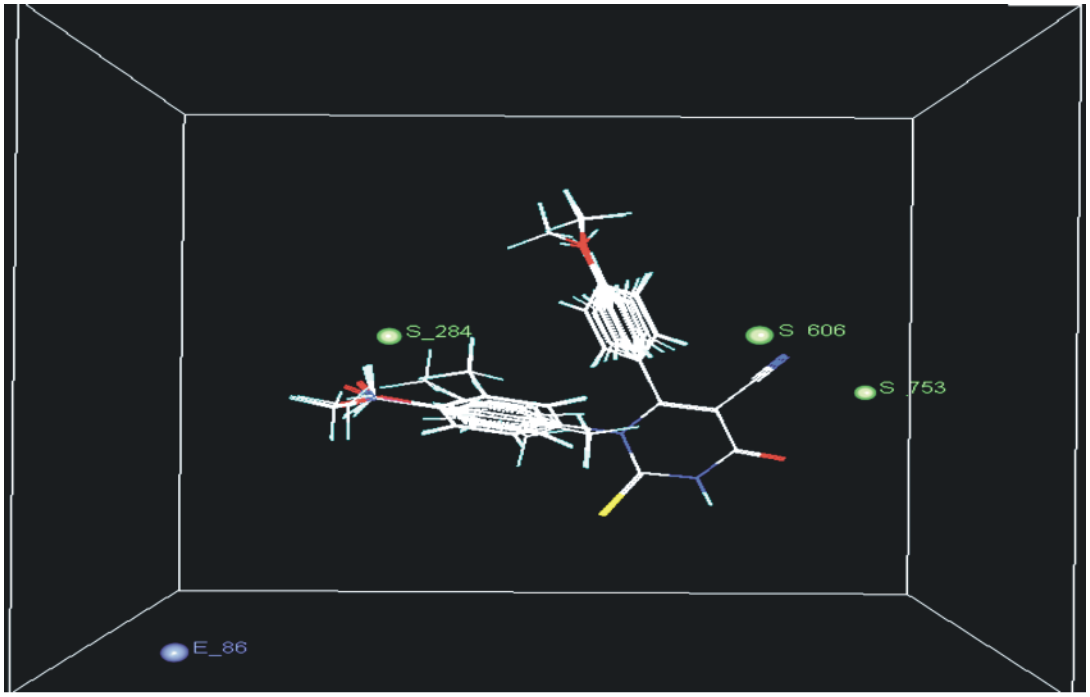


Figure 3. Relative positions of the local fields around aligned molecules for 3D model 1.

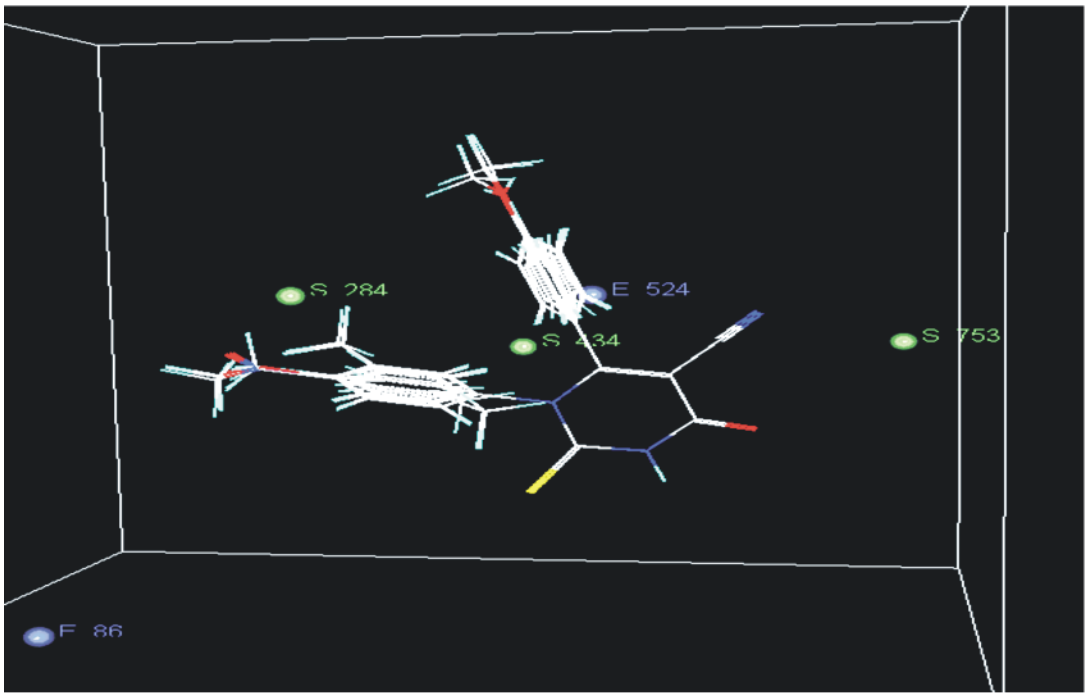


Figure 4. Relative positions of the local fields around aligned molecules for 3D model 2.

field descriptor E_86 and steric field descriptors S_753, and S_284 are common in both models. The positive range for E_86 and E_524 indicate that positive electrostatic potential is favorable for increase in activity and hence a less electronegative substituent should be preferred in these regions.

The steric field descriptors S_284, S_606, and S_434 contribute positively which indicates that positive steric potential is favorable for an increase in activity and hence a more bulky substituent is preferred in these regions. However, a negative contribution of S_753 indicates that the bulky substituent can not be tolerated and steric interactions should be reduced in that region for optimal biological activity.

Thus these relative positions and ranges of the corresponding important electrostatic and steric fields in the above models could be helpful in design of new molecules with improved locomotor activity.

4. Conclusions

The present QSAR investigations demonstrated that the generated 3D QSAR models were statistically significant in terms of their correlation with biological activity and internal and external predictive abilities. It could be concluded that the locomotor activity of thiouracil derivatives was positively contributed by electrostatic interaction fields. In addition the increase in steric interactions (bulky aromatic or cycloaliphatic substituents) at certain lattice points could be beneficial for improved potency of the thiouracil derivatives. Thus the developed 3D QSAR models may raise a scope for the design of new thiouracil derivatives with improved biological profiles (locomotor activity).

Acknowledgements

Authors are thankful to Principal, Bharati Vidyapeeth's College of Pharmacy, Kolhapur, Maharashtra, India, for providing software facilities for QSAR studies. Authors are very much thankful to Principal, Govt. College of Pharmacy, Karad, Maharashtra, India, for providing laboratory facilities and constant encouragement.

References

1. Sunduru N, Agarwal A, Katiyar SB, Nishi, Goyal N, Gupta S, Chauhan PM. Synthesis of 2,4,6-trisubstituted pyrimidine and triazine heterocycles as antileishmanial agents. Bioorg Med Chem. 2006; 14:7706-7715.
2. George S, Parameswaran MK, Chakraborty AR, Ravi TK. Synthesis and evaluation of the biological activities of some 3-{[5-(6-methyl-4-aryl-2-oxo-1,2,3,4-tetrahydropyrimidin-5-yl)-1,3,4-oxadiazol-2-yl]-imino}-1,3-dihydro-2H-indol-2-one derivatives. Acta Pharm. 2008; 58:119-129.
3. Gangjee A, Yu J, Kisliuk LR, Halie HW, Sobrero G, McGuire JJ. Design, synthesis and biological activities of classical N-4-[2-(2-amino-4-ethylpyrrolo[2,3-d]pyrimidin-5-yl)ethyl]benzoyl-L-glutamic acid and its 6-methyl derivative as potential dual inhibitors of thymidylate synthase and dihydrofolate reductase and as potential antitumor agents. J Med Chem. 2003; 46:591-600.
4. Rostom SF, Fahmy HY, Saudi MS. Synthesis and *in vitro* anti-HIV screening of certain 2-(benzoxazol-2-ylamino)-3H-4-oxopyrimidines. Sci Pharm. 2003; 71:57-74.
5. Gordon LB, Donald EA, Banitt SL, Kenneth LB, Stephen AM, Palmer JR, Tustin JM. Synthesis of novel 2,4-diamino pyrrolo[2,3-d]pyrimidines with antioxidant neuroprotective and antiasthama activity. J Med Chem. 1995; 38:4161-4163.
6. Matosiuk D, Tkaczynski T, Stefanczyk J. Synthesis and CNS activity of new 1-alkyl-2-aryl-7-hydroxy-5(1H)oxo-imidazo[1,2-a]pyrimidines. Acta Pol Pharm. 1996; 53:209-212.
7. Jarmuła A, Cieplak P, Krygowski TM, Rode W. The effect of 5-substitution in the pyrimidine ring of dUMP on the interaction with thymidylate synthase: Molecular modeling and QSAR. Bioorg Med Chem. 2007; 15:2346-2358.
8. Shrikant K, Debnath B, Jha T. QSAR study on adenosine kinase inhibition of pyrrolo[2,3-d]pyrimidine nucleoside analogues using the hansch approach. Bioorg Med Chem Lett. 2002; 12:899-902.
9. Chakraborti AK, Gopalakrishnan B, Sobhia ME, Malde A. 3D-QSAR studies on thieno[3,2-d]pyrimidines as phosphodiesterase IV inhibitors. Bioorg Med Chem Lett. 2003; 13:1403-1408.
10. Prabhakar YS, Ram VJ. A QSAR study on the antileishmanial activity of some substituted pyrimidines and pyrazolo [1,5-a] pyrimidines. Indian J Pharm Sci. 1997; 59:286-291.
11. Jain P, Soni LK, Gupta AK, Kashkedikar SG. QSAR analysis of 2,4-diaminopyrido[2,3-d]pyrimidines and 2,4-diaminopyrrolo[2,3-d]pyrimidines as dihydrofolate reductase inhibitors. Indian J Biochem Biophy. 2005; 42:315-320.
12. Pore Y, Kuchekar B, Bhatia M, Ingale K. Quantitative structure activity relationship (QSAR) studies on 5-cyano, *N*1,6-disubstituted, 2-thiouracil derivatives as central nervous system depressants. Digest Journal of Nanomaterials and Biostructures. 2009; 4:373-382.
13. Coi A, Fiamingo FL, Livi O, Calderone V, Martelli A, Massarelli I, Bianucci AM. QSAR studies on BK channel activators. Bioorg Med Chem. 2009; 17:319-325.
14. Golbraikh A, Tropsha A. Predictive QSAR modeling based on diversity sampling of experimental data sets for the test and training set selection. J Comp Aid Molec Des. 2002; 16:357-369.
15. Golbraikh A. Molecular data set diversity indices and their applications to comparison of chemical databases and QSAR analysis. J Chem Inf Comput Sci. 2000; 40:414-425.
16. Guha R, Serra JR, Jurs PC. Generation of QSAR sets with a self-organizing map. J Mol Graph Model. 2004; 23:1-14.
17. Chakraborti AK, Gopalakrishnan B, Sobhia ME, Malde A. 3D-QSAR studies of indole derivatives as phosphodiesterase IV inhibitors. Eur J Med Chem. 2003; 38:975-982.
18. Puratchikody A, Nagalakshmi G, Doble M. Experimental and QSAR studies on antimicrobial activity of

benzimidazole derivatives. Chem Pharm Bull. 2008; 56:273-281.
19. Tropsha A, Gramatica P, Gombar VK. The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. QSAR Comb Sci. 2003; 22:69-77.
20. Golbraikh A, Tropsha A. Beware of q^2! J Mol Graph Model. 2002; 20:269-276.
21. Afantitis A, Melagraki G, Sarimveis H, Igglessi-Markopoulou O, Kollias G. A novel QSAR model for predicting the inhibition of CXCR3 receptor by 4-*N*-aryl-[1,4] diazepane ureas. Eur J Med Chem. 2009; 44:877-884.
22. Ravichandran V, Mourya VK, Agrawal RK. QSAR prediction of HIV-1 reverse transcriptase inhibitory activity of benzoxazinone derivatives. Internet Electron J Mol Des. 2007; 6:363-374.
23. VLifeMDS3.5 Molecular Design Suite, Vlife Sciences Technologies Pvt. Ltd., Pune, India (*www.vlifesciences.com*) (2008).

(Received February 18, 2010; Accepted March 14, 2010)