International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:7, No:12, 2013

QSAR Studies of Certain Novel Heterocycles Derived from Bis-1, 2, 4 Triazoles as Anti-Tumor Agents

Madhusudan Purohit, Stephen Philip, Bharathkumar Inturi

***Abstract*—**In this paper we report the quantitative structure activity relationship of novel bis-triazole derivatives for predicting the activity profile. The full model encompassed a dataset of 46 Bis-triazoles. Tripos Sybyl X 2.0 program was used to conduct CoMSIA QSAR modeling. The Partial Least-Squares (PLS) analysis method was used to conduct statistical analysis and to derive a QSAR model based on the field values of CoMSIA descriptor. The compounds were divided into test and training set. The compounds were evaluated by various CoMSIA parameters to predict the best QSAR model. An optimum numbers of components were first determined separately by cross-validation regression for CoMSIA model, which were then applied in the final analysis. A series of parameters were used for the study and the best fit model was obtained using donor, partition coefficient and steric parameters. The CoMSIA models demonstrated good statistical results with regression coefficient (r^2) and the cross-validated coefficient (q^2) of 0.575 and 0.830 respectively. The standard error for the predicted model was 0.16322. In the CoMSIA model, the steric descriptors make a marginally larger contribution than the electrostatic descriptors. The finding that the steric descriptor is the largest contributor for the CoMSIA QSAR models is consistent with the observation that more than half of the binding site area is occupied by steric regions.

***Keywords*—**3D QSAR, CoMSIA, Triazoles.

I. Introduction

CANCER, the uncontrolled, rapid and pathological proliferation of abnormal cells, is one of the most formidable afflictions in the world [1]. Cancer continues to be a worldwide killer, despite the enormous amount of research and rapid developments during the past decade. According to research statistics [2], cancer accounts for about 23% of the total deaths in the USA and is the second most common cause of death after heart disease. Therefore, there is an increasing need for new therapies, especially of cancer biology as well as those taking advantage of the cancer cell phenotype, described by Hanahan and Weinberg [3].

Over the last few years the increasing number of neoplastic diseases together with the accompanied high mortality rates has stimulated an unparalleled level of research directed towards the development of new lead molecules that might be of use in designing novel anti-neoplastic agents.

M. Purohit is with Dept. of Pharmaceutical Chemistry, JSS College of Pharmacy, JSS University, Mysore- 570015, India (phone: 0091 9886698475; fax: 0821 2548359; e-mail: mnpurohit04@yahoo.com).

S. Philip and B. Inturi are with Dept. of Pharmaceutical Chemistry, JSS College of Pharmacy, JSS University, Mysore- 570015, India (e-mail: stephen3124@hotmail.co.uk, bharatinturi@gmail.com).

Triazoles are known to have a large spectrum of potential anticancer, antimitotic and antifungal properties [4]. With an improved understanding of the genes and pathways responsible for cancer initiation and progression, cancer drug development has undergone a paradigm change in the recent years, from predominantly cytotoxic agent based therapy to therapy aimed at molecular and genetic targets. The derivatives of 1, 2, 4 triazoles have high potential for biological activity. The following 1, 2, 4 triazoles derivatives are used in medicine: alprazolam (tranquiliser), estazolam (hypnotic), benatradin (diuretic) and trazodon (anti-depressant). The derivatives of 1, 2, 4 triazole possess a wide range of anti-microbial and anti-tumor properties. Recently, the compounds containing 1, 2, 4 triazoles were discovered as a novel class of potent tubulin polymerization inhibitors.

Quantitative structure–activity relationships (QSAR), an important area of chemoinformatics have been widely utilized to study the relationship between chemical structures and biological or other functional activities. QSAR has become increasingly helpful in understanding many aspects of chemical–biological interactions in drug and pesticide research as well as in many other areas [5]. The Schematic representation of the QSAR model is shown in Fig. 1.

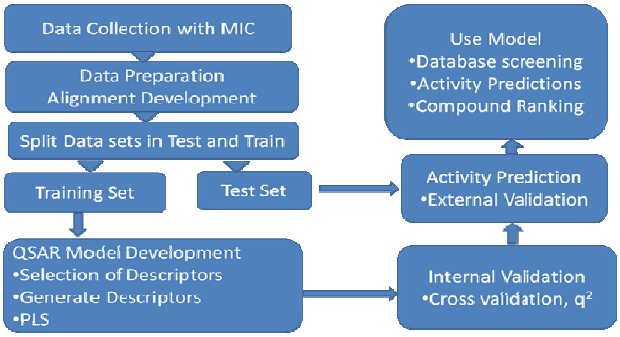


Fig. 1 Schematic Representation of 3D-QSAR model building and validation

II. Materials and Methods

A. Data Set

The dataset of MIC related to antitumor activity was collected from our earlier work, structures are shown in Table I [6], [7] and converted into pIC_{50} for convenience. CoMSIA (comparative molecular similarity index analysis) studies were

carried out using SYBYL X 2.1 (CERTARA, Portugal).

The generation of consistent statistical models depends on the proper selection of both training and test sets in terms of structural diversity and property values distribution. From the data 80% compounds were selected as members of the training set for model construction, and the other 20% compounds as members of the test set for external model validation. The compounds of the training and test sets were selected by based unity finger prints and dissimilarity, which is available in sybyl X.2.1. A graphical representation confirmed that the composition of both training and test sets is representative of the whole data set.

B. Molecular Alignment

Since the results of 3D-QSAR studies are sensitive to alignment of molecules, therefore, the alignment of 3D structures plays a vital role during CoMSIA analysis [8]. The lowest energy conformer of the most active molecule in the data set (**4b**) was chosen as template structure for the data set molecular alignment (Fig. 2). The molecules in their respective lowest energy conformations were superimposed on the template using the rigid-body fit option in SYBYL-X 2.1.

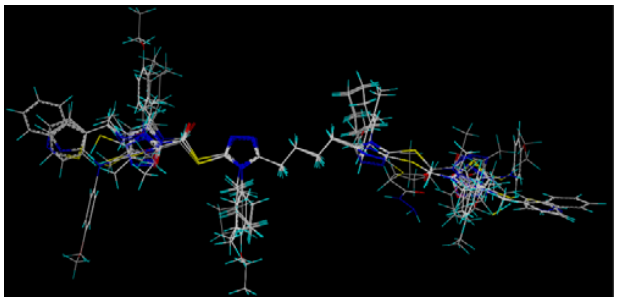


Fig. 2 3D-QSAR structure alignment and superposition of 46 compounds using compound 4b as the template

CoMSIA is a technique in which similarity indices are calculated at different points on a regularly spaced grid for pre-aligned molecules. In this approach, five different similarity fields are calculated: Steric, Electrostatic, Hydrophobic, Hydrogen bond Donor and Hydrogen bond Acceptor. These fields are selected to cover the major contributions to ligand binding. In CoMSIA fields, singularities were avoided at atomic positions because a Gaussian type distance dependence of each physicochemical property was adopted and thus no arbitrary cutoffs were necessary. The attenuation factor was set to the default value of 0.3. Cross-validated regression coefficient (q^2) values were calculated by using partial least-squares (PLS) methodology [11]-[13]. Leave-one-out (LOO) cross-validation was used to obtain optimum number of components (ONC). The final non-cross-validated model was developed with ONC to yield conventional regression coefficient (r^2) value, statistical significance value (F) and standard error of estimate (SEE).

C. Partial Least Square (PLS) Analysis and Validations

The standard CoMFA procedure, as implemented in SYBYLX.2.1, was performed. A 3D cubic lattice with a grid spacing of 2 Å was created automatically in all X, Y and Z directions by the program to encompass all the aligned ligands. A default sp^3 C probe atom with a Van-Der Waals radius of 1.52 Å and a charge of þ1.0 was used to generate steric field energies and Electrostatic (Coulombic potential) fields with a distance dependent dielectric at each lattice point. The computed field energies were truncated to 30 kcal/ mol for both Steric and electrostatic fields. CoMSIA is a technique in which similarity indices are calculated at different points on a regularly spaced grid for pre-aligned molecules. In this approach, five different similarity fields are calculated: Steric, Electrostatic, Hydrophobic, Hydrogen bond Donor and Hydrogen bond Acceptor. These fields are selected to cover the major contributions to ligand binding. In CoMSIA fields, singularities were avoided at atomic positions because a Gaussian type distance dependence of each physicochemical property was adopted and thus no arbitrary cutoffs were necessary. The attenuation factor was set to the default value of 0.3. Cross-validated regression coefficient (q^2) values were calculated by using partial least-squares (PLS) methodology [9]-[11]. Leave-one-out (LOO) cross-validation was used to obtain optimum number of components (ONC). The final non-cross-validated model was developed with ONC to yield conventional regression coefficient (r^2) value, F value, and S value (standard error of estimate).

$$q^2_{cv} = 1 - \frac{\sum (y_{obs} - y_{pred})^2}{\sum (y_{obs} - y_{mean})^2} \quad (1)$$

$$SEP = \sqrt{\sum \frac{(y_{obs} - y_{pred})^2}{N}} \quad (2)$$

where Y_{pred} = predicted value, Y_{obs} = experimental value, Y_{mean} = average value and N = number of objects (molecules)

D. CoMSIA Statistical Results

For the CoMSIA model, some possible combinations of different fields were performed to determine the best CoMSIA model. The highest cross-validated q^2 was obtained by using the combination of steric, H-bond donor and ClogP fields (q^2-0.575, r^2-0.830, SEE-0.163) with six components. The corresponding field contributions are 81.1%, 12.1%, and 6.9%, respectively. Residual values (the difference between predicted and actual values) were shown in Table II. The relationship between the predicted and the experimental pIC_{50} values for the CoMSIA model is depicted in Fig. 3. From the cross-validation results, it can be seen that the CoMSIA model has a good predictive ability, suggesting that a reliable CoMSIA model is successfully constructed.

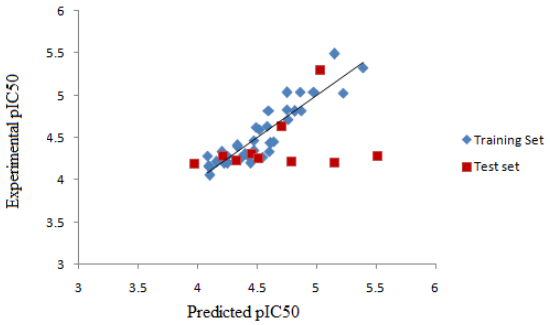
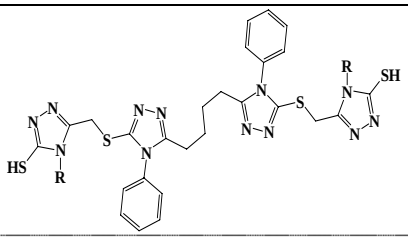
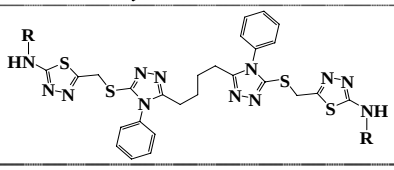
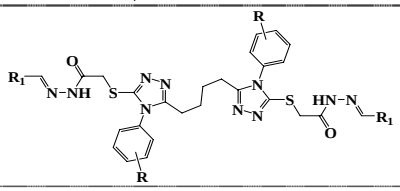
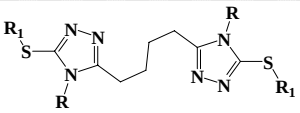
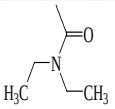
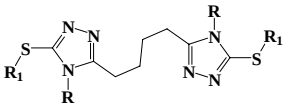
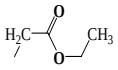
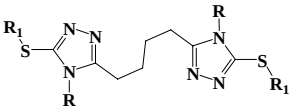
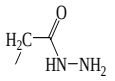
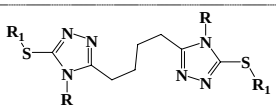
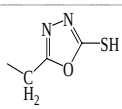


Fig. 3 Calculated pMIC Vs. experimental pMIC values for the 36 training set molecules obtained by PLS analysis using CoMSIA

TABLE I
MOLECULAR STRUCTURES TAKEN FOR THE STUDY

Cpd ID	COMPOUND	R	R_1
1a		Phenyl	-
1b		p- tolyl	-
1c		m-tolyl	-
1d		p-ethoxy phenyl	-
1e		Cyclohexyl	-
1f		n-butyl	-
1g		p-methoxy phenyl	-
2a		Phenyl	-
2b		p-tolyl	-
2c		p-methoxy phenyl	-
3a		-H	4-pyridyl
3b		-H	2-thienyl
3c		-H	3-indolyl
3d		-3-CH_3	4-pyridyl
3e		-3-CH_3	2-thienyl
3f		-3-CH_3	3-indolyl
4a		Phenyl	
4b		p-tolyl	
4c		m-tolyl	
4d		p-ethoxy phenyl	
4e		Cyclohexyl	
4f		n-butyl	
5a		Phenyl	
5b		p-tolyl	
5c		m-tolyl	
5d		p-ethoxy phenyl	
5e		Cyclohexyl	
5f		n-butyl	
6a		Phenyl	
6b		p-tolyl	
6c		m-tolyl	
6d		p-ethoxy phenyl	
6e		Cyclohexyl	
6f		n-butyl	
7a		Phenyl	
7b		p-tolyl	
7c		m-tolyl	
7d		p-ethoxy phenyl	

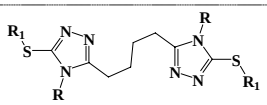
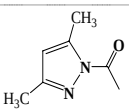
7e		Cyclohexyl	
7f		n-butyl	
8a		Phenyl	
8b		p-tolyl	
8c		m-tolyl	
8d		p-ethoxy phenyl	
8e		Cyclohexyl	
8f		n-butyl	

TABLE II
EXPERIMENTAL, PREDICTED AND RESIDUAL VALUES OF THE COMSIA MODEL

Cpd ID	IC_{50}	pIC_{50}	CoMSIA	Residual
		Training Set		
1b	53.37	4.2727	4.5469	0.2742
7e	69.94	4.1553	4.1049	-0.0504
1d	14.73	4.8318	4.7476	-0.0842
7f	52.3	4.2815	4.0831	-0.1984
1e	40.25	4.3952	4.3283	-0.0669
8a	35.25	4.4528	4.6346	0.1818
1f	25.19	4.5988	4.5192	-0.0796
1g	15.09	4.8213	4.5964	-0.2249
8c	9.23	5.0348	4.7468	-0.288
2a	61.26	4.2128	4.242	0.0292
8d	4.73	5.3251	5.3889	0.0638
2b	46.94	4.3285	4.2039	-0.1246
8e	63.37	4.1981	4.2506	0.0525
2c	88.36	4.0537	4.0998	0.0461
3a	51.81	4.2856	4.2456	-0.04
3b	23.88	4.622	4.4888	-0.1332
3d	56.19	4.2503	4.3577	0.1074
3e	45.81	4.339	4.601	0.262
3f	15.37	4.8133	4.8157	0.0024
4a	19.31	4.7142	4.7572	0.043
4b	3.23	5.4908	5.149	-0.3418
4c	15.25	4.8167	4.8694	0.0527
4e	51.81	4.2856	4.4219	0.1363
4f	63.06	4.2002	4.4418	0.2416
5a	34.47	4.4626	4.4703	0.0077
5b	9.13	5.0395	4.862	-0.1775
5c	23.27	4.6332	4.5824	-0.0508
5d	9.3	5.0315	5.2246	0.1931
5f	61.5	4.2111	4.1548	-0.0563
6a	63.16	4.1996	4.2216	0.022
6b	36.94	4.4325	4.6133	0.1808
6c	39.15	4.4073	4.3337	-0.0736
6d	9.13	5.0395	4.9759	-0.0636
6e	69.33	4.1591	4.0909	-0.0682
7a	49.91	4.3018	4.3985	0.0967
1a	45.25	4.3444	4.4732	0.1288
		Test Set		
7d	63.06	4.2002	5.1528	0.9526
1c	49.23	4.3078	4.4528	0.145
8b	5.09	5.2933	5.0264	-0.2669
8f	58.86	4.2302	4.3192	0.089
3c	23.27	4.6332	4.7036	0.0704
4d	51.87	4.2851	5.5115	1.2264
5e	51.87	4.2851	4.2091	-0.076
6e	65.26	4.1854	3.9675	-0.2179
7b	61.23	4.213	4.7902	0.5772
7c	55.25	4.2577	4.5106	0.2529

III. DISCUSSION

To view the field effect on the target property, CoMSIA contour maps were generated for the best predictive models. The contour maps can indicate the important regions in 3D space around the molecules where any change in the steric field may affect the biological activity. The field energies of all fields were calculated with the weight of "stdev*coeff" (the standard deviation and the coefficient). The contour maps obtained from the CoMSIA are illustrated together with template ligand 4b. The steric contour map of CoMSIA model is displayed in Fig. 4. The green (sterically favorable) and yellow (sterically unfavorable) contours represent 80% and 20% contributions. The green contours characterized by tolyl groups are favorable for bulky steric modifications where as diethyl carbamoyl groups are unfavorable for the steric modifications.

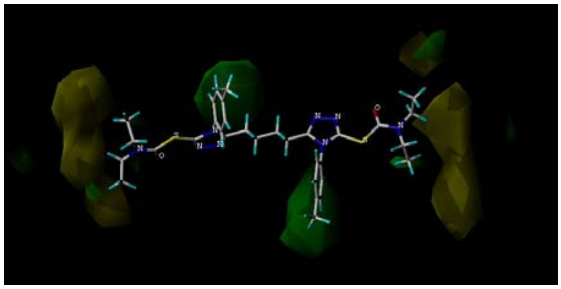


Fig. 4 CoMSIA contour maps of compound **4b** Steric contours: Green contours indicate regions where steric interaction is favored. Yellow contours are areas where the steric interactions unfavored

IV. CONCLUSION

The 3D QSAR study was carried out for a library of 46 molecules comprising of bis- 1, 2, 4 triazoles with different substitutions such as carbethoxy, hydrazine, diethyl carbamoyl etc.

CoMSIA (q^2-0.575, r^2-0.830, S-0.163) models gave good statistical results in terms of q^2 and r^2 values when studied along with donor, partition coefficient and steric parameters, and provided significant insights that could be used in further design of novel, potent and selective antitumor agents. Studies on the mechanism of action of these compounds are in progress and will be reported in the future.

ACKNOWLEDGMENTS

Authors are thankful to The Principal, JSS College of Pharmacy, Mysore, India for providing necessary facilities. Authors are beholden to DBT for the financial assistance (Ref.

No: BT/PR5594/MED/29/540/20l2).

REFERENCES

[1] N. A. Bruce and L. S. Gold, "Endogenous Mutagens and the Causes of Aging and Cancer," *Mutat. Res.* vol. 250, pp. 3-16, 1991.
[2] P. Anand, A. B. Kunnumakara, C. Sundaram, K. B. Harikumar, S. T. Tharakan, O. S. Lai, B. Sung and B. B. Aggarwal, "Cancer is a preventable disease that requires major lifestyle changes" *Pharm. Res*. vol. 25, 2008 pp. 2097-2116.
[3] A. Jemal, R. Siegel , E. Ward , T. Murray , J. Xu and M. J. Thun, "Cancer statistics," *C. A. Cancer J. Clin*, Vol. 57, Jan. 2007, pp. 43-66.
[4] D. Hanahan and R. A. Weinberg "The hallmarks of cancer," *Cell*, vol. 100, pp. 57-70.
[5] Yashwant, B. K. Sharma, B. Srivastava and S. Vandana, "Recent advancements of triazoles as anticancer agents," *IJPCR*, vol. 2, 2010, pp. 95-97.
[6] Madhusudan Purohit and Y. C. Mayur, "Synthesis, *in vitro* cytotoxicity, and anti-microbial studies of 1, 4-bis (4-substituted-5-mercapto-1, 2, 4-triazol-3-yl) butanes," *Medicinal Chemistry Research*, Vol. 21, 2012, pp 174-184.
[7] Madhusudan Purohit, V. V. S. Rajendra Prasad and Y. C. Mayur, "Synthesis and Cytotoxicity of Bis-1, 3, 4-Oxadiazoles and Bis-Pyrazoles derived from 1, 4-Bis [5-Thio-4-Substituted-1, 2, 4-Triazol-3-Yl]-Butane and their DNA Binding," *Archiv der Pharmazie - Chemistry in Life Scienc*,. Vol. 11, 2011, pp. 248-254.
[8] Madhusudan Purohit, Chandrashekhar Venkaraddi Mangannavar and Y. C. Mayur, "Synthesis and in vitro Cytotoxicity Studies of Certain Novel Heterocycles Derived from Bis 1, 2, 4-Triazoles and their DNA Damage Studies," *Med Chem*, Vol. 9, 2013, pp. 1063-1072.
[9] S. Marcelo and Castilho, "Two- and three-dimensional quantitative structure–activity relationships for a series of purine nucleoside phosphorylase inhibitors," *Bioorg. Med. Chem.* Vol. 14, 2006, pp. 516-527.
[10] P. V. and C. Hansch. "An approach toward the problem of outliers in QSAR. Bioorg," *Med. Chem.* Vol. 13, 2005, pp. 4597-4621.
[11] D. Richard and Cramer III, "Cross validation, Bootstrapping, and Partial Least Squares Compared with Multiple Regression in Conventional QSAR Studies," Quantitative Structure-Activity Relationships, Vol. 7, 1988, pp. 18-25.
[12] S. Wold, "The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses," *SIAM J. Sci. and Stat. Comput*, Vol. 5, 1984, pp. 735–743.
[13] Partha Pratim Roy, "On Some Aspects of Variable Selection for Partial Least Squares Regression Models,"*QSAR & Combinatorial Science*, Vol. 27, 2008, pp. 302-313.