



Scholars Research Library

J. Comput. Method. Mol. Design, 2011, 1 (1): 52-58
(<http://scholarsresearchlibrary.com/archive.html>)



3D-QSAR of 2,5-disubstituted-1,3,4-thiadiazole derivatives as diuretic agents: A comparative molecular field analysis study

Sanmati K. Jain^{*1} and Pradeep Mishra²

^{*1}SLT Institute of Pharmaceutical Sciences, Guru Ghasidas Vishwavidyalaya, Bilaspur, India
²GLA Institute of Pharmaceutical Research (GLAIPR), Mathura, India

ABSTRACT

A three-dimensional quantitative structure activity relationship (3D-QSAR) study using comparative molecular field analysis (CoMFA) method was performed on 2,5-disubstituted-1,3,4-thiadiazole derivatives as diuretic agents. This study was performed using 40 compounds, in which the CoMFA model was developed using a training set of 30 compounds. Ten compounds (selected at randomly served as a test set), which were not used in model generation, were used to validate the CoMFA model. CoMFA derived QSAR model shows a good conventional squared correlation coefficient r^2 and cross validated correlation coefficient r^2_{cv} 0.976 and 0.579 respectively. In this analysis steric and electrostatic field contribute to the QSAR equation by 73.8% and 26.2% respectively, suggesting that variation in biological activity of the compounds is dominated by differences in steric (van der Waals) interactions. To visualize the CoMFA steric and electrostatic field from PLS analysis, contour maps are plotted as percentage contribution to the QSAR equation and are associated with the differences in biological activity.

Keywords: 3D-QSAR, CoMFA, 2,5-Disubstituted-1,3,4-thiadiazole derivatives, Diuretic agents.

INTRODUCTION

Comparative Molecular Field Analysis (CoMFA) is a three-dimensional quantitative structure activity relationship (3D-QSAR) approach, introduced in 1988 by Cramer [1,2]. It developed slowly. From the very first formulation of a lattice model to compare molecules by aligning them with a putative pharmacophore and by mapping their surrounding fields to a three-dimensional grid, CoMFA approach was an application of the dynamic lattice oriented molecular modeling system (DYLOMMS), as it was called till 1987. A real advance resulted in 1987, the method was still named DYLOMMS but now it used grids including several thousands of points, partial least

squares (PLS) analysis and most important, a cross-validation procedure to check the predictive ability of different models. CoMFA is by far the most often employed receptor- independent (RI) 3D-QSAR approach, reflecting a novel, conceptually satisfying scientific approach reduced to practice as a well-written and versatile software package. In this method a relationship is established between the biological activities of a set of compounds and their steric and electrostatic properties [3-6]. For establishing relationship between structure and biological activities of the synthesized compounds [7-9] quantitatively, three-dimensional quantitative structure activity relationship (CoMFA) study was carried out.

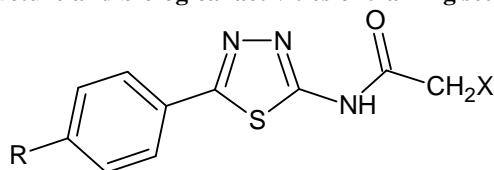
MATERIALS AND METHODS

Experimental

Data Set: A dataset of 40 molecules synthesized [7-9] earlier (2,5-disubstituted-1,3,4-thiadiazole derivatives) having diuretic activity has been taken for the present study (Table-1). Selected data set, their biological activity are shown in Table-1 and 2 forming the training and test set respectively. For CoMFA study, logarithmic value of biological activity (log BA) was taken, while BA is calculated using the following formulae [10]. BA is expressed percent urine excretion per micro mole of drug per kilo gram of body weight.

$$BA = \% \text{ urine excretion} \times \text{Mol. Wt.} / \text{dose (g)} \times 10^6$$

Table-1: Structure and biological activities of training set molecules (30)



Compound No.	R	X	AA*	Mol. Wt.	BA**	log BA
1	H	<i>n</i> -Butyl methyl amino	52.8	304.41	0.1607	-0.79
2	H	Di-(<i>n</i> -butyl) amino	85.2	346.49	0.2952	-0.53
3	H	Di- <i>iso</i> -butyl amino	79.4	346.49	0.2751	-0.56
4	H	Morpholino	75.8	304.37	0.2307	-0.64
5	H	4-Methyl piperidino	55.8	316.42	0.1766	-0.75
6	H	Piperidino	82.8	302.42	0.2504	-0.60
7	H	N- Methyl piperazino	49.9	317.41	0.1584	-0.80
8	H	Dicyclohexyl amino	101.6	398.57	0.4049	-0.39
9	CH ₃ O-	Di-(<i>n</i> -propyl) amino	46.6	348.46	0.1624	-0.79
10	CH ₃ O	<i>n</i> -Butyl methyl amino	20.55	334.44	0.0687	-1.16
11	CH ₃ O	Di-(<i>n</i> -butyl) amino	80.46	376.52	0.3030	-0.52
12	CH ₃ O	Di- <i>iso</i> -butyl amino	85.9	376.52	0.3234	-0.49
13	CH ₃ O	Morpholino	70.47	334.39	0.2441	-0.61
14	CH ₃ O	4-Methyl piperidino	55.57	346.45	0.1925	-0.72
15	CH ₃ O	Piperidino	75.17	332.42	0.2498	-0.60
16	CH ₃ O	N- Methyl piperazino	44.63	347.44	0.1551	-0.81
17	CH ₃ O	Pyrrolidino	50.8	318.39	0.1617	-0.79
18	CH ₃ O	Dicyclohexyl amino	79.28	428.59	0.3398	-0.47
19	CH ₃ O	2-Pyrrrolidinono	70.8	332.38	0.2353	-0.63
20	CH ₃	Di-(<i>n</i> -propyl) amino	37.3	332.46	0.1240	-0.91
21	CH ₃	Di- <i>iso</i> -butyl amino	83.9	360.52	0.3025	-0.52

22	CH ₃	4-Methyl piperidino	54.9	330.45	0.1814	-0.74
23	CH ₃	Piperidino	71.5	316.42	0.2262	-0.65
24	CH ₃	N- Methyl piperazino	42.4	331.44	0.1405	-0.85
25	CH ₃	2-Pyrrolidinono	65.3	316.38	0.2066	-0.68
26	Cl	Di-(<i>n</i> -propyl) amino	21.6	352.88	0.076	-1.12
27	Cl	Di- <i>iso</i> -butyl amino	56.0	380.94	0.2133	-0.67
28	Cl	Morpholino	42.6	338.81	0.1443	-0.84
29	Cl	Piperidino	55.9	336.84	0.1883	-0.73
30	Cl	2-Pyrrolidinono	41.3	336.80	0.1391	-0.86

* = Percent urine excretion at 100 mg/kg body weight orally; ** = Percent urine excretion per micromole of drug per kilogram of body weight.

Table-2: Structure and biological activities of test set molecules (10)

Compound No	R	X	AA	Mol. Wt.	BA	log(BA)
1	H	Di- <i>iso</i> -propyl amino	91.0	318.44	0.2897	-0.54
2	H	Pyrrolidino	50.1	288.37	0.1445	-0.84
3	CH ₃ O-	Di- <i>iso</i> -propyl amino	49.92	348.46	0.1740	-0.76
4	CH ₃	Di-(<i>n</i> -butyl) amino	75.9	360.52	0.5628	-0.56
5	CH ₃	Morpholino	67.2	318.39	0.2139	-0.67
6	CH ₃	Dicyclohexyl amino	70.5	412.59	0.2988	-0.54
7	Cl	Di-(<i>n</i> -butyl) amino	45.7	380.94	0.1740	-0.76
8	Cl	N- Methyl piperazino	36.7	351.85	0.1291	-0.89
9	Cl	Pyrrolidino	50.1	322.81	0.1617	-0.79
10	Cl	Dicyclohexyl amino	60.5	433.01	0.2619	-0.58

* = Percent urine excretion at 100 mg/kg body weight orally.

** = Percent urine excretion per micromole of drug per kilogram of body weight.

Molecular Modeling

Molecular Modeling and CoMFA studies were performed on Silicon Graphics Octane computer using molecular modeling package SYBYL 6.5 using the standard TRIPOS force field. Structural manipulations were performed with molecular modeling package SYBYL 6.5 using the standard TRIPOS force field. Partial atomic charges of ligands were calculated using within MOPAC. The structures were then optimized by energy minimization using the Powell algorithm to a final root mean square gradient of 0.05 kcal / mol.

Alignment

The alignment, i.e. molecular conformation and orientation, is one of the sensitive inputs for CoMFA. One of the most active compounds used as a reference compound. The compounds were fitted to the active analogue compound.

GRID Size

Once the molecules are aligned a grid or lattice is established which surrounds the set of analogs in potential receptor space. Current CoMFA studies seldom use grid resolution less than 1 Å and, most often, 2 Å. The choice of grid resolution represents a compromise between computational practicality and detailing of the fields. If the grid resolution is too small, the number of field-points (cells) becomes too large to perform a timely analysis. Moreover spatial information on field preference can be lost, through a 'smearing out' effect, if the cells become

too small. The grid resolution in the 1 to 2 Å° range corresponds to, at best, differentiating single carbon-carbon (1.54 Å°) from one another.

CoMFA Interaction Energy

The steric and electrostatic (potential fields) energies were calculated at each lattice intersection of a regularly spaced grid box. The lattice spacing was set a value of 2.0 Å°. CoMFA region was defined automatically which extends the lattice walls beyond the dimensions of each structures by 4.0 Å° in all directions. The Lennard-Jones Potential and coulombic term which represent, respectively steric and electrostatic fields, were calculated using the TRIPOS force fields.

An sp³ carbon atom with a van der Waals radius of 1.52 Å° and a +1.0 charge served as the probe atom to calculate steric and electrostatic fields. The default value of 30.0 kcal/mol was used as the maximum electrostatic and steric energy cutoff.

Partial least squares (PLS) and Cross-validation in CoMFA

The last step in a CoMFA is a partial least square analysis to determine the minimal set of grid points which is necessary to explain the biological activities of the compounds. Partial least-squares is an iterative procedure that applies two criteria to produce its solution. First, to extract a new component, the criterion is to maximize the degree of commonality between all of the structural parameter columns (independent variable) collectively and the experimental data (dependent variable). Second, in the evaluation phase of a PLS iteration, the criterion for acceptance of the principal component just generated is an improvement in the ability to predict, not to reproduce, the dependent variable.

The technique used in PLS to assess the predictive ability of a QSAR is cross-validation [11]. Cross-validation is based on the idea that the best way to assess predictive performance is to predict. When cross-validating, one pretends that one or more of the unknown experimental value is, in fact, unknown. The analysis being cross-validated is repeated, excluding the temporarily 'unknown' compounds and then using the resulting equation to predict the experimental measurement of the omitted compound(s). The cross-validation cycle is repeated until each compound has been excluded and predicted exactly once. The results of cross-validation are the sum of the squared prediction errors, sometimes called the predicted residual sum of squares (PRESS). For evaluation of the overall analysis, the PRESS is commonly expressed as a cross-validated correlation coefficient r^2_{cv} , or $xv - r^2$, value.

RESULTS AND DISCUSSION

The results of the CoMFA studies are summarized in **Table-3**. From this table it is evident that the CoMFA derived QSAR shows a good cross validated r^2 , (0.579) and conventional r^2 , 0.976, therefore indicates a considerable predictive and correlative capacity of the model. In this analysis both steric and electrostatic field contribute to the QSAR equation by 73.8% and 26.2%, respectively, suggesting that variation in biological activity of compounds is dominated by differences in steric (van der Waals) interactions.

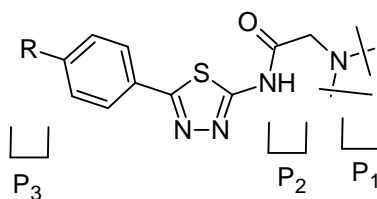


Figure-1

Table-3: Summary of CoMFA results

r^2 conventional	0.976
Standard error of estimate	0.031
F value	154.423
P value	0.000
r^2 cross-validated	0.579
Standard error of predictions	0.116
No. of components	6
Steric contribution	0.738
Electrostatic contribution	0.262

* Results from leave one out (LOO) cross validation analysis using six components.

Table-4: Data from PLS Cross-validated analysis (For Training Set)

Compound No.	Actual log (BA)	Calculated log (BA)	Residual
01.	-0.79	-0.82	0.03
02.	-0.53	-0.53	0.00
03.	-0.56	-0.52	-0.04
04.	-0.64	-0.64	0.00
05.	-0.75	-0.74	-0.01
06.	-0.60	-0.61	0.01
07.	-0.80	-0.78	-0.02
08.	-0.39	-0.44	0.05
09.	-0.79	-0.78	-0.01
10.	-1.16	-1.17	0.01
11.	-0.52	-0.54	0.02
12.	-0.49	-0.47	-0.02
13.	-0.61	-0.64	0.03
14.	-0.72	-0.74	0.02
15.	-0.60	-0.61	0.01
16.	-0.81	-0.79	-0.02
17.	-0.79	-0.76	-0.03
18.	-0.47	-0.44	-0.03
19.	-0.63	-0.66	0.03
20.	-0.91	-0.93	0.02
21.	-0.52	-0.52	0.00
22.	-0.74	-0.75	0.01
23.	-0.65	-0.61	-0.04
24.	-0.85	-0.80	-0.05
25.	-0.68	-0.68	0.00
26.	-1.12	-1.10	-0.02
27.	-0.67	-0.70	0.03
28.	-0.84	-0.83	-0.01
29.	-0.73	-0.79	0.06
30.	-0.86	-0.84	-0.02

The real test for model predictiveness is to predict the activity of ligands, which were not used in the model generation. Our test set has 10 ligands or compounds, which were randomly kept aside as a test set. The CoMFA models exhibited a good predictiveness on these ligands (Table-4).

To visualize the CoMFA steric and electrostatic fields from PLS analysis, contour maps of the product of the standard deviation associated with the CoMFA column and coefficient (SD X coeff.) at each lattice point were generated. The contour maps are plotted as percentage contribution to the QSAR equation and are associated with the differences in biological activity. In **Figure-2a** the regions of high and low steric tolerance are shown in green and yellow polyhedral, respectively. The areas of high bulk tolerance (80% contribution) are observed near P1 and P2 positions of the ligands (Figure-1). The active analogue (SM-6) shown in Figure-1a, shows that cyclohexyl ring embedded in the green region at P1 site. The diuretic activity shown by the compounds SM- 4, SM-5, SM-7, SM-16, SM-17, SM-18, SM-20 and SM-29 was due to the presence of bulky groups in P1 position surrounded by green contours in steric field plot. In the present sterically unfavored yellow regions were observed near the P3 position. The steric bulk in this region has a negative effect on the activity as represented by low activity of the compounds SM-13, SM-14, SM-25 and SM-26. Sterically unfavored yellow contours are also present at P1 position, embedded in the surrounding green contours, suggesting that there is a definite requirement of a substructure with appropriate shape to exhibit high activity.

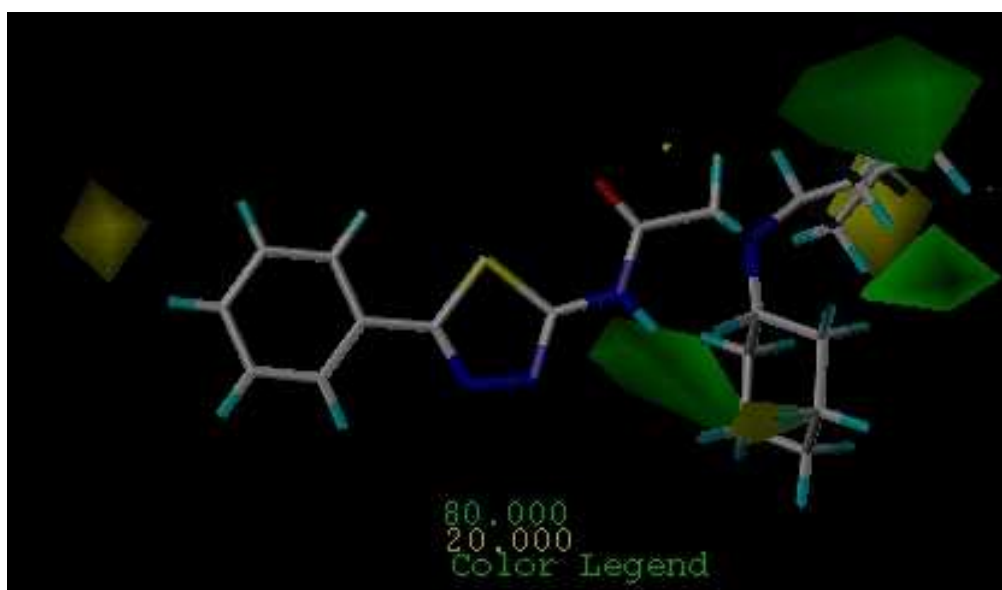


Figure-2a: Steric contour plot: favored (contribution level 80%) and unfavored (contribution level 20%) areas are represented as green and yellow contours, respectively.

CoMFA electrostatic fields are shown as blue and red polyhedral in **Figure- 2b**. A low electron density within the molecules near blue and red polyhedral, respectively, increases or decreases the activity and vice versa. Presence of a blue contour at P1 and P3 position suggesting that a low electron density in this area will have a positive effect on the biological activity and substructures with high electron density will reduce the activity. Presence of red contours at P2 and P1 position suggest that high electron density in this region increases the activity.

Though the electrostatic field contributions are less, a small change in electrostatic interactions will have a considerable effect on the activity.

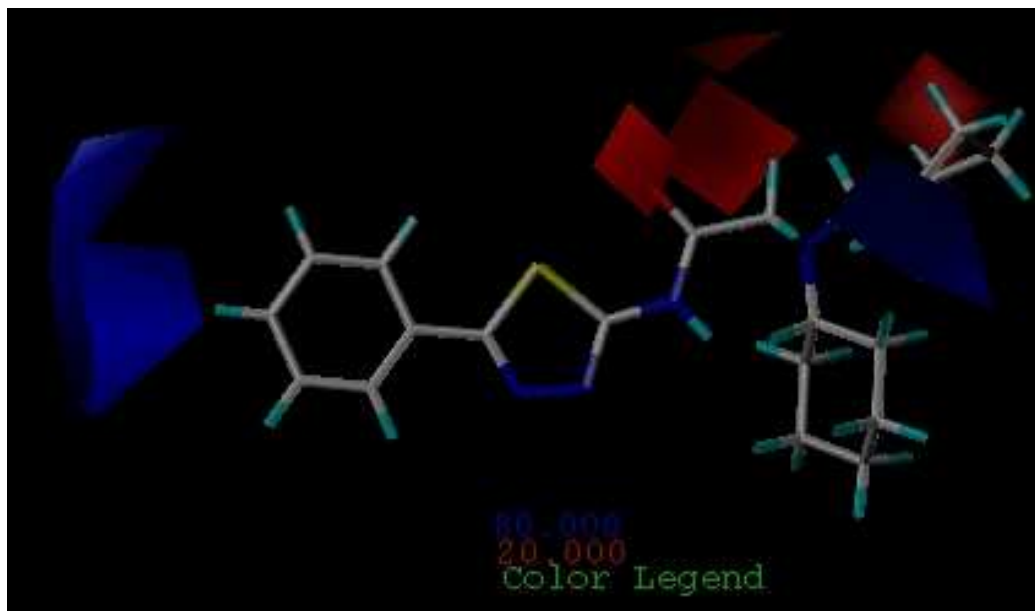


Figure-2b: Electrostatic contour plot: positive (contribution level of 80%) and negative (contribution level of 20%) charge favoring areas are represented as blue and red contours, respectively.

Acknowledgement

We are thankful to the Head, Department of Pharmaceutical Sciences, Dr. H. S. Gour University, Sagar (MP) and Molecular Modeling Division, IICT, Hyderabad (AP) for their help. SKJ is thankful to UGC, New Delhi for the award of JRF during his PhD.

REFERENCES

- [1] R.D. Cramer, D.E. Patterson, J.D. Bunce, *J. Am. Chem. Soc.*, **1988**, 110, 5959-5967.
- [2] M. Clark, R.D. Cramer, D.M. Jones, D.E. Patterson, P.E. Simeroth, *Tetrahedron Comput. Methodology*, **1990**, 3, 47-59.
- [3] P.S. Charifson (ed): "Practical Application of Computer-Aided Drug Design", Marcel Dekker, Inc., New York, **1997**.
- [4] S. S. Kulkarni, L. K. Gediya, V.M. Kulkarni, *Bioorg. Med. Chem.*, **1999**, 7, 1475.
- [5] S. S. Kulkarni, V. M. Kulkarni, *J. Med. Chem.*, **1999**, 42, 373.
- [6] M. L. Brown, C. C.Zha, C. C. Van Dyke, G. B. Brown, W. J. Brouillette, *J. Med. Chem.*, **1999**, 42, 1537.
- [7] S. K. Jain, P. Mishra, *Indian J. Chem.*, 43B, **2004**, 184.
- [8] S. K. Jain, PhD thesis, Dr. H.S. Gour University (Sagar, India, 2001).
- [9] S. K. Jain, P. Mishra, *Asian J. Chem.*, 23 (3), **2011**, 1305.
- [10] W.W. Wilkerson, *Eur. J. Med. Chem.*, 30, **1995**, 191.
- [11] R.D. Cramer, J.D. Bunce, D.E. Patterson, *Quant. Struct. Act. Relat.*, **1988**, 7, 18.