3D-QSAR Studies on 2-(indol-5-yl)thiazole Derivatives as Xanthine Oxidase (XO) Inhibitors

Santhosh Kumar Nagarajan and Thirumurthy Madhavan[†]

Abstract

Xanthine Oxidase is an enzyme, which oxidizes hypoxanthine to xanthine, and xanthine to uric acid. It is widely distributed throughout various organs including the liver, gut, lung, kidney, heart, brain and plasma. It is involved in gout pathogenesis. In this study, we have performed Comparative Molecular Field Analysis (CoMFA) on a series of 2-(indol-5-yl) thiazole derivatives as xanthine oxidase (XO) inhibitors to identify the structural variations with their inhibitory activities. Ligand based CoMFA models were generated based on atom-by-atom matching alignment. In atom-by-atom matching, the bioactive conformation of highly active molecule 11 was generated using systematic search. Compounds were aligned using the bioactive conformation and it is used for model generation. Different CoMFA models were generated using different alignments and the best model yielded a cross-validated q^2 of 0.698 with five components and non-cross-validated correlation coefficient (r^2) of 0.992 with Fisher value as 236.431, and an estimated standard error of 0.068. The predictive ability of the best CoMFA models was found to be r^2_{pred} 0.653. The CoMFA study revealed that the R_3 position of the structure is important in influencing the biological activity of the inhibitors. Electro positive groups and bulkier substituents in this position enhance the biological activity.

Keywords: Xanthine Oxidase, Gout, 3D-QSAR, CoMFA

1. Introduction

Xanthine oxidase is an enzyme that generates reactive oxygen species. The enzyme oxidizes hypoxanthine to xanthine, and xanthine to uric acid, producing hydrogen peroxide. XO is widely distributed throughout various organs. This includes the liver, gut, lung, kidney, heart, brain and plasma[1] with the highest levels being found in the gut and the liver[2]. It is localized to the capillary endothelial cells, in the myocardium[3]. XO is the only enzyme capable of catalyzing the formation of urate in man[4]. There is a large variability in human XOR expression which can be up to three-fold and on average 20% higher in men than in women[5].

The active form of the enzyme is a homodimer of molecular mass 290 kDa, with each of the monomers acting independently in catalysis. Each subunit mole-cule is composed of an N-terminal 20-kDa domain containing two iron sulfur centers, a central 40-kDa FAD domain, and a C-terminal 85-kDa molybdopterin-binding domain with the four redox centers aligned in an almost linear fashion. The hydroxylation of xanthine takes place at the molybdopterin center, and the electrons thus introduced are rapidly transferred to the other linearly aligned redox centers[6,7].

Xanthine oxydase is involved in gout pathogenesis. Gout is a disease in which defective metabolism of uric acid causes arthritis, especially in the smaller bones of the feet, deposition of chalk-stones, and episodes of acute pain. Gout is a common disease with prevalence of >2% in men older than 30 years and in woman older than 50 years according to the Third National Health and Nutrition Examination Survey (1988-1994)[8,9]. Gout occurs in individuals who have high serum uric acid levels, in response to precipitation of monosodium urate monohydate crystals in various tissues, followed by an inflammatory response. Typical symptoms of gout include acute recurrent gouty arthritis, a tophinod-ular collection of monosodium urate crystals and uric acid urolithiasis[10]. XO also plays an important role in

Department of Bioinformatics, School of Bioengineering, SRM University, SRM Nagar, Kattankulathur, Chennai 603203, India

[†]Corresponding author : thiru.murthyunom@gmail.com, thirumurthy.m@ktr.srmuniv.ac.in

various forms of ischemic and other types of tissue and vascular injuries, inflammatory diseases, and chronic heart failure[11].

A xanthine oxidase inhibitor is any substance that inhibits the activity of xanthine oxidase. Xanthine oxidase inhibitors are of two kinds: purine analogues and others. Purine analogues include allopurinol, oxypurinol, and tisopurine. Others include febuxostat, topiroxostat, and inositols (phytic acid and myo-inositol). These commercially available drugs have certain disadvantages. For Example, Allopurinol has two important disadvantages. First, its dosing is complex[12]. Second, some patients are hypersensitive to the drug,[13] therefore its use requires careful monitoring. Allopurinol has rare but potentially fatal adverse effects involving the skin. The most serious adverse effect is a hypersensitivity syndrome consisting of fever, skin rash, eosino-
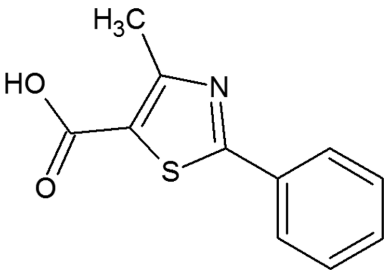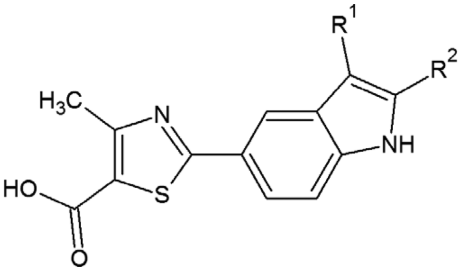
philia, hepatitis, worsened renal function, and, in some cases, allopurinol hypersensitivity syndrome[13]. The adverse effects of febuxostat include nausea, diarrhea, arthralgia, headache, increased hepatic serum enzyme levels and rash[14]. Hence, the discovery of structurally diverse XO inhibitors becomes imoportant.

2. Materials and Methods

2.1. Data Set

A series of 21 compounds with their biological activities were taken from the literature[15]. The compounds were divided into training set and test set molecules. The test set molecules were selected manually in order to cover low, middle, and high biological activity from the dataset. For QSAR analysis, the given inhibitory concentration values (IC_{50}) were

Table 1. Structures and biological activities (pIC_{50}) of Xanthine Oxidase inhibitors

The Xanthine Oxidase inhibitor scaffold

a) Compound 1-7

Compound	R^1	R^2	pIC50 values
1	H	H	6.9586
2	Cl	H	7.9706
3	Cl	H	8.5529
4	Cl	CH$_3$	8.0458
5	NO$_2$	H	7.9101
6	NO$_2$	H	8.2518
7	CN	H	8.2218

Table 1. Continued

b) Compound 8-11
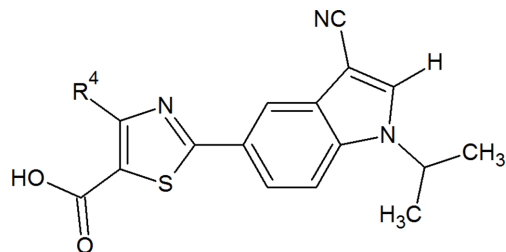
Compound	R^4	pIC50 values
8	H	8.4089
9	CF_3	8.0655
10	OCH_3	7.0457
11	CH_3	8.2441

c) Compound 12-21
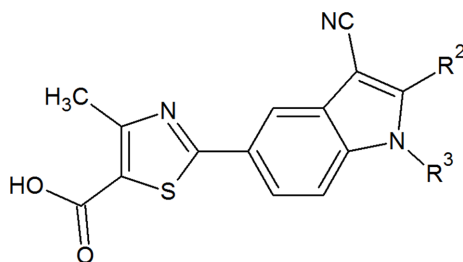
Compound	R^2	R^3	pIC50 values
12	H	2-methylpropane	7.7959
13	CH_3	1-fluoro-2-methylpropane	8.3767
14	H	2-methylpropan-1-ol	7.9208
15	H	1-methoxy-2-methylpropane	8.2596
16	H	1-(methylsulfonyl)propane	8.3565
17	H	N-propylacetamide	6.0482
18	H	N-propylmethanesulfonamide	7.0458
19	H	ethylbenzene	8.4559
20	H	1-ethyl-2,4-difluorobenzene	8.320

d) Febuxostat

| Compound	Structure	pIC50 values
21		8.2596

*Test set compounds

changed to minus logarithmic scale value (pIC$_{50}$), as a dependent variable for 3D-QSAR analysis.

$$pIC_{50}= -\log (IC_{50}).$$

The structures and biological activities of all compounds including both training set and test set molecules is shown in Table 1.

2.2. Partial Atomic Charge Calculation and Bioactive Conformation Determination

For each compound, the partial atomic charges were assigned by utilizing Gasteiger Hückel method available in SYBYLX 2.1 package (Tripos Inc., St. Louis, MO, USA). The bioactive conformation of the highly active compound 11 was determined using systematic search method. The obtained conformation from these methods was used as the template to perform alignment for CoMFA calculations. During systematic search all rotatable bonds were searched with incremental dihedral angle from 120° and the lowest energy conformer was assumed as bioactive conformer and was chosen for subsequent QSAR modeling. Bioactive conformations for other compounds were determined by systematic conformer search for the additional moieties, by keeping core part constrained. It was seen
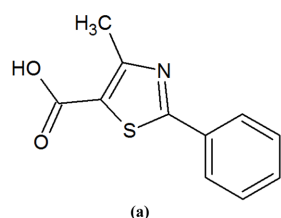
(a)
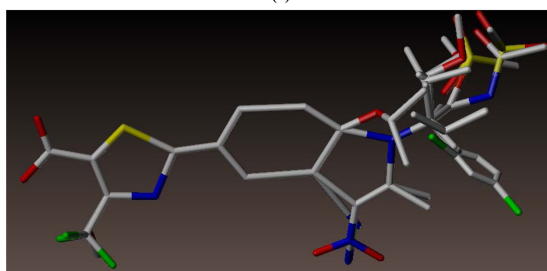
(b)

Fig. 1. (a) Maximum common substructure present in all molecules. (b) Alignment of molecules based on systematic search conformation of highly active compound 11.

that, the common scaffold of every inhibitor occupies the same area in 3D space. The matching atoms were selected by maximum common substructure (MCS) as shown in Fig. 1(a) and the alignment of all the molecules are represented in Fig. 1(b).

2.3. CoMFA Model Calculation

CoMFA was performed with the generally used steric and electrostatic fields in SYBYL-X 2.0. The aim of CoMFA analysis is to derive a correlation between the biological activity of a set of molecules and their 3D shape, electrostatic and hydrogen bonding characteristics. This correlation is derived from a series of super-imposed conformations aligned using the atom-by-atom matching. Each conformation is taken in turn, and the molecular fields around it are calculated. The electrostatic and steric (van der Waals interactions) fields are measured at the lattice points of a regular Cartesian 3D grid. The lattice spacing is typically 2 Å. The lattice was defined automatically and is extended 4 units past Van der Waals volume of all molecules in x, y, and z directions. Using Tripos force field, the Van der Waals potential and columbic terms, which represent steric and electrostatic fields, respectively, were calculated. A distance-dependent dielectric constant was used. The CoMFA method was performed using steric and electrostatic fields with standard ±30 kcal/mol cutoffs. A series of models were constructed with an increasing number of partial least squares (PLS) analysis factors. The numbers of components in the PLS models were optimized by using the cross-validated correlation coefficient (q^2), non-cross-validated correlation coefficient (r^2), standard error estimate (SEE) and Fisher's values (F), etc., which were obtained from the leave-one-out (LOO) cross-validation procedures[16,17]. After the optimal number of components was determined, a non-cross-validated analysis was carried out without column filtering.

2.4. Partial Least Squares (PLS) Analysis

The CoMFA equation was established by partial least-squares (PLS) analysis between biological activity and molecular parameters. By using the PLS method[18,19] the CoMFA descriptors were linearly correlated with the biological activity values. The CoMFA cutoff values were set to 30 kcal/mol for both the steric and electrostatic fields. Leave-one-out method (LOO

Method) was used to perform the cross-validation analysis. The cross-validated correlation coefficient (q^2) was calculated using the following equation:

$$q^2 = 1 - \frac{\sum_{\gamma}(\gamma_{pred} - \gamma_{actual})^2}{\sum_{\gamma}(\gamma_{actual} - \gamma_{mean})^2}$$

where γ_{pred}, γ_{actual}, and γ_{mean} are the predicted, actual, and mean values of the target property (pIC_{50}), respectively.

2.5. Predictive Correlation Coefficient (r^2_{pred})

The predictive power of CoMFA models were determined from the set of eight test molecules which was excluded during model development.[20-24] In the structural preparation of test set molecules, sketching and optimization was same as the training set molecules. The activity of the test set was predicted by using model derived from training set. The predictive correlation coefficient (r^2_{pred}), based on the test set molecules, is defined as:

$$r^2_{pred} = \frac{(SD - PRESS)}{SD}$$

where, PRESS is the sum of the squared deviation between the predicted and actual activity of the test set

molecules, and SD is defined as the sum of the square deviation between the biological activity of the test set compounds and the mean activity of the training set molecules.

2.6. CoMFA Contour Maps

CoMFA analysis generates color-coded contour maps that represent regions in three dimensional space where changes in the steric and electrostatic fields of a compound correlate strongly with changes in its biological activity. Contour maps were generated as a scalar product of coefficients and standard deviation (StDev * Coeff) associated with each column. Favored levels were fixed at 70% and disfavoured levels were fixed at 30%. The contours for steric fields are shown in green (more bulk favored) and yellow (less bulk favored). The electrostatic field contours are displayed in red (electronegative substituents favored) and blue (electropositive substituents favored) colors.

3. Results and Discussion

In the present work, ligand-based CoMFA models was generated on a series of 2-(indol-5-yl) thiazole derivatives as xanthine oxidase inhibitors. Different combinations of 15 training and 6 test compounds were

Table 2. Statistical results of CoMFA models obtained from systematic search conformation based alignment

PLS statistics	Ligand-based CoMFA model (systematic search)				
	Model 1	Model 2	Model 3	Model 4	Model 5
q^2	0.698	0.673	0.670	0.623	0.606
N	5	5	5	5	3
r^2	0.992	0.993	0.995	0.996	0.978
SEE	0.068	0.066	0.056	0.047	0.108
F-value	236.431	250.376	346.078	483.371	160.944
r^2_{pred}	0.653	0.581	0.601	0.545	0.512
Field contribution					
Steric	0.552	0.604	0.568	0.540	0.577
Electro static	0.448	0.396	0.432	0.460	0.423

q^2= cross-validated correlation coefficient; N= number of statistical components; r^2= non-cross validated correlation coefficient; SEE=standard estimated error; F=Fisher value; $r^2_{predictive}$= predictive correlation coefficient for test set. The model chosen for analysis is highlighted in bold fonts.
Test set compounds
Model 1 - compound no 1,3,9,13,15,16
Model 2 - compound no 1,3,9,13,16,21
Model 3 - compound no 1,3,9,12,13,16
Model 4 - compound no 1,3,9,13,16,19
Model 5 - compound no 1,5,9,13,15,16

used for model generation. Alignment was based on atom-by-atom matching method. Many CoMFA models were obtained, of those only 5 models from systematic search conformation were selected based on higher q^2 and r^2_{pred} values.

3.1. CoMFA Model Analysis

Ligand-based CoMFA models was generated using different combinations of training set and test compounds. Based on higher q^2 and r^2_{pred} values only 5 models from systematic search were selected. The

Table 3. Predicted activities and experimental pIC_{50} values obtained from CoMFA models

Compound	Actual pIC_{50}	Predicted	Residual
1*	6.9586	7.504	0.396
2	7.9706	7.965	0.0056
3*	8.5529	8.533	-0.0 101
4	8.0458	8.08	-0.0342
5	7.9101	7.93	-0.0199
6	8.2518	8.21	0.0418
7	8.2218	8.271	-0.0492
8	8.4089	8.343	0.0659
9*	8.0655	8.449	-0.3835
10	7.0457	8.246	-0.0019
11	8.2441	8.044	0.0927
12	7.7959	7.898	-0.1021
13*	8.3767	8.824	-0.4473
14	7.9208	7.916	0.0048
15*	8.2596	7.898	0.3616
16*	8.3565	7.827	0.5295
17	6.0482	6.014	0.0342
18	7.0458	7.086	-0.0402
19	8.4559	8.395	0.0609
20	8.320	8.383	-0.0732
21	8.2596	8.245	0.0146

*Test set compounds

(a) CoMFA - Training Set   **(b) CoMFA - Test Set**
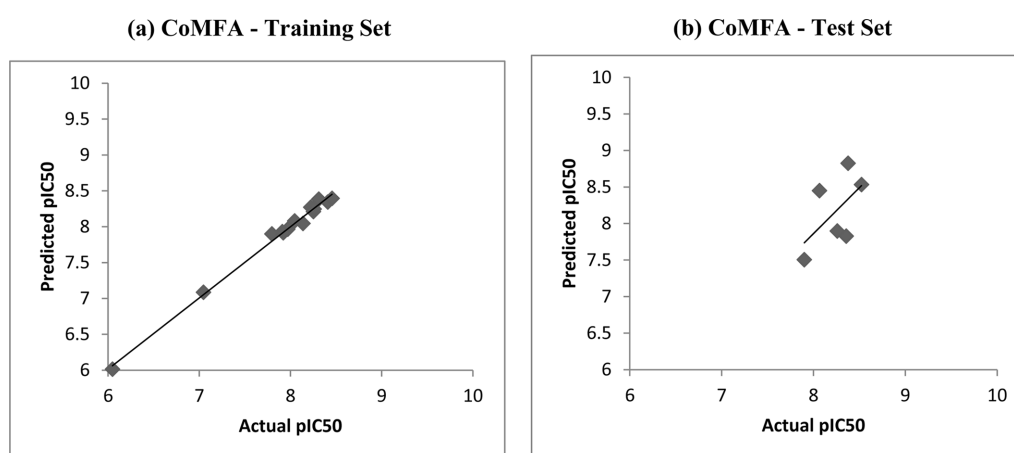

Fig. 2. (a and b) Plot of actual versus predicted pIC50 values for the training set and test set for the CoMFA values atom-by atom matching alignment by systematic search.

statistical values of the 5 models from different alignments are tabulated in Table 2 for systematic search. The best statistics were found in Model 1 having a cross-validated q^2 of 0.698 with five components and non-cross-validated correlation coefficient (r^2) of 0.992 with Fisher value as 236.431, and an estimated standard error of 0.068. The predictive ability of the best CoMFA models was found to be r^2_{pred} 0.653. Predicted and experimental pIC$_{50}$ and their residual values in Table 3 and their corresponding scatter plot is depicted in Fig. 2.

3.2. Mapping of CoMFA Contour Maps

The CoMFA contour map for the generated models

was analysed to explore the important structural features. The steric and electrostatic contour map obtained for systematic search alignment based CoMFA model are displayed in Fig. 3 and 4. The steric interactions are represented by green and yellow contours. The green steric contour near the R_3 position indicates that bulkier substituents are preferred at this position. Thus, compounds 20, 21 and 6 with bulkier substituents at these positions are more active than compounds which have smaller substituents at the R_3 position. The electrostatic interactions are represented by red and blue contours. The blue contours represent that the electropositive substituents are favored in that region. The
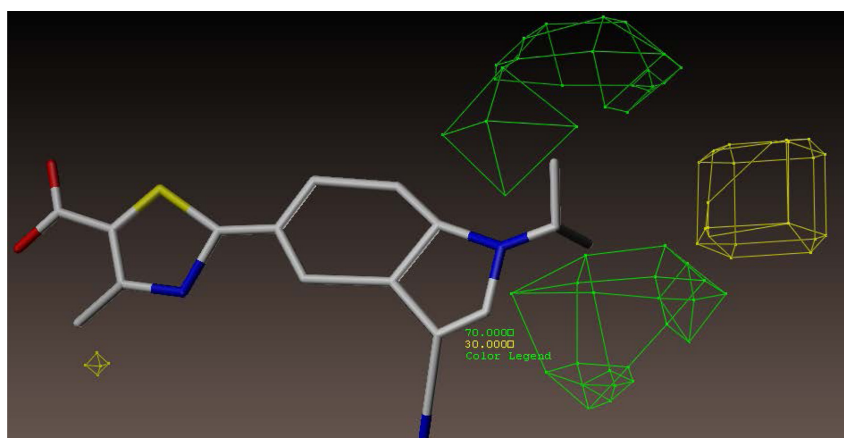
Fig. 3. CoMFA steric contour map with highly active compound 11 for systematic search based alignment. Here green contour indicates region where bulky group increases activity and yellow contours indicates bulky group decreases activity.
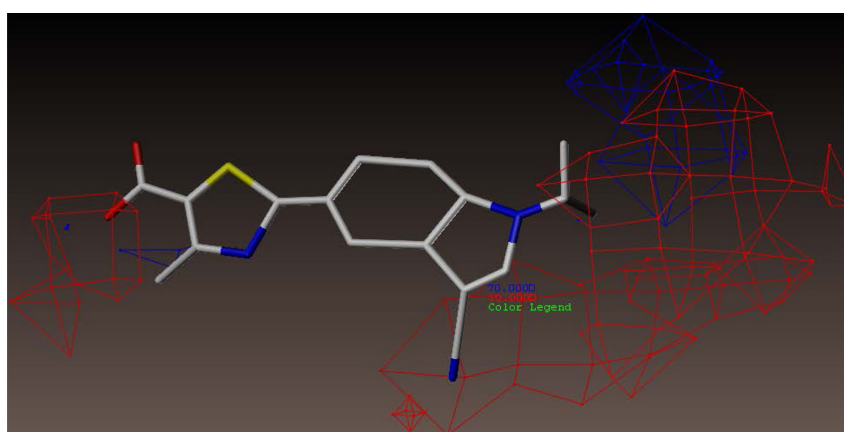
Fig. 4. CoMFA electrostatic contour map with highly active compound 11 for systematic search based alignment. Here blue contour indicates regions where electropositive groups increases activity and red contours indicates regions where electronegative groups increases activity.

electrostatic contour plot shows that there is a big blue colored region near the R_3 position. Electro positive group linked to this position will enhance the biological activity. For example, compounds 16, 17 and 18, which are substituted with 1-(methylsulfonyl)propane, N-propylacetamide and N-propylmethanesulfonamide at that position is in direct contact with the big blue contour map. The red contours represent that the electronegative substituents are favored in that region. For example, the substituents at the R_3 position of compounds 1 and 20 are in direct contact with the red contour map, which reduce the biological activity.

4. Conclusion

In this study, we have developed CoMFA models of 2-(indol-5-yl) thiazole derivatives as xanthine oxidase inhibitors. CoMFA models with good cross-validated q^2 values and predictive ability were obtained. The statistical results showed the important structural alignments required for the inhibitory activity of the 2-(indol-5-yl) thiazole derivatives. The results correlate the structural features with the inhibitory activities against xanthine oxidase and provide valuable information about characteristics of inhibitors. The present CoMFA study revealed that the R_3 position of the structure is important in influencing the biological activity of the inhibitors. Electro positive groups and bulkier substituents linked to this position enhance the biological activity. The results obtained from this study have thrown light on the important structural and chemical features in designing and developing new potent novel inhibitors for xanthine oxidase.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] P. Pacher, A Nivorozhkin, and C. Szabó, "Therapeutic effects of xanthine oxidase inhibitors: Renaissance half a century after the discovery of allopurinol", Pharmacol. Rev., Vol. 58, pp. 87-114, 2006.

[2] D. A. Parks and D. N. Granger, "Xanthine oxidase: Biochemistry, distribution and hysiology", Acta

Physiologica Scandinavica. Supplementum, Vol. 548, pp. 87-99, 1986.

[3] M. Cicoira, L. Zanolla, A. Rossi, G. Golia, L. Franceschini, G. Brighetti, P. Zeni, and P. Zardini, "Elevated serum uric acid levels are associated with diastolic dysfunction in patients with dilated cardiomyopathy", Am. Heart J., Vol. 143, pp. 1107-1111, 2002.

[4] K. D. Pfeffer, T. P. Huecksteadt, and J. R. Hoidal, "Xanthine dehydrogenase and xanthine oxidase activity and gene expression in renal epithelial cells. cytokine and steroid regulation", J. Immunol., Vol. 153, pp. 1789-1797, 1994.

[5] R. Guerciolini, C. Szumlanski, and R. M. Weinshilboum, "Human liver xanthine oxidase: nature and extent of individual variation", Clin. Pharmacol. Ther., Vol. 50, pp. 663-672, 1991.

[6] R. Harrison, "Structure and function of xanthine oxidoreductase: Where are we now?", Free Radical Bio. Med., Vol. 33, pp. 774-797, 2002.

[7] R. Harrison, "Physiological roles of xanthine oxidoreductase". Drug Metab. Rev., Vol. 36, pp. 363-375, 2004.

[8] H. M. Kramer and G. Curhan, "The association between gout and nephrolithiasis: the national Hhealth and nutritieExamination survey III, 1988-1994", Am. J. Kidney Dis., Vol. 40, pp. 37-42, 2002.

[9] H. K. Choi and G. Curhan, "Gout: Epidemiology and lifestyle choices", Curr. Opin. Rheumatol., Vol. 17, pp. 341-345, 2005.

[10] R. L. Wortmann, "Recent advances in the management of gout and hyperuricemia", Curr. Opin. Rheumatol., Vol. 17, pp. 319-324, 2005.

[11] J. George and A. D. Struthers, "The role of urate and xanthine oxidase Inhibitors in cardiovascular disease", Cardiovascular Drug Reviews, Vol. 58, pp. 59-64, 2008.

[12] N. Dalbeth and L. Stamp, "Allopurinol dosing in renal impairment: Walking the tightrope between adequate urate lowering and adverse events", Seminars in Dialysis, Vol. 20, pp. 391-395, 2007.

[13] T.-F. Tsai and T.-Y. Yeh, "Allopurinol in dermatology", Am. J. Clin. Dermatol., Vol. 11, pp. 225-232, 2010.

[14] B. L. Love, R. Barrons, A. Veverka, and K. M. Snider, "Urate-lowering therapy for gout: focus on febuxostat", Pharmacotherapy, Vol. 30, pp. 594-608, 2010.

[15] J. U. Song, S. P. Choi, T. H. Kim, C.-K. Jung, J.-Y. Lee, S.-H. Jung, and G. T. Kim, "Design and

synthesis of novel 2-(indol-5-yl)thiazole derivatives as xanthine oxidase inhibitors", Bioorg. Med. Chem. Lett., Vol. 25, pp. 1254-1258, 2015.

[16] K. Abe, H. Shimokawa, K. Morikawa, T. Uwatoku, K. Oi, Y. Matsumoto, T. Hattori, Y. Nakashima, K. Kaibuchi, K. Sueishi, and A. Takeshit, "Long-term treatment with a Rho-kinase inhibitor improves monocrotaline-induced fatal pulmonary hypertension in rats", Circ. Res., Vol. 94, pp. 385-393, 2004.

[17] T. Ishizaki, M. Uehata, I. Tamechika, J. Keel, K. Nonomura, M. Maekawa, and S. Narumiya, "Pharmacological properties of Y-27632, a specific inhibitor of Rho-associated kinases", Mol. Pharmacol., Vol. 57, pp. 976-983, 2000.

[18] P. Geladi and B. R. Kowalski, "Partial least-squares regression: a tutorial", Anal. Chim. Acta, Vol. 185, pp. 1-17, 1986.

[19] S. Wold, "Cross-validatory estimation of the number of components in factor and principal components models", Technometrics, Vol. 20, pp.

397-406, 1978.

[20] B. Sathya and T. Madhavan, "Comparative molecular field analysis of caspase-3 Inihibitors", J. Choun Natural Sci., Vol. 7, pp. 166-172, 2014.

[21] P. Singh and T. Madhavan, "Histone deactylase inhibitors as novel target for cancer, diabetes, and inflammation", J. Chosun Natural Sci., Vol. 6, pp. 57-63, 2013.

[22] S. Kulkarni and T. Madhavan, "Application of docking methods: an effective in Silico tool for drug design", J. Chosun Natural Sci., Vol. 6, pp. 100-103, 2013.

[23] M. Shalini and T. Madhavan, "Homology modeling of CCR 4: novel therapeutic target and preferential maker for Th2 Cells", J. Chosun Natural Sci., Vol. 7, pp. 234-240, 2014.

[24] B. Sathya and T. Madhavan, "Comparative molecular mimilarity indices analysis of caspase-3 inhibitors", J. Chosun Natural Sci., Vol. 7, pp. 227-233, 2014.