International Journal of Engineering, Science and Mathematics
Vol. 6 Issue 2, June 2017,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijesm.co.in, Email: ijesmj@gmail.com

QSAR and Pharmacophore Model Generation from Corticosteroid Derivatives to Design Active Molecules for the Treatment of Asthma

Sarithamol S, V.L.Pushpa*

P G and Research Department of Chemistry, Sree Narayana College, Kollam

Abstract

Corticosteroids are the popular medications used for the disease, asthma. The side effects reported with the usage of these candidates made the research world to think of new candidates with least adverse effects and more therapeutic efficiency. The present study has performed a 2D QSAR analysis on a set of corticosteroid derivatives with their glucocorticiod receptor agonistic property. A statistically significant model with R^2 =0.8261, Q^2= 0.6, p and F in the significant limits were obtained. The model pointed out that the descriptor properties, SHdsCH and Weta3.mass has highly correlated with the GR agonistic property of the compounds. The model validation has carried out by leave one out cross validation methods. Atom based 3D QSAR was developed and found that their occurs a relationship between higher dimensional descriptors such as hydrogen bond donor effect, hydrophobic effect and electron withdrawing effects of the molecules in the dataset we studied. A five feature pharmacophore were derived from the compounds with their glucocorticoid receptor agonistic activity. Thus the statistically significant QSAR models created through the present study and pharmacophoric feature has engendered some valuable information for the designing of promising candidates with very good therapeutic efficacy and least side effects.

Introduction

Asthma is one of the most serious inflammatory lung disorders characterized by airway remodeling, hyperresponsiveness, and thick mucus secretion and difficult to breath. A large variety of targets are involved in the disease generation[1]. This makes the treatment more difficult. No medicine is reported for making a complete curing effect. All the popular medicines including bronchodilators, corticosteroids, antibody treatments etc are creating a short term relief from the severe condition[2]. Corticosteroids are the most useful medicines among them. They

along with bronchodilators as combination medicines will create a long term control against the disease. Eventhough corticosteroids are accepted as promising medicines for the present severe condition, the adverse effects reported in the clinical practices has created a discomfort with steroid safety, especially in children[3]. The adverse effects in children include bone metabolism, growth retardation and skin thinning. Therefore many studies have been reported for increasing the therapeutic efficacy of these medicines. Computer aided drug designing (CADD) is a popular technique widely used by the pharmaceutical companies. Quantitative structure activity relationship is the most accepted method in CADD[4]. This technique can be utilized for designing new compounds with desirable features and very little side effects through statistically significant models by the advent of an experimentally derived activity data of structurally similar compounds. The present study has executed a 2D and 3D QSAR study on a set of corticosteroid derivatives with their glucocoricoid receptor agonistic activity to derive good activity correlated models. Pharmacophore feature analysis was also executed. The models thus generated is used to design new steroid based compounds having promising therapeutic efficiency.

Materials and methods

Insilico data

A set of inhaled corticosteroid derivatives with their glucocorticoid agonistic activity was collected from the literature for the insilico QSAR analysis[5]. 2D QSAR and 3D QSAR studies were carried out on the set to derive the structural correlation with their biological activities. The activity concentration varied from 0.43nM to 21.10nM. The activity data was expressed as their negative logarithmic formation for statistical significance. The structure and activity values are represented in supplementary information (Table 1).

2D QSAR studies

Descriptor calculation

The whole dataset was subjected for calculation of physicochemical properties using the software, padel. The descriptor set include topological, geometrical, whim based, polarizability and correlation based type descriptors. The insignificant descriptors were removed from the

collection of 2000 descriptors using the software, Phakiso. The final resulted 400 descriptors were considered for the correlation studies with the dataset.

Regression analysis

The dataset was divided into training and test set in the ratio 80: 20. There were 12 compounds in the training set and 3 compounds in the test set. Multiple linear regression analysis was executed and genetic function algorithm was applied to find the well correlated descriptors from the set of 400 properties. 2000 iterations were given for GA calculations. All operations were carried out in the software, QSARINS [6].

QSAR model validation studies

The regression coefficient, R^2 stands for goodness of fit. The terms such as analysis of variance, F and statistical significance p also contribute for the validity of the model. External test set prediction was checked by leave one out cross correlation method. The cross correlation coefficient $Q^2 \geq 0.5$ indicate external predictability of the model[7]. For a better model, the following criteria should be followed, $\mathbf{r^2 < 0.6, Q2< 0.6, r_{pred}^2 < 0.5, p<0.0001.}$ The cross correlation coefficient Q^2 can be calculated by the equation,

$$\mathbf{Q^2 = 1-\{\sum(y_{obs}-y_{pred})^2/\sum(y_{obs}-y_m)^2\}}$$

Where y^t_{obs} and y^t_{pred} are the observed and predicted activities of the test set molecules and y_m is the mean activity of the training set molecules[6].

Pharmacophore modeling

Pharmacophore is the 3 dimensional functionality of a compound responsible for a concerned activity property[8]. Pharmacophore modeling studies can create a clear path for the virtual screening scenario. The present study has generated pharmacophores from a set of 15 corticoid derivatives with their glucocorticoid receptor agonistic activity. Phase module of Schrodinger software has been used for the analysis. The entire 15 compound were energy minimized by the ligprep module using the force field of OPLS_2005 at pH 7.4. All the possible conformations were generated from the set using Conf Gen method and keeping all other parameters in a default

state[9]. The method assigned that the compounds with pIC50 above 9.1 as active and below 8.1 as inactive. Thus the top 5 active compounds and least three inactive compounds were considered for the pharmacophore generation. The scoring of the features was done on the basis of site, vector and volume scores.

3D QSAR analysis

Atom based 3D QSAR analysis is a higher dimensional QSAR method and the descriptors here considering include hydrogen bond donor effect, hydrogen bond acceptor effect, hydrophobic effect, negative and positive ionic effects, and electron withdrawing effects. All the descriptors can be calculated easily by considering the compounds as whole. Partial least square regression method was done for the statistical analysis[10]. The IC50 values of compounds were represented into negative logarithmic values for the statistical significance. The whole dataset was split into 11 training and 4 test set compounds. The training and test set splitting has done in such a way that training set contain all the diverse set of compounds and test set should be a subset of training set. The whole dataset was aligned on the pharmacophore derived from the activity profile of the compounds.

Result and discussion

2D QSAR studies

Multiple linear regressions for statistical analysis and genetic algorithm for selecting most appropriate descriptor from the list of 400 descriptors were executed using the software, QSARINS. Three better models obtained were shown below.

1) Activity = 11.6143- 0.7686 SHdsCH - 4.4593 Weta3.mass,

R2= 0.8333, Q2loo= 0.5525, F = 22.4959, R2-Q2loo = 0.2808, Predictions by LOO: Exp(x) vs. Pred(y): R2: 0.7260

2) Activity = 12.2679 -0.3766 SdsCH -4.7446 Weta2.eneg

R2: 0.7241, F: 10.4981, Q2loo: 0.5726, R2-Q2loo: 0.1515,

Predictions by LOO:

Exp(x) vs. Pred(y): R2: 0.6103.

3) Activity = 11.3295 -0.7210 SHdsCH -4.0637 Weta3.mass

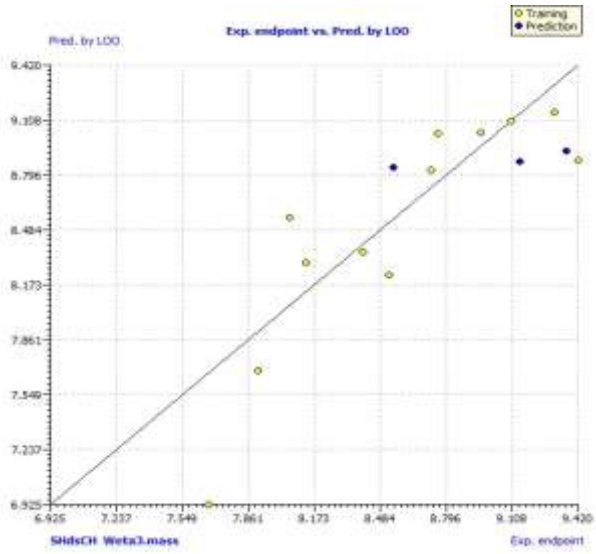
R2: 0.8261, R2adj: 0.7875, Q2loo: 0.6019, R2-Q2loo: 0.2242, F: 21.3812

Predictions by LOO: Exp(x) vs. Pred(y): R2: 0.7295

All the three models were created using the descriptors, SHdsCH, Weta3.mass and Weta2.eneg from 15 corticoid derivatives containing 12 training and 3 test set compounds. Among the three, third model is found to have high R^2 and Q^2 value. The high F varient value, high R^2 and R^2 adjusted indicate that the model fit is not a chance of occurance. The model validation is done by leave one out method. The cross correlation coefficient of the third model is 0.6019 which indicate the prediction ability and stability of the model. The prediction ability has been proved by the external prediction value, R^2= 0.7295. GR agonistic property has been highly correlated with the descriptor properties, SHdsCH and Weta3.mass through the model 3. The first and third model has correlated activity with these two descriptors. The models has explored that SHdsCH and Weta3.mass can create an impact on the GR agonistic activity of the steroid derivatives and these can take into consideration for finding out of more promising candidates having all the qualities of steroids. SHdsCH is a 2 dimensional descriptor coming under the category of Atom type electrotopological state descriptors, Sum of atom-type H E-State: =CH- and Weta3.mass is a whim based 3D descriptor coming under the category of geometrical descriptors, through which molecular 3D information regarding molecular mass with respect to invariant reference frames was indicated. Both the descriptors were creating a negative impact on the activity property. Decrease in values of both the properties will cause an increase in GR agonistic property. In detail, the increase in molecular mass and =CH group will create a negative impact on GR agonistic activity property. The third model has identified 10^{th} compound as outlier. Compound 10 is the least active compound also on considering the SHdsCH descriptor value for the compound 10; it is very high in the range of 3 while all others have values in the range of 1.6 , signifies the influence of =CH group in the structure once again. The lowest value of

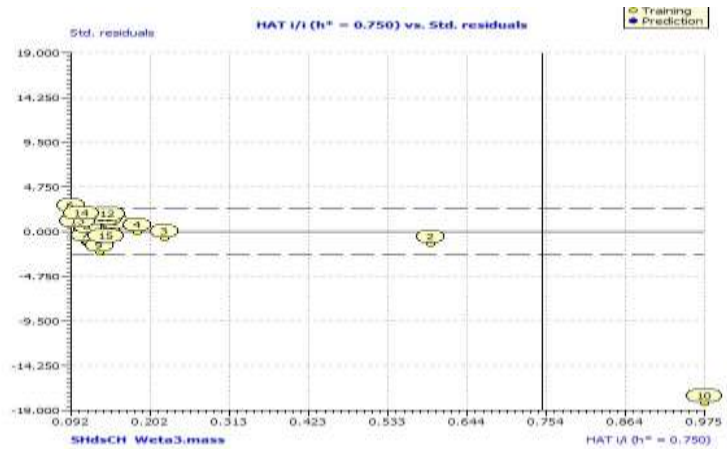
weta3.mass also made the compound as outlier. The details of predicted activity is listed in supplementary data (table 2)

Fig 1: model regression graph by LOO method



The outlier can be represented by considering Williams plot, utilized to represent applicability domain. Williams plot is the plot of residual vs leverage. Leverage is the distance of a compound from the centroid of X. X is the descriptor matrix derived from the descriptor values of training set. The data point lies outside the area created by the plot can be considered as outlier. In the present study, compound 10, the least active compound among, was in the outlier region.

Fig 2 : Williams plot for the data set (applicability domain)



Correlation matrix of descriptors

	SHdsCH	Weta3.mass
SHdsCH	1	
Weta3.mass	-0.1613	1

Correlation matrix indicates that there is no correlation between the descriptors. If there is a correlation above 0.7 happens, then, any one of the correlated descriptors at a time should be considered. Here the descriptors SHdsCH and Weta3.mass did not intercorrelated. For a significant QSAR model, there should not beany correlation between the descriptors; they should be independent of each other.

Pharmacophore modeling

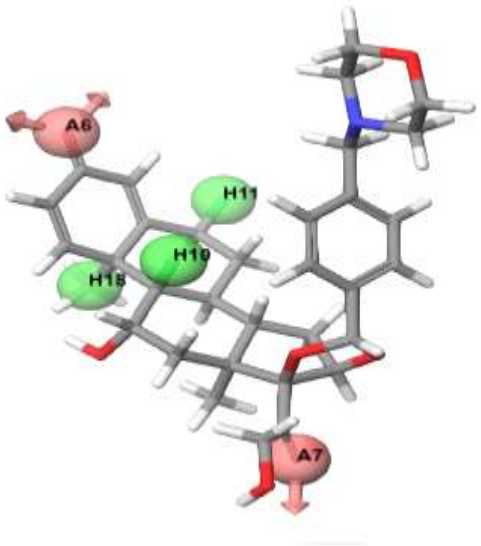
A potential three dimensional feature responsible for the glucocoricoid receptor agonistic activity has derived from 15 corticoid derivatives. The feature composed of 5 feature pharmacophore AAHHH. Phase module of schrodinger software has utilized to derive the pharmacophoric feature. A total of 136 pharmacophoric features were derived from the set of compounds. Among them, the feature having highest survival active and survival inactive scores has chosen as the most activity contributive feature. The top five features are represented in the table below.

Table 1: Pharmacophore scoring table

Pharmacophoric feature	Survival score	Survival inactive score	matches	Activity	Energy
AAHHH.15805	5.965	3.451	5	9.149	3.047
AAHHH.18307	5.931	3.689	5	9.367	1.5
AAHHH.18313	5.926	3.362	5	9.367	1.5
AAHHH.19004	5.911	3.339	5	9.149	2.274
AAHHH.17119	5.896	3.543	5	9.149	2.274

The pharmacophoric feature, AAHHH.15805, with the survival score 5.965 and survival inactive score 3. 451 have chosen as the potential feature.

Fig 3: potential pharmacophore feature, A- hydrogen bond acceptor, H- hydrophobic effect

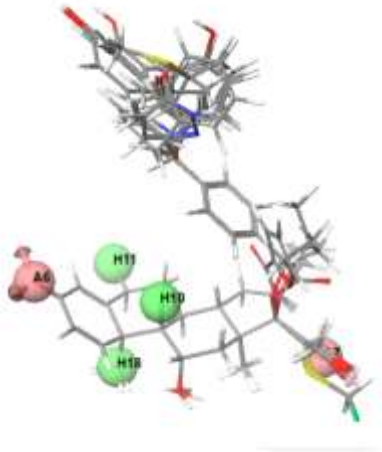


The feature composed of two hydrogen bond acceptor features and three hydrophobic features. The hydrogen bond acceptor feature has shown by carbonyl group attached to the cyclohexadienone ring portion of steroid skeleton and carbonyl group attached to the hexahydrocyclopentadioxole ring portion. The hydrophobic feature of methyl group attached to the cyclohexadienone ring group was responsible for one of the hydrophobic feature of the pharmacophore. The second and third hydrophobic features have generated by two fluorine atoms attached to the cyclohexyl ring of the steroid skeleton. The feature signifies the importance of steroid skeletal portion of those compounds towards the glucocorticoid receptor agonistic property.

3D QSAR analysis

Atom based 3D QSAR analysis has executed on the dataset of 15 corticoid derivatives. All the compounds were shown very good alignment with the pharmacophoric feature AAHHH.15805, furnished by the dataset. The aligned structure is shown in the figure.

Fig 4: dataset aligned with pharmacophoric feature



Partial least square statistics was applied to derive 3D correlation between the activity and structure of the dataset. The PLS factor given was 2 and the parameters obtained were SD= 0.1915, R^2= 0.9102, Q^2= 0.6724, F= 40.5, p= 6.50E-05, RMSE= 0.29 and Pearson r =0.9219. SD, R^2, F and p were stands for training set correlation and Q^2, RMSE and Pearson r stands for test set correlation.

Table 2: Statistical score table of QSAR models

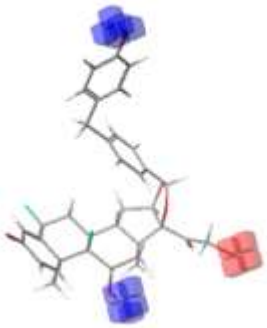
PLS factor	R^2	SD	F	P	RMSE	Q^2	Pearson - r
1	0.6478	0.3576	16.6	0.0028	0.41	0.3427	0.5963
2	0.9102	0.1915	40.5	6.50E-05	0.29	0.6724	0.9219

R^2- regression coefficient, SD- standard deviation, F- variance ratio, p- statistical significance, RMSE- root mean square error, Q^2- cross correlation coefficient, Pearson –r- correlation between the predicted and experimental activity of test set.

The contour plots for the QSAR model explaining about the structure activity is detailed below

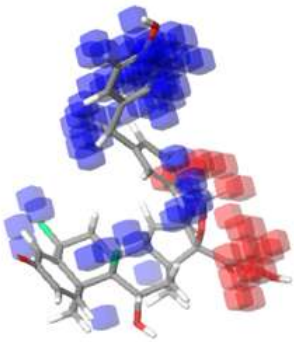
Fig 5: QSAR contour plots of active and inactive compounds, blue cubes – active contribution, red cubes – inactive contribution towards GR agonistic activity

Hydrogen bond donor effect- compound

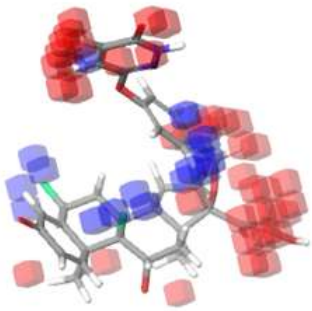


Hydrogen bond donor effect – compound- 10

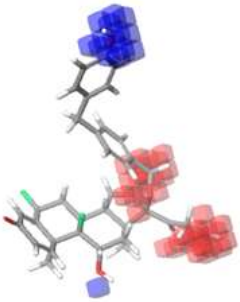
Hydrophobic effect- compound 9



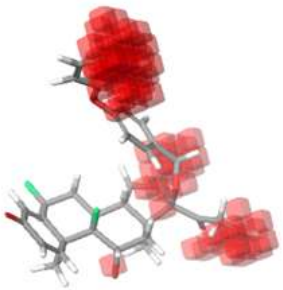
Hydrophobic effect – compound 10



Electron withdrawing – compound 9



Electron withdrawing- compound 10



The model suggested that compound 10 is not a good ligand, all the descriptor values corresponding to this compound revealed its inactive nature. Compound 9 is the active compound.

Hydrogen bond donor effect

The hydrogen bond donor effect of para hydroxyl phenyl group and hydroxyl group attached to the cyclohexyl group of steroid frame work favors the activity of the compounds and hydroxyl group attached to the hexa hydrocyclopenta dioxole group of steroid structural segment deactivates the compounds. Compound 10 is a better inactive compound by considering the hydrogen bond donor effect. The hydrogen bond donor effects of hydroxyl groups attached to the cyclohexyl ring and hydro cyclo penta dioxole group of steroid fragment and hydroxyl group attached to the pyridazine group deactivate the compound.

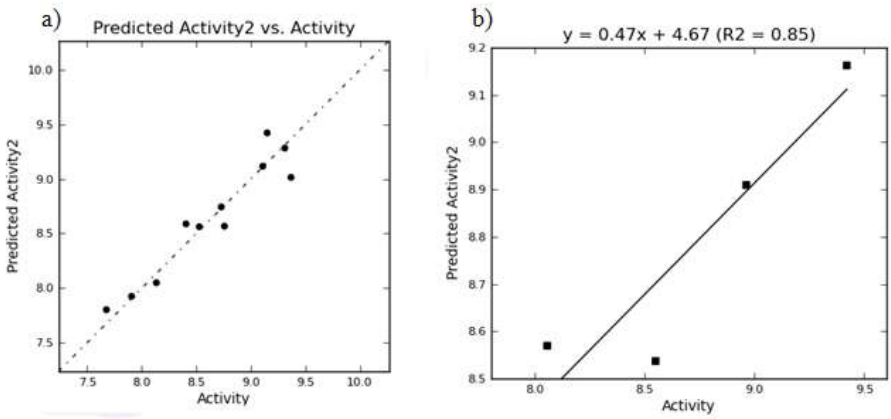
Hydrophobic effect

The hydrophobic effect of steroid ring fragment activates the compounds while considering compound 10, the steroid fragment is not much contributive towards the activity. The hydrophobic effect of 2- hydroxyl acetaldehyde group attached to the hexa hydrocyclopenta dioxole ring system is not favoring the activity of the compounds. The hydrophobic effect of phenolic ring in compound 9 is contributing towards the activity. The pyridazine ring substituent in compound 10 has generated a negative impact on the activity.

Electron withdrawing effect

The electron withdrawing effect of phenolic oxygen plays an important role for activating the compound 9 while corresponding hydroxyl oxygen at the pyridazine ring deactivates compound 10. Electron withdrawing effect of groups present in compound 10 make it more inactive.

Fig 6: Correlation plot for a) training and b) test set compounds



The QSAR model thus created introducing compound 9 as better active and compound 10 as better inactive compound. Virtual screening can be done more effectively by considering these two candidates as references.

Conclusion

Corticosteroids are the most accepted molecule used for the treatment of asthma. The 2D QSAR analysis on a set of 15 corticosteroid derivatives furnished a statistically significant model. The best three models were discussed throughout the work. All the models with regression coefficient values above 0.7 indicate goodness of fit of the models. Q^2 above 0.5 stands for external predictability of the model. The three models suggested that the descriptorsSHdsCH, Weta3.mass and Weta2.eneg are responsible for the glucocorticoid agonistic activity. All the three descriptors are making a negative impact on the activity property. These descriptors signify the importance of alkyl groups and molecular weight to the activity of these set of compounds. According to the third model we can exclude compound 10, the least active compound among, as outlier. Atom based 3D QSAR analysis has created another higher dimensional structure activity correlation model. Partial least square regression analysis on the dataset of 11 training and 4 test set compounds has generated R^2= 0.9102 and Q^2= 0.67 indicated goodness of fit and external prediction ability of the model. F and p values of the QSAR model were in the significant limits. Pharmacophore modeling has created a five feature pharmacophore, AAHHH from the activity of the derivatives. All the QSAR modeling studies

and pharrmacophore modeling studies has shed light on the effect of steroid frame work towards the glucocorticoid receptor agonistic activity of the compounds. The significance of hydrogen bond donor effect, hydrophobic effect and electron withdrawing effects were well illustrated through the 3D QSAR model analyses. Thus, the awareness and understanding about the descriptors involved in the glucocorticoid receptor agonistic activity of these compounds could provide a very good opportunity for the design of more active compounds.

Acknowledgement

University of Kerala for the University JRF, Sree Narayana College, Kollam, as research centre and Sree Narayana College , Chengannur for the computational lab facility.

Reference

[1] P. J. Barnes, "Therapeutic approaches to asthma-chronic obstructive pulmonary disease overlap syndromes," *J. Allergy Clin. Immunol.*, vol. 136, no. 3, pp. 531–545, 2015.

[2] P. J. Barnes and I. M. Adcock, "How Do Corticosteroids Work in Asthma?," *Ann. Intern. Med.*, vol. 139, no. 5 I, pp. 359–370, 2003.

[3] P. J. Barnes, S. Pedersen, and W. W. Busse, "Efficacy and safety of inhaled corticosteroids. New developments," *Am J Respir Crit Care Med*, vol. 157, no. 3 Pt 2, pp. S1-53, 1998.

[4] S. E. Abechi and E. I. Edache, "Application of Genetic Algorithm-Multiple Linear Regression (Ga-Mlr) for Prediction of Anti-Fungal Activity," vol. 7, no. 4, pp. 204–220, 2016.

[5] Y. M. Millan DS, Ballard SA, Chunn S, Dybowski JA, Fulton CK, Glossop PA, Guillabert E, Hewson CA, Jones RM, Lamb DJ, Napier CM, Payne-Cook TA, Renery ER, Selby MD, Tutt MF, "Design and synthesis of long acting inhaled corticosteroids for the treatment of asthma.," *Bioorg Med Chem Lett.*, vol. 21, no. 19, pp. 5826–5830, 2011.

[6] P. Gramatica, S. Cassani, and N. Chirico, "QSARINS-chem: Insubria datasets and new QSAR/QSPR models for environmental pollutants in QSARINS," *J. Comput. Chem.*, vol. 35, no. 13, pp. 1036–1044, 2014.

[7] H. M. Soni, P. K. Patel, M. T. Chhabria, D. N. Rana, B. M. Mahajan, and P. S. Brahmkshatriya, "2D-QSAR Study of a Series of Pyrazoline-Based Anti-Tubercular Agents Using Genetic Function Approximation," no. October, pp. 45–53, 2015.

[8] S.-Y. Yang, "Pharmacophore modeling and applications in drug discovery: challenges and recent advances.," *Drug Discov. Today*, vol. 15, no. 11–12, pp. 444–450, 2010.

[9] U. A. Shah, H. S. Deokar, S. S. Kadam, and V. M. Kulkarni, "Pharmacophore generation and atom-based 3D-QSAR of novel 2-(4-methylsulfonylphenyl)pyrimidines as COX-2 inhibitors," *Mol. Divers.*, vol. 14, no. 3, pp. 559–568, 2010.

[10] D. K. Behera, P. M. Behera, L. Acharya, and A. Dixit, "Development and validation of pharmacophore and QSAR models for influenza PB2 inhibitors," *Chem. Biol. Lett.*, vol. 4, no. x, pp. 1–8, 2016.