



Computational Studies on Human CDK9 Inhibitors

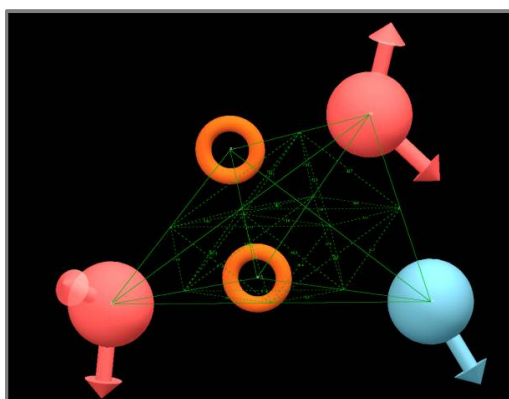
P. Revathi, S. Guru raj, S. Sree Kanth and P. Sarita Rajender*

Department of Chemistry, Nizam College [OU], Hyderabad-500 001, **INDIA**
Email: saritarajender@gmail.com

ABSTRACT

CDK9 (cyclin dependent kinase) is a protein used as a target in the treatment of cancer. CDK9 is a component of the multi-protein complex TAK/P-TEFb, an elongation factor for RNA polymerase II-directed transcription and functions by phosphorylating the C-terminal domain of the largest subunit of RNA polymerase II. It forms a complex with regulatory subunit of cyclin-T or cyclin-K. CDK9 is also known to associate with other proteins such as TRAF2, and is involved in differentiation of skeletal muscle. A 5-Point AADRR.63 pharmacophore model was developed using wogonin derivatives as CDK9 inhibitors. The generated pharmacophore model was used to derive a predictive-atom based 3D Quantitative structure activity relationship analysis (3D QSAR) model for the studied dataset. The obtained 3D-QSAR model has an excellent correlation coefficient value ($r^2=0.9332$) along with good statistical significance as shown by higher fisher ratio ($F=130.4$). The model also exhibited good predictive power confirmed by the high value of cross validated correlation coefficient ($q^2=0.6843$). Virtual screening was carried out further to identify potential CDK9 inhibitors. The QSAR model suggests the electron withdrawing character is crucial for the CDK9 inhibitory activity. In addition to the electron-withdrawing character, hydrogen bond donating groups, hydrophobic and negatively charged groups contribute to the CDK9 inhibition. These findings provide promising guidelines for designing compounds with better CDK9 inhibitory potential.

Graphical Abstract



Geometry of best pharmacophore hypothesis
AADRR.63 with (a) angles.

Keywords: Kinase, RNA polymerase II, CDK9, Wogonin derivatives.

INTRODUCTION

Cancer is one of the most prominent life threatening diseases and is a challenge to researcher's worldwide. Researchers have been studying natural products since a long time and proved them to be potent anti cancer agents [1-4]. Flavonoids are one of them; they are polyphenolic compounds which are present abundantly in plants as their pigments [5]. Flavonoids act as anti cancer agents by preventing cell proliferation [6] and progression. In the present study we tried to put forth some of the flavonoids as potent anti cancer agents by performing PHASE studies on certain wogonin derivatives [7].

Wogonin derivatives have a flavonoid nucleus and acts as anti-cancer agents by inhibiting CDK9. Cyclin-dependent kinase (CDK9) is a catalytic member of a group of enzymes which facilitate cell differentiation. It combines with cyclin T1, forms a Positive Transcription Elongation Factor b (p-TEFb) complex responsible for cell regulation, transcription, elongation and mRNA maturation. The activity of CDK9 depends on its ability to associate with cyclin-T to form the positive transcription elongation factor b(P-TEFb). As cited by QiangZhou *et al*, [8] The carboxy-terminal domain (CTD) of the largest subunit of RNA polymerase (Pol) II undergoes a cycle of phosphorylation and dephosphorylation during the transcription cycle. Shortly after transcription begins, the CTD becomes phosphorylated on Ser5 of the heptapeptide (YSPTSPS) repeats by the CDK7 kinase of the general transcription factor TFIID. This signals the polymerase to clear the promoter and shifts into an elongation mode. It also allows the CTD to recruit capping activities to the 5' end of the pre-mRNA. During the elongation stage of the transcription cycle, phosphorylation of the CTD Ser2 predominates. This enables the polymerase to resist pausing caused by negative elongation factors.

Molecular modeling is an interdisciplinary tool in the hands of a researcher to expedite the process involved in drug discovery. The extent of in silico studies on CDK9 inhibitors is increased, after understanding the function of CDK9 at the molecular level. The present work helps us to identify the pharmacophoric features of the wogonin derivatives taken from the literature using PHASE module of the Schrodinger software and build models with at least five such features. Best model is selected based on the statistics Similarity search is done against databases for the best hypothesis and hit molecules having the same features as the best hypothesis were identified.

MATERIALS AND METHODS

Data set selection and Molecule preparation: A set of fifty molecules with definite inhibitory activity against CDK9 was used for the QSAR analysis. In vitro inhibitory concentrations (IC_{50}) of the molecules were converted into corresponding pIC_{50} [$-\log(IC_{50})$] and were used as dependent variables in the QSAR calculations. PHASE module [9-14] of Schrodinger molecular modeling software was used to generate pharmacophore models. All the selected molecules were sketched in Maestro build panel. Further, geometry optimization of built molecules was carried out using the Ligprep application applying OPLS_2005 force field. All possible conformations for a molecule were generated by conformational analysis using Monte-Carlo Multiple Minimum method implemented in the Schrodinger software. The ligands were designated as active and inactive by giving suitable activity threshold value. The threshold value was fixed for active and inactive ligands. The activity threshold value was selected on the basis of dataset activity distribution and the active ligands were chosen to derive a set of suitable pharmacophores.

Creation of Pharmacophore sites and Scoring: The chemical features of all ligands were defined by six pharmacophoric features: H-bond acceptor (A), H-bond donor (D), hydrophobic group (H), negatively charged group (N), positively charged group (P), and aromatic ring (R). An active analogue approach was used to identify the common pharmacophore hypotheses (CPHs). The resulting pharmacophores were then scored and ranked to identify the best hypothesis. The scoring algorithm includes the contributions from the alignment of site points and vectors, volume overlap, selectivity,

number of ligands matched, relative conformational energy, and activity.

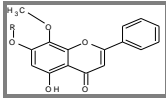
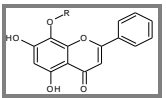
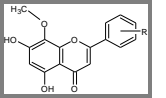
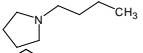
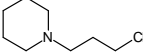
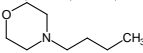
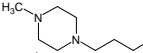
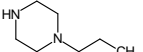
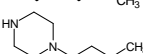
Building 3D-QSAR models: An atom-based QSAR model generation was carried out using the selected hypothesis by dividing the dataset into training and test set in a random manner in 70:30 ratio. Partial Least Square (PLS) analysis was carried to build atom-based QSAR model for the selected hypothesis using the training set molecule with a grid spacing of 1.0 Å. The best QSAR model was validated by predicting activities of the compounds. A best model (PLS factor) with good statistics was obtained for the dataset keeping the number of PLS factors in each model by maintaining PLS factor in a ratio of 1:5 of the total number of training set molecules.

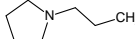
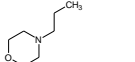
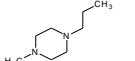
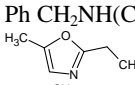
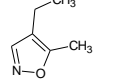
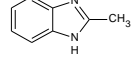
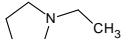
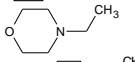
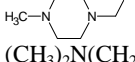
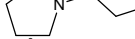
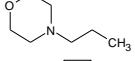
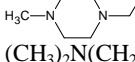
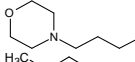
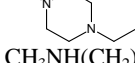
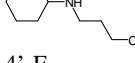
Virtual Screening based on Phase Hypothesis: Virtual screening was carried out using best hypothesis to search a 3D database of ASINEX gold-platinum library [15] used for structures that match the pharmacophoric features of the model. To achieve the best 3D similarity search, a constraint of 0.7 RMSD, 10 rotatable bond cut-off and molecular weight range of 180–500 Dalton was applied. A molecule which fits well with the pharmacophoric features of the best hypothesis was retrieved as a virtual hit molecule.

RESULTS AND DISCUSSION

The aim of this study was to elucidate the 3D structural features of Wogonin derivatives that are crucial for binding, by generating 3D pharmacophore and to quantify the structural features of CDK9 inhibitors essential for biological activity by generating atom-based 3D QSAR model. For this, a set of 50 derivatives were used in the pharmacophore modelling and QSAR studies performed in Phase module of Schrodinger suite. The structure of the molecules is provided in table 1 along with their activity values (pIC₅₀). Based on the dataset activity distribution (4.007-5.970) 11 molecules as active and 12 molecules as inactive was selected. Active (pIC₅₀>4.5) molecules were considered for the generation of pharmacophores model. To have optimum combination of sites or features common to the most active compounds, five minimum and maximum sites were set for common pharmacophore generation. Common pharmacophores models were generated with different combination of variants in which all the models were considered. Among these pharmacophores, the models which are showing better alignment with active compounds were identified by mapping to them and calculating the survival score.

Table 1. Molecule structure, Activity, PLS-3 Fitness and predicted activity data for test and training set

Molecule	R	QSAR Set	Activity	Predicted Activity	Pharm Set	Fitness
						
	1	2	3			
1a	CH ₃ (CH ₂) ₂ N(CH ₂) ₄	training	4.764	4.61		2.61
1b	(HOCH ₂ CH ₂) ₂ N(CH ₂) ₄	training	4.578	4.65		2.68
1c		training	5.442	5.29	active	2.85
1d		training	5.705	5.72	active	2.98
1e		test	5.481	5.78	active	2.99
1f		training	5.95	5.97	active	2.97
1g		training	5.039	5.07	active	2.69
1h		test	5.97	5.34	active	3
1i	Ph CH ₂ NH(CH ₂) ₄	training	4.34	4.31		2.67
1j	(CH ₃ CH ₂) ₂ N(CH ₂) ₃	training	5.477	5.33	active	2.7
1k	(HOCH ₂ CH ₂) ₂ N(CH ₂) ₃	training	5.421	5.53	active	2.63

1l		training	4.812	4.84		2.79
1m		test	5.187	5.21	active	2.82
1n		training	4.5	4.75		2.83
1o	Ph CH ₂ NH(CH ₂) ₃	training	4.09	4.14	inactive	2.65
1p		training	4.693	4.59		2.7
1q		training	4.664	4.54		2.71
1r		training	4.531	4.6		2.59
2a	(CH ₃) ₂ N(CH ₂) ₂	test	4.398	4.3		2.66
2b	(CH ₃ CH ₂) ₂ N(CH ₂) ₂	test	4.167	4.27	inactive	2.63
2c		test	4.164	4.28	inactive	2.64
2d		training	4.399	4.3		2.66
2e		training	4.247	4.28	inactive	2.65
2f	(CH ₃) ₂ N(CH ₂) ₃	test	4.379	4.29		2.66
2g	CH ₃ CH ₂) ₂ N(CH ₂) ₃	training	4.17	4.13	inactive	2.61
2h		training	4.232	4.39	inactive	2.69
2i		training	4.667	4.51		2.71
2j		test	4.37	4.59		2.69
2k	(CH ₃) ₂ N(CH ₂) ₄	test	4.447	4.49		2.72
2l		training	4.82	4.65		2.7
2m		test	4.82	4.62		2.68
2n	CH ₃ NH(CH ₂) ₂	training	4.192	4.29	inactive	2.63
2o	CH ₃ (CH ₂) ₂ NH(CH ₂) ₂	test	4.007	4.46	inactive	2.73
2p	CH ₃) ₂ CHNH(CH ₂) ₂	training	4.417	4.41		2.66
2q	CH ₃ (CH ₂) ₃ NH(CH ₂) ₂	training	4.242	4.43	inactive	2.71
2r	HO(CH ₂) ₃ NH(CH ₂) ₂	test	4.244	4.4	inactive	2.68
2s	HOCH ₂ CH(CH ₃)NH(CH ₂) ₂	training	4.361	4.17		2.64
2t	CH ₃ NH(CH ₂) ₄	test	4.013	4.68	inactive	2.69
2u	CH ₃ (CH ₂) ₂ NH(CH ₂) ₄	test	5	4.65	active	2.65
2v	(CH ₃) ₂ CHNH(CH ₂) ₄	training	4.524	4.54		2.68
2w	CH ₃ (CH ₂) ₃ NH(CH ₂) ₄	training	4.41	4.58		2.46
2x		training	4.961	4.93	active	2.64
3a	4'-F	training	4.721	4.56		2.64
3b	4'-Br	test	4.417	4.59		2.62
3c	2'-Br, 4'-F	training	4.474	4.57		2.56
3d	3'-F	test	4.226	4.3	inactive	2.54
3e	4'-OCH ₃	training	4.372	4.55		2.61
3f	4'-CH ₃	Training	4.603	4.59		2.62
3g	4'-Cl	test	4.465	4.33		2.54

The survival scoring function identifies the best candidate hypothesis from the generated models and provides an overall ranking of all the hypotheses. The scoring algorithm includes contributions from the alignment of site points and vectors, volume overlap, selectivity, number of ligands matched, relative conformational energy and activity. However, these pharmacophore models should also differentiate the active (most active) and inactive (less active) molecules. To identify the pharmacophore models with more active and less inactive features among these models, they were mapped to inactive compounds and scored. If inactive ligands score well, the hypothesis could be

invalid because it does not discriminate between active and inactive ligands. Therefore, adjusted survival score was calculated by subtracting the inactive score from survival score of these pharmacophores. After careful analysis, best pharmacophore hypothesis were selected that had maximum adjusted survival score and lowest relative conformational energy (Table 2). Based on the selected hypothesis, pharmacophore (atom)-based alignment of all the CDK9 inhibitors was performed. Atom-based QSAR models were generated using the 32-member training set and validated on 18 test set compounds for prediction of activities. A three-component (PLS factor) model with good statistics was obtained.

Table 2. Scores of different parameters of the AADRR.63 hypothesis

ID	Survival	Survival – inactive	Post-hoc	Site	Vector	Volume	Selectivity	Matches	Energy	Activity	Inactive
AADRR.63	3.792	3.792	3.792	0.99	0.997	0.803	1.414	11	0.861	5.97	4.13

Model AADRR.63 has been chosen because it produced good predictive power above other models. The special arrangement of features present in five-featured pharmacophore, AADRR.63 with distances and angles between different sites of the model AADRR.63 was shown in figure 1. The alignment generated by the best pharmacophore model AADRR.63 was used for QSAR model generation. Figure 2a presents good alignment of the active ligands and scattered alignment of inactive ligands on the developed pharmacophore model. As depicted in the figure, the two ring aromatic features, two acceptor features and one donor features are mapped with all 11 active inhibitors. Alignments of in actives are shown in figure 2b.

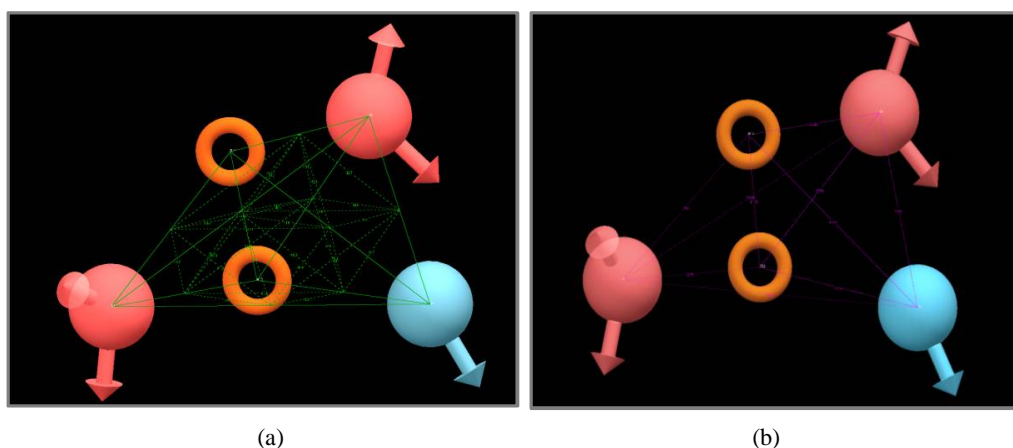


Figure 1. Geometry of best pharmacophore hypothesis AADRR.63 with (a) angles and (b) distances.

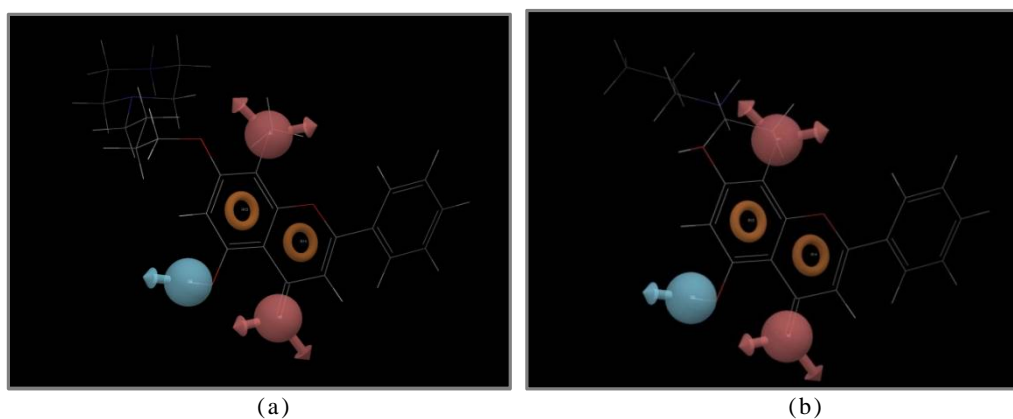


Figure 2. Alignment of (a) Best active molecule and (b) inactive molecules on hypothesis AADRR.63.

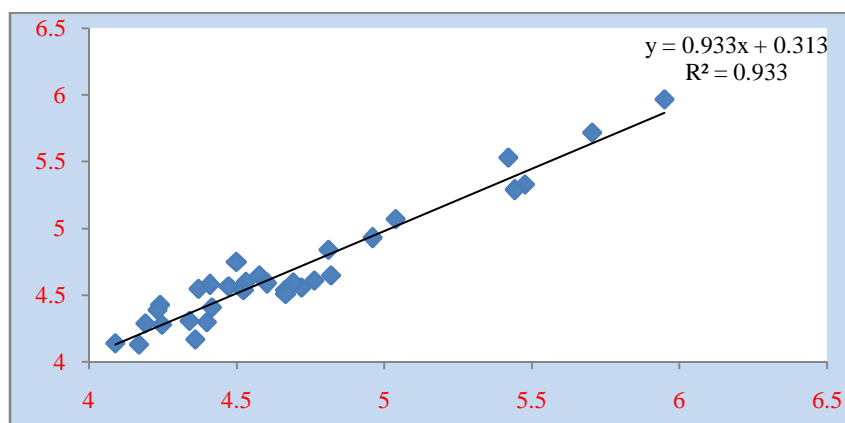
From figures (2 a and b), we can easily identify that active ligand is having good alignment than inactive one. A three-PLS factor model with good statistics and predictive ability was generated for the dataset (Table 3).

Table 3. PLS statistical parameters of selected 3D-QSAR model

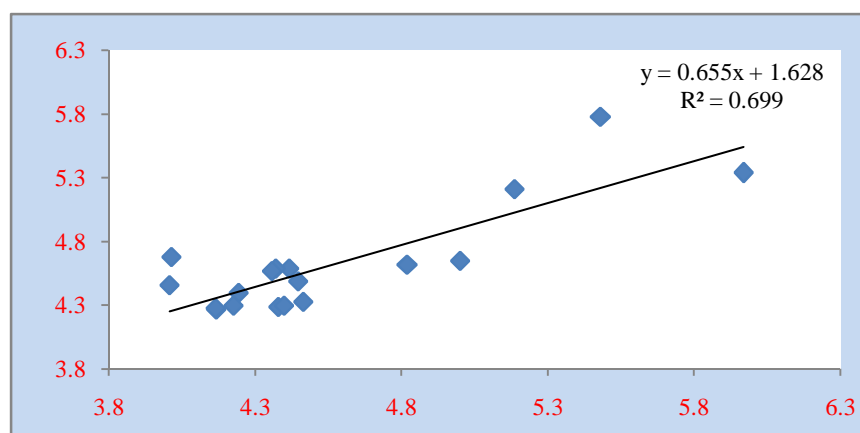
ID	Factors	SD	R ²	F	P	Stability	RMSE	Q ²	Pearson-R
AADRR.63	1	0.3025	0.5951	44.1	2.36E-07	0.611	0.2993	0.6639	0.8493
	2	0.1822	0.858	87.6	5.12e-013	0.1428	0.2864	0.6922	0.836
	3	0.1272	0.9332	130.4	1.48e-016	0.0948	0.29	0.6843	0.8365

(SD, standard deviation of the regression; R, squared value of R² for the regression; F, variance ratio. Large values of F indicate a more statistically significant regression, P, significance level of variance ratio. Smaller values indicate a greater degree of confidence; RMSE, root-mean-square error, Q, squared value of Q² for the predicted activities, Pearson-R, Pearson R value for the correlation between the predicted and observed activity for the test set.)

The statistical significance and predictivity was observed for an incremental increase in PLS factor up to three. The F value (130.4) indicates a statistically significant regression model, which is also supported by the small value of the variance ratio (P), an indication of a high degree of confidence. Further, small values of standard deviation (0.1272) of the regression and RMSE (0.29) make an obvious implication that the data used for model generation are best for the QSAR analysis. Validity of the model is authenticated by cross-validated correlation coefficient ($q^2 = 0.6843$) that was obtained for the test set and was not included in the model generation. The scatter plot of experimental and PHASE predicted activity of the Wogonin derivatives is shown in figure 3.



(a)



(b)

Figure 3. Scatter plot of experimental activity (on X axis) vs PHASE predicted activity (on Y axis) of [a] Training set [b] Test.

The QSAR model displays 3D characteristics (Figure-4a, b and c) as cubes that represent the model and color according to the sign of their coefficient values, which by default is blue for positive coefficients and red for negative coefficients. Positive coefficients indicate increase inactivity, negative coefficients a decrease. The visualization of the coefficients is useful to identify characteristics features of the molecule structures that tend to increase or decrease the activity. This might give a clue to what functional groups are enviable or adverse at certain positions in a molecule. The blue cubes in 3D plots of the 3D pharmacophore regions refer to position at which the specific feature in a molecule is important for better activity, whereas the red cubes demonstrates that particular structural feature or functional group, which is not essential for the activity or likely the reason for decreased binding potency. Visual analysis of above figure 4a demonstrates that the presence of the blue cubes at alkyl chain is pointing out the positive potential of electron withdrawing characteristic of the molecules at this particular place and is requisite for the activity. It can be suggested that addition of appropriate electron withdrawing groups at this region will enhance the CDK9 inhibition. Figure 4b illustrates that H-bond donor characteristic is necessary at D2 site. The red cubes at the alkyl chain demonstrate a negative potential of H-bond donor at that position on the compounds. Figure-3c demonstrates the effect of hydrophobic groups on CDK9 inhibition. It can be deduced from the figure that hydrophobic group's are well tolerated near alkyl chain (blue cubes).

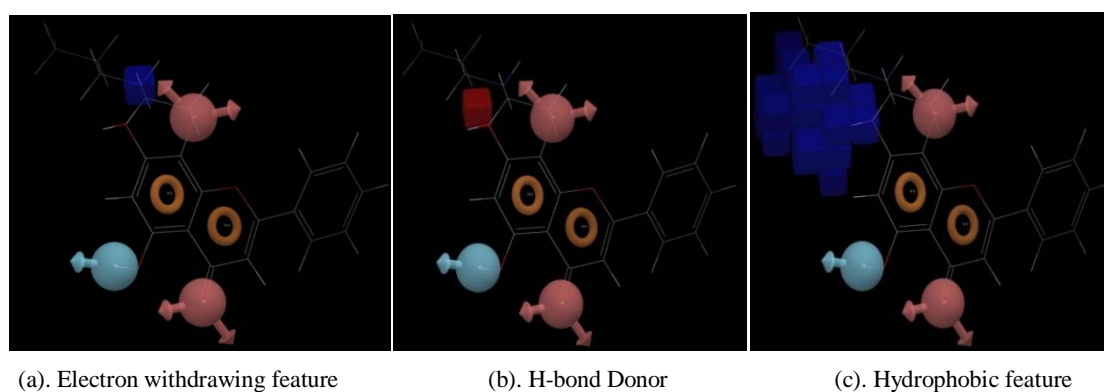


Figure 4. QSAR visualization of various substituent effects.

APPLICATION

3D QSAR Model-Virtual Screening: The pharmacophore hypothesis AADRR.63 was used as a 3D structural query for retrieving potent molecules from the in house database. From that 1000 molecules were retrieved as a hit which were further screened based on their drug like property. The most promising hit molecule structure is provided in figure 5. The hit molecules can be further screened by molecular docking studies, where in molecular interaction between the hit molecule and CDK9 protein can be analysed.

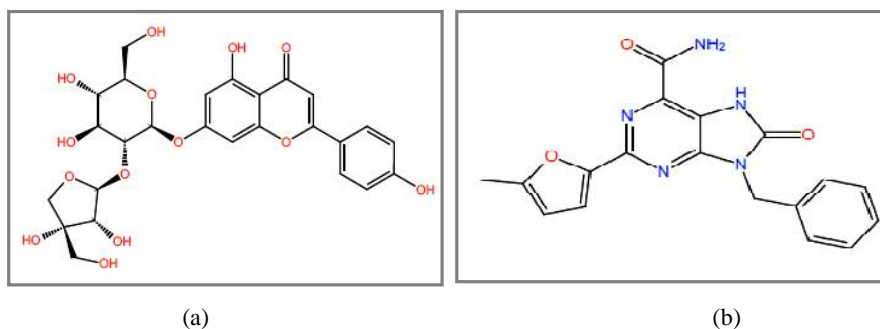


Figure 5. New scaffolds retrieved as CDK9 inhibitors based on 3D-QSAR similarity feature search (PHASE hypothesis based virtual screening).

CONCLUSION

A ligand-based pharmacophore model was generated for the series of Wogonin with CDK9 inhibitory activity to reveal the structural features responsible for biological activity. With the help of pharmacophore-based alignment of CDK9 inhibitors, a meaningful 3D-QSAR was derived to identify the three-dimensional arrangements of various substituents and their impact on CDK9 inhibition. The selected model as shown by the correlation statistics and predictive statistics is very much significant to draw unambiguous inferences. Further, the generated 3D-QSAR model also explain show and at what extent electron withdrawing, hydrophobic and H-Donor moieties in the molecular structure influences CDK9 inhibition showed by the title compounds. The model shows that CDK9 inhibition can be increased, if the electron-withdrawing character at alkyl side chain is supplemented by appropriate functional groups and by the incorporation of H bond acceptor and hydrophobic groups at specific positions in the molecules.

Finally, six potential hits with good fitness value and predicted activity were identified by virtual screening and docking, whose activity can be further improved with the help of this QSAR model. This study provides a set of guidelines which will greatly help in designing the newer and more potent CDK9 inhibitors of the Wogonin based scaffold.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support provided by Department of Chemistry, Nizam College, OU, for carrying out this research project.

REFERENCES

- [1]. P. Wei, X. Zhang, S. Tu, S. Yan, H. Ying, P. Ouyang. New potential inhibitors of DNA topoisomerase. Part II: Design and synthesis of α -lapachone derivatives under microwave irradiation, *Bioorg Med Chem Lett*, **2009**, 19, 828–830.
- [2]. J. Wiemann, L. Heller, R. Csuk., Targeting cancer cells with oleanolic and ursolic acid derived hydroxamates, *Bioorg Med Chem Lett.*, **2016**, 26, 907–909.
- [3]. Wellington K., Understanding cancer and the anticancer activities of naphthoquinones – a review, *RSC Adv.*, **2015**, 5, 20309–20338.
- [4]. Singh M, Kaur M, Silakari O., Flavones: an important scaffold for medicinal chemistry, *Eur. J. Med. Chem.*, **2014**, 84, 206–239.
- [5]. Diane F. Birt, Suzanne Hendrich, Weiqun Wang, Dietary agents in cancer prevention: flavonoids and isoflavonoids, *Pharmacol Ther.*, **2001**, 90(2-3), 157-77.
- [6]. Theodore Fotsis, Michael S. Pepper, Erkan Aktas, Stephen Breit, Sirpa Rasku, Herman Adlercreutz, Kristiina Wähälä, Roberto Montesano and Lothar Schweigerer, Flavonoids, Dietary-derived Inhibitors of Cell Proliferation and in Vitro Angiogenesis, *J. Cancer Res.*, **1997**, 57, 2916-2921.
- [7]. Jinlei Bian, Tinghan Li, Tianwei Weng, Jubo Wang, Yu Chen, Zhiyu Li, Synthesis, evaluation and quantitative structure–activity relationship (QSAR) analysis of Wogonin derivatives as cytotoxic agents, *Bioorg Med Chem Lett.*, **2017**, 27(4), 1012-1016.
- [8]. J. H. Yik, R. Chen, R. Nishimura, J. L. Jennings, A. J. Link, Q. Zhou, Inhibition of P-TEFb (CDK9/Cyclin T) Kinase and RNA Polymerase II Transcription by the Coordinated Actions of HEXIM1 and 7SK snRNA, *Mol Cell.*, **2003**, 12(4), 971-82.
- [9]. Schrödinger, LLC, New York, NY, PHASE, Version 3.0, **2010**.
- [10]. S. L. Dixon, A. M. Smondyrev, E. H. Knoll, S. N. Rao, D. E. Shaw, R. A. Friesner, PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results, *J. Comput. Aided Mol. Des.*, **2006**, 20, 647–671.

- [11]. S. L. Dixon, A. M. Smondyrev, S. N. Rao, , PHASE: A Novel Approach to Pharmacophore Modeling and 3D Database Searching, *Chem. Biol. Drug Des.*, **2006**, 67, 370-372
- [12]. W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives , Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids, *J. Am. Chem. Soc.*, **1996**, 118, 11225-11236.
- [13]. H. Dureja, V. Kumar, S. Gupta, A. K. Madan, Topochemical models for the prediction of lipophilicity of 1,3-disubstituted propan-2-one analogs, *J. Theor. Comput. Chem.*, **2007**, 6(3), 435-448.
- [14]. W. L. Jorgensen, J. Tirado-Rives, The OPLS [optimized potentials for liquid simulations potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin, *J. Am. Chem. Soc.*, **1988**, 110(6), 1657-1666.
- [15]. <http://www.asinex.com/libraries-html/>.